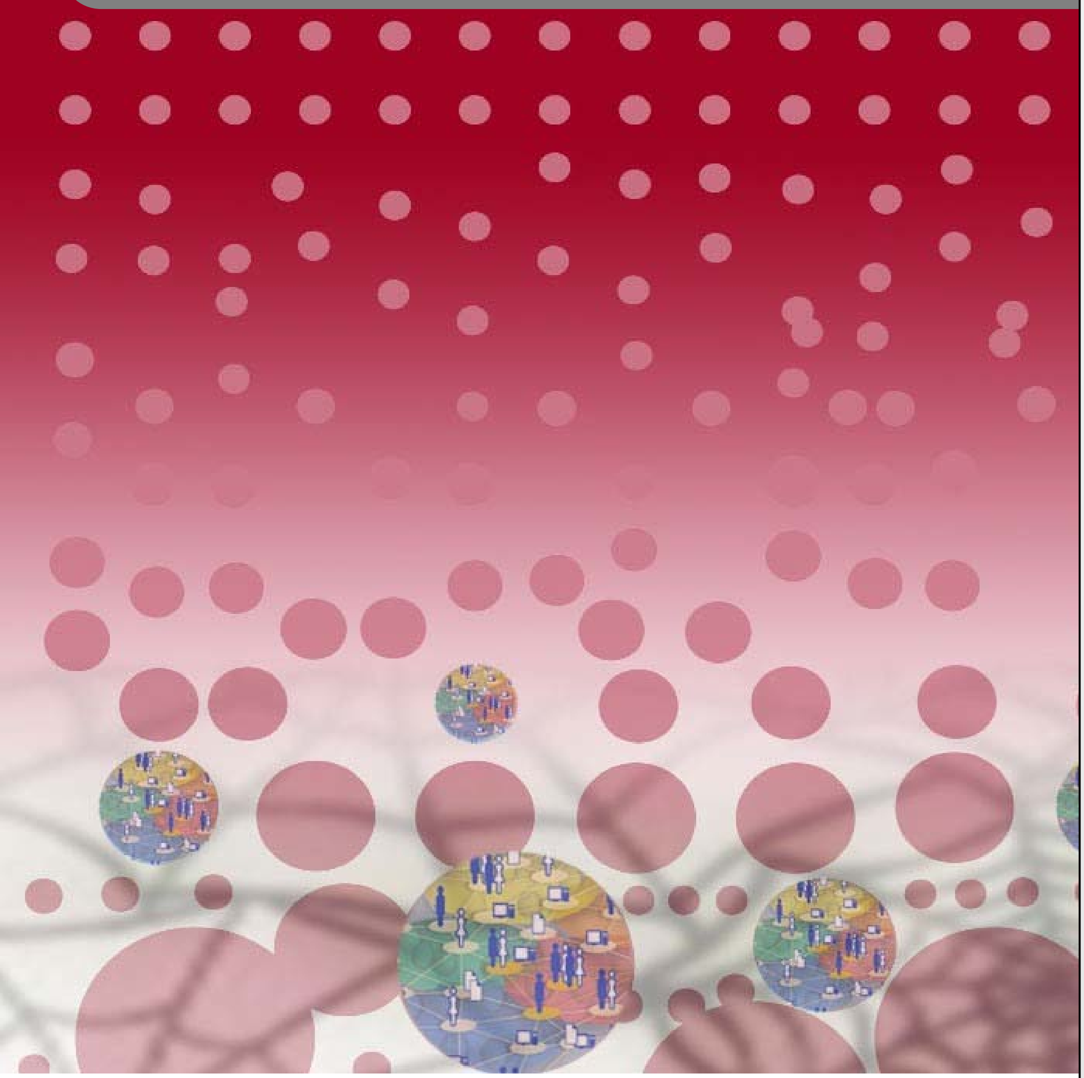
Human centered processes

Patrick Brézillon, Gilles Coppin, Philippe Lenca

HCP-2008



Human Centered Processes

Third International Conference

Proceedings – Part I

Main Conference

Editors:
Brézillon Patrick,
Coppin Gilles,
Lenca Philippe

June 8-12, 2008
Delft
The Netherlands

HCP-2008 Proceedings

Part I

Main Conference

Third International Conference on

Human Centered Processes

June 8-12, 2008

Delft, The Netherlands

Editors

Brézillon Patrick, Université de Paris 6, France
Coppin Gilles, Institut TELECOM/TELECOM Bretagne, France
Lenca Philippe, Institut TELECOM/TELECOM Bretagne, France

Managing Editor

Uguen Gurvan, Institut TELECOM/TELECOM Bretagne, France

TELECOM Bretagne ISBN: 978-2-908849-22-6

Preface

The papers met in this volume correspond to the communications done during the Third International Conference on Human Centered Processes (HCP-2008), which held in Delft (The Netherlands) during June 2008. HCP-2008 is in line with the series of Human Centered Processes conferences held previously at Brest (France, 1999) and in Luxembourg (2003) under the umbrella of the EURO Working Group Human Centered Processes (http://recherche.telecom-bretagne.eu/hcp/). HCP-2008 is a platform with a plenary interdisciplinary conference and workshops in order to foster a shared interest across disciplines and stimulate within-discipline progress and information exchange.

Today the human world and the digital information world form a complex changing environment. Any combination of human actors and software systems must behave like a process able to maintain its integrity, efficiency and adaptability in such dynamical environments. Humans engage in interaction and cooperation in many different contexts that often coincide in the same space-time interval. We need a better understanding and implementation of effective solutions to support humans in their adaptive, interactive and cooperative processes. The dynamics of cognitive processes and cross-learning are of much interest, especially when decision makers are associated with decision support systems: the advent of a paradigm for human-machine sustainability.

The first HCP conference was mainly devoted to the application of cognitive approaches in various fields of process control and other complex industrial problems usually managed by Operational researchers. The second HCP conference was dedicated to distributed decision and man-machine cooperation. Both of these conferences put special emphasis upon cognitive models of decision making. The third conference (2008) focuses on the human actor and software agent collaboration in safety and time critical systems-of-systems. This initial focus included topics ranging from: architectural requirements for such systems, cognition-based approaches for awareness and decision support, context awareness and the effects on learning, reasoning and decision making in both humans and software agents, the interaction and socially inspired process models for such collaborating systems and techniques to support them, the design, build and testing of such systems, what design patterns are adequate, what methods and tools to use.

The papers carefully selected for the main conference cover most of these themes. Three keynotes speakers, namely J.J. Meyer (Universiteit Utrecht, The Netherlands), R.R. Hoffman (IHMC, USA) and D.L. Olson (University of Nebraska-Lincoln, USA) have accepted to share their experience with participants in the context of the HCP-2008 conference. Three workshops completed in a nice way the general presentations done at the main conference.

The first workshop organized by Anders Kofod-Petersen, Jörg Cassens, David B. Leake and Marielba Zacarias) proposed to attack the HCP matter from the modeling and reasoning in context. It was the fourth incarnations of a series of workshops held jointly with different conferences of high quality. Where traditional software applications “know” by design in which situations they are to function, applications in pervasive computing and ambient intelligence do not necessarily have this luxury. Due to the very nature of the dynamism in the world with which these systems interact, they have to dynamically adapt their behavior in run time. To do this, they must be able to somehow interpret the environment in which they are situated. This ability is often referred to as being context aware, or even situation aware. Being aware of the environment facilitates the ability to adapt behavior by being context sensitive.

The second workshop considered the ethical aspects of HCP in decision making. It was organized by Fred Wenstop (Chairman of the EURO Working Group ETHICS) and Jean-Pierre Brans (Initiator of this EURO Working and Past President of EURO). Many themes were discussed: The contributions focussed on methodological aspects (System Thinking, Multi-Actors Systems and Ethics, MCDA, Control, Ethical principles and Ethical solutions, Sustainable Development, Environment, Society models, Large Scale Social Systems, Peace Education, and Power of Mathematics in Decision Processes) and on applications and cases studies (Fossil Fuel and CO2 emissions, Ethical Approaches for Health care and Disabled persons, Hospitals, Government Negotiations).

The third workshop was focused on the problem of supervisory control and the question of relationship between automation and human operators. It was organized by Jill Drury (MITRE) and Yves Boussemart / Carl Nehme (MIT/HAL). Problems such as the management of adaptive interaction or human operators supervising strategies were discussed. The first part of the workshop consisted in short papers presentation, while a second part was devoted to an interactive discussion that led to a workshop synthesis reporting about the most important and challenging issues in the domain, main recommendations from the work group and future interesting research directions.

A slot of time was left at the main conference for the workshop organizers for summing up the main results of their discussions in order to share with others. Otherwise the main conference has been organized in the following way, and these Proceedings kept the same paper organization. The papers have been divided in seven groups in order to exhibit their cohesion. Thus, the sessions of the conference and the chapters of these Proceedings are organized in the following way: (1) Information Systems and HCP, (2) Data mining for HCP, (3) HCP-based applications, (4) The human side of decision, (5) Knowledge in HCP, (6) Methodological aspects of HCP, and (7) Societal aspects of HCP.

We sincerely hope that you will enjoy the reading of these Proceedings and would like participate to the next HCP Conference!

Patrick Brézillon, Conference Chair
Gilles Coppin, Steering Committee Chair
Philippe Lenca, Workshop General Chair

HCP-2008 platform governance structure

Steering committee

Coppin Gilles, Institut TELECOM/TELECOM Bretagne, France Chair of the Committee

Bisdorff Raymond, University of Luxembourg, Luxembourg
Brans Jean-Pierre, University of Brussels, Belgium
Lenca Philippe, Institut TELECOM/TELECOM Bretagne, France
Krarup Jakob, University of Copenhagen, Denmark

Program Committee

Brézillon Patrick, Université de Paris 6, France Chair of the Committee

Adam Frédéric University College Cork, IR
Bisdorff Raymond, University of Luxembourg, LX
Borges Marcos, UFRJ-NCE, BR
Brans Jean-Pierre, VUB, BE
Brézillon Juliette, Université de Paris 6, FR
Burstein Frada, Monash U., AU
Crevits Igor U. Valenciennes ,FR
Cummings Missy, MIT, USA
Dagienne Valentina, Institute of Mathematics, LT
Giboin Alain, INRIA, FR
Gonzalez Avelino, U. Central Florida, US
Greco Alberto, U. of Genova, IT
Grisvard Olivier, Thalès, FR
Guerrero Luis, U. of Chili, CL
Guillet Fabrice, Université de Nantes, FR
Hoffman Robert, IHMC, US
Humphreys Patrick, London School of Econ & Poli Sci, UK
Kouadri Mostéfaoui, Ghita, U. of Oxford, UK
Lenca Philippe, Institut TELECOM/TELECOM Bretagne, FR
Linger Henry, Monash U., AU
Médini Lionel, Université de Lyon I, FR
Multisilta Jari, Tampere U. of Technology, FI
Nieuwenhuis Kees C. H. M., D-CIS Lab / Thales Research, NL
Pino José, University.of Chili, CL
Pomerol Jean-Charles, U. Paris 6, FR
Poulet François, ESIEA, FR
Roth-Berghofer Thomas U. Kaiserslautern, DE
Salgado Ana Carolina, UFPE, BR
Schultz Stefan, Carleton U., CA
Slowinski Roman, Poznan U. of Technology, PL
Tijus Charles, Université de Paris 8, FR
Turner Roy, U. Maine, US
Wijngaards, Niek, D-CIS Lab / Thales Research, NL
Zaslavsky Arkady, Monash U., AU

Workshop Committee

Lenca Philippe, Institut TELECOM/TELECOM Bretagne, France	Chair of the Committee
Boussemart Yves, MIT/HAL, USA	WS SCCSM
Brans Jean-Pierre, VUB, Belgique	WS OR and Ethics
Cassens Jörg, NTNU, Norway	WS MRC
Drury Jill, MITRE, USA	WS SCCSM
Kofod-Petersen Anders, NTNU, Norway	WS MRC
Leake David, Indiana University, USA	WS MRC
Nehme Carl, MIT/HAL, USA	WS SCCSM
Wenstøp Fred, BI,Norway	WS OR and Ethics
Zacarias Marielba, Algarve University, Portugal	WS MRC

Organization Committee

Nieuwenhuis Kees C. H. M., DECIS Lab, The Netherlands	Local organization chair
Wijngaards Niek J. E., DECIS Lab, The Netherlands	Local organization chair
Charfi Sofien, INT & LIP6, France	Web site Chair

HCP-2008 Sponsorships

D-CIS Lab

Université Pierre & Marie Curie (UPMC)

Institut TELECOM

Laboratoire d'informatique de Paris 6 (LIP6)

TELECOM Bretagne

Association Française pour le Contexte (AFC)

Association Extraction et Gestion des Connaissances (EGC)

Table of Contents

“Information Systems and HCP” session

“Data Mining for HCP” session

“HCP-based Applications” session

“The Human Side of Decision” session

“Knowledge in HCP” session

“Methodological Aspects of HCP” session

“Societal Aspects of HCP” session

“Information Systems and HCP” session

11:00 – 12:30 am

Tuesday, June 10, 2008

Situated cognitive engineering: Developing adaptive track handling support for naval command and control centers

M.A. Neerincx[1,2], G.M. te Brake[1], J.G.M. van de Ven[1], H.F.R. Arciszewski[1], T.E. de Greef[1,2], and J. Lindenberg[1]

[1]*TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands <mark.neerincx><guido.tebrake>@tno.nl*

[2]*Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*

ABSTRACT. *Future naval missions set high operational, human factors, and technical demands for a system that enhances human-machine teams' capabilities to cope with complex and potentially hazardous situations. This paper presents a situated Cognitive Engineering (sCE) methodology for an integrated analysis of these three types of demands, the derivation and maintenance of a requirements baseline with its rationale, and the refinement and validation processes of the requirements (including human-in-the-loop evaluations of simulation-based prototypes). Application of this method provided a coherent and concise compilation of design knowledge for an Adaptive Automation (AA) module for track identification: a theoretical and empirical founded Requirements Baseline with its design rationale consisting of core functions, claims, scenarios and use cases. The core functions focused on coping with workload dynamics and maintaining situation awareness; for each core function we specified claims on the operational effects and requirements for the system, which were added to the use cases specifications. The evaluation showed that the specified AA-module helps to reduce workload and maintain adequate situation awareness during critical naval missions.*

1. Introduction

Naval missions are changing substantially, setting new demands for advanced cognitive support systems. There is a general increase in operation diversity, and in complexity and amount of information flows that have to be processed, while team sizes are decreasing. The smaller teams must have the capabilities to cope with ambiguous and dynamic situations under high time pressure (e.g., for surveillance of littoral regions). In such situations, the team members have to collaborate optimally within their team and with ‘external’ partners who may come from other defense units and countries. To cope with all these changing demands, the Royal Netherlands Navy (RNlN) aims at flexible crews and systems that adapt to dynamic demands. The research program ‘Human-System-Task Integration’ (HSTI) identified and developed a broad set of design knowledge—consisting of theories, concepts, models and methods—to support the (re)design of crews, work and systems.

This paper presents a situated cognitive engineering method that helps to systematically acquire, refine and validate such design knowledge for the derivation of a theoretical and empirical founded Requirements Baseline. This method has originally been developed for the space domain, and was recently used to establish a Requirements Baseline for a system that supports distributed human-machine teams during planetary space missions (Neerincx et al., 2006). The method seems transferable to a lot of other domains as well. The naval domain is a case in point, because of the correspondences between space and defense missions. Both take place in hostile and complex environments where human operations and safety are highly dependent on adequate technology performance. Our situated cognitive engineering method fits in the trend of simulation-based acquisition for the development of complex and critical systems (US Army, 2007). Especially in the military domain, use of simulation technology and modeling across integrated acquisition phases and programs is considered important to save costs and develop products that meet users’ expectations.

2. Situated Cognitive Engineering

Technological developments are causing a fundamental change of the machine’s role in complex work environments, e.g. in the defense and space domain. Machines are becoming part of cognitive systems that consist of human and synthetic actors who collaborate for successful attainment of their joint operation objectives (e.g., Hoc, 2001). Neerincx & Lindenberg (2008) developed a situated Cognitive Engineering (sCE) method for the design of such Human-Machine Collaboration (HMC). It follows an iterative human-centered development process corresponding to recent human-factors engineering methods and standards (e.g. ISO 13407 “Human-centered design processes for interactive systems”), aiming at an incremental development of advanced technology. Corresponding to the ‘classical’ CE methods (Hollnagel and Woods, 1983; Norman, 1986; Rasmussen, 1986), it consists of an iterative process of generation, evaluation and refinement. In addition, the sCE method combines operational, human factors and technological analyses to establish a sound and practical design rationale, and applies simulation-based acquisition and assessment techniques for refinement and validation. Figure 1 shows the general structure that consists of three components: the Work Domain and Support (WDS) analysis, the generation and maintenance of the Requirements Baseline (RB) with its rationale, and the review and refinement activities to improve and validate this Requirements Baseline. In general, the cognitive engineering activities start with a—possibly preliminary—WDS analysis, followed by a first RB specification and, subsequently, a series of refinement processes. It should be noted that during this process, new insights in the operational demands, human factors or technology can be acquired and used to further improve the RB. Furthermore, the methodology allows for an incremental development approach, in which successive prototypes include more functions.

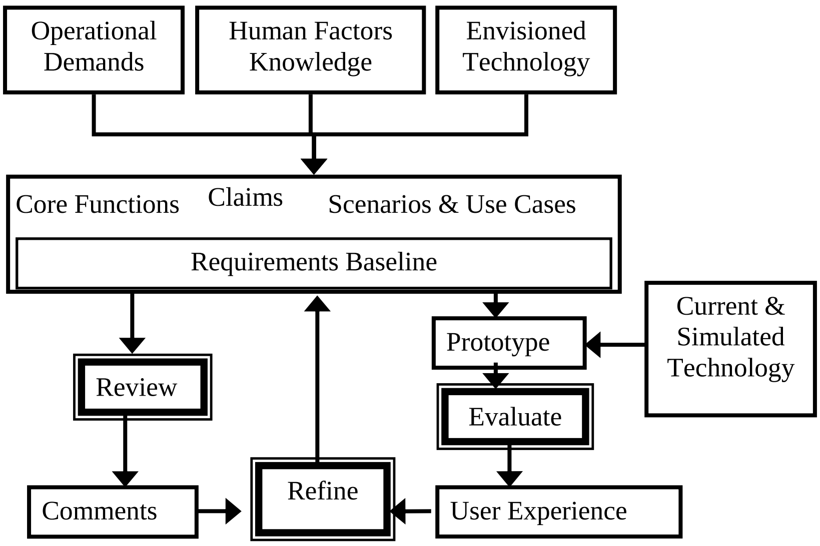


Figure 1: *The iterative process of requirements analysis*.

First, the WDS analyses identify operational, human factors and technological demands that set specific objectives or constraints for the support of future naval missions. Current and future *operational demands* can be acquired via scenario analysis (Rosson & Carroll, 2001; McCrickard & Chewar, 2006) and cognitive work analysis (Vicente, 1999; Naikar and Sanderson, 2001). With respect to *human factors knowledge*, the methodology advocates to systematically address state-of-the-art practical theories and models, which include accepted features of human cognitive and affective processes and which can be supported by technological artifacts (cf. Neerincx, 2003). This paper will provide an example of such a theory on situation awareness (Endsley, 2000). The analysis of *envisioned technology* is included to set a feasible design space, and to systematically address the adaptive nature of both human and synthetic actors with their reciprocal dependencies (Maanen et al., 2005).

The second component of the sCE methodology consists of the construction and maintenance of the Requirements Baseline (a table with all requirements) and the general design rationale that consists of the core functions, claims, and scenarios & use cases. For the specification of this rationale and the actual requirements baseline, we distinguish three steps that should be followed both from top and bottom (i.e., a top-down, goal-directed approach to work out core functions and a bottom-up, event-driven approach to address contextual or situational demands in scenarios (Neerincx, 2003):

- The *core functions* of the system are derived from the WDS analysis. For example, to better cope with workload dynamics during track identification in the Command and Information Center, an adaptive track-handling support system can be proposed which brings about an adequate operator's situation awareness. Corresponding core functions can be the prevention of mode errors or automation surprises.
- For each core function, one or more testable *claims* on its operational effects have to be specified; such a claim can be assessed unambiguously in review or evaluation processes. Both positive and negative claims can be specified. An example of the proposed system is that the operators will show faster adequate responses to all critical tracks in high workload conditions. Furthermore, for each core function, one or more requirements have to be specified for the future system (i.e., what the system must do). The complete set of requirements comprises the *requirements baseline*.
- *Scenarios & use cases*. Scenarios are coherent and situated stories about how a specific team of operators and the system behave or will behave in specific circumstances with the operational consequences, e.g. for track identification in littoral waters. Scenarios are often visualized via annotated storyboards to present and discuss the design rationale, and used for the evaluation of prototypes. Use cases describe the general behavioral requirements for software systems or business processes, and have a specific specification format. According to our methodology, each use case should explicitly refer both to one or more requirements and to one or more claims. In addition, each claim and each requirement should be included in one or more use cases. A scenario can be viewed as an instance of one or more use cases. Use cases and scenarios are very useful when discussing a not-yet-existing system with different stakeholders. With minor help most people are able to understand these design specifications.

Third, the sCE methodology distinguishes two types of refinement and validation processes. Domain, human factors and technical experts have to *review* the requirements baseline and its rationale, checking for consistency and validity. Furthermore, simulation-based prototypes can be built for human-in-the-loop *evaluations* of the claims. Both the environmental conditions and the system functions can be simulated in such evaluations, partly controlled by a human operator (e.g., to trigger alarms and change the level of automation according to a fixed set of rules). Review and evaluation results should be used to confirm, refine or remove design specifications.

The HSTI program applied the sCE methodology to the development of an Adaptive Automation (AA) module for track identification in order to establish a theoretically sound and empirically proven Requirements Baseline, and to realize a good transfer of research activities to the future system development and software engineering activities. It is important to note that this study focused on a future naval system that is in an early development stage, in which human factors aspects, operational demands and technology are systematically explored. According to current system development approaches, requirements are assessed and refined from rather high-level specifications in early development stages to detailed definitions in late development stages. Consequently, the present requirements baseline contains relatively abstract specifications that will be assessed and refined further.

3. Work Domain and Support Analysis

3.1. *Operational demands*

Missions for navies around the world are becoming increasingly complex, more-and-more taking place in littoral waters with involvement of several parties that may have unclear intentions and loyalties (see Joint Vision 2010 and 2020). Furthermore, reductions in crew sizes may cause reduced resilience to high workload situations (e.g., due to diminished opportunities for task re-planning). To deal with these trends, the amount of automated systems on board of naval ships is increasing, especially for command and control tasks, the development of situation awareness (SA), and weapon deployment. In future naval command and control, humans and systems will need to collaborate in an effective manner to obtain optimal performance. Problematic of automation in these environments is the large variability in workload. For prolonged periods, the amount of work may be rather small, but there will also be intense periods with high workload and stress. Overload might appear due to a competition for the operator's attention that is going on between different information items. If many tasks are handled by automated systems, the operator can deal with high workload circumstances, but will suffer from severe underload during quiet periods, probably losing his or her situational awareness.

3.2. Human Factors knowledge

In addition to the operations analyses, a literature study was conducted on Human Factors of complex high-demanding task environments, in which well-trained human operators may act in extreme and hostile situations (such as the defense and safety domains). This study provided key issues that the system should address to cope with the workload dynamics:

- *Cognitive task load* (Neerincx, 2003). Due to the large variety over time in number and complexity of tasks, the momentary mental load of the operators can be suboptimal. The system should support an adequate load scheduling over time and available human-machine resources based on a model of cognitive task load that distinguishes three load factors: percentage time occupied, number of task switches and task complexity.
- *Situation awareness* is most relevant in naval command and control environments (Endsley, 2000). A large part of sensors, automation, and operators, are used to build a common operational picture showing all relevant tracks, and providing the officers with the information required to make sense of the situation (Weick, 1995).
- *Trust* (Parasuraman, 1997; Adams et al., 2003). If users rely too much or too little on human or technology, performance will be suboptimal. Appropriate

trust depends on understanding of capabilities of the system, colleagues, and oneself. Users are not very good at estimating how much to trust a machine.

- *Diversity of cognitive capacities* (Scerbo, 2001). The human operators have different expertise and experiences, and will perform their various tasks in different environments. This causes differences in the momentary capabilities, levels of attention and available modalities. The system needs to be aware of these factors to be able to help the operator effectively by tailoring the level of automation to the available attention resources and modalities.
- *Decision making* (Klein, 1998). Rational decision making ought to be supported by exhaustive evaluation of options, collecting and providing an overview, ranking the options, and possibly proposing the best. Naturalistic decision-making is supported by functions that assess the situation based on patterns capturing experience and preference of the crew, recommend actions based on the patterns, check that the execution of the course of actions is according to expectancies, and test assumption underlying human naturalistic decision-making.
- *Collaboration* (Mohammed and Dumville, 2001). Adaptive teams help to avoid gaps and overlap in individuals' assigned work (i.e. support coordination), to obtain mutual benefits of human and machine actors by sharing or partitioning work (i.e. support cooperation) and to achieve collective results that the participants would be incapable of working alone. Furthermore, it helps to support the generation and maintenance of a shared mental model within human-machine teams, which contains both team knowledge as well as situation knowledge. By mediating between actors, insight will be provided into the other actors' goals, intentions behavior and needs.

3.3 Envisioned technology

In addition to the operations and human-factors analysis, we conducted a technology assessment. On naval ships, the combat management system (CMS) assists the crew in its assessment of the tactical situation, its decision making with respect to the evolving situation, and the execution of its subsequent actions. Four types of technology development will influence future CMSs.

- Current cross-disciplinary Research and Development (R&D) communities bring forth enabling technology that can be integrated, and possibly embedded, in such environments in order to sense, interpret and anticipate individual human conditions and behaviors (e.g. to improve safety and health) around themes like affective computing and augmented cognition (e.g., Picard, 1997; Satyanarayanan, 2001; Schmorrow et al., 2005).

- Context-sensing systems are being developed that employ diverse smart sensors to assess the momentary context of operation.
- Developments in Artificial Intelligence, such as Multi-Agent Systems, bring forth real or virtual machines that can act autonomously in dynamic environments and take the initiative in joint human-machine operations (e.g. Clancey, 2004). The demand for unmanned vehicles (aerial, land, above, under water) is rising sharply. It's a response to the increasing danger from highly mobile threats, and the need to rapidly engage high-value targets of opportunity, especially deep in unfriendly territory and dangerous environments.
- Future naval operations will be network-centric. Sensors and shooters are connected in a big network, enabling ships and planes to use equipment on other platforms. This capability will come with a sharp increase of the amount of available information and planning complexity.

Future combat management systems and sensors will contain many intelligent algorithms for correlating, fusing, and interpreting data. Many low-level tasks related to correlation and sensor fusion will be handled automatically, and more and more automation will enter the realm of more cognitive tasks. This trend will produce domain algorithms that can handle tasks such as classification, identification, navigation, planning, et cetera. The systems will be able to deal with incompleteness and uncertainty in information. A major challenge is to design for failure: realize resilience for both machine and human failures (Grant et al., 2007).

Adaptive automation has been proposed for dynamic task environments such as naval command and control. Adaptive systems adjust their level of support to the circumstances: low when possible, high when required, with the goal of keeping the operator near optimal workload levels in all circumstances (Scerbo, 1996, Rouse, 1988). If an operator is starting to get overwhelmed by the situation, a CMS that is capable of autonomous decision making could intervene and reallocate less critical parts of the work to itself so that the workload of the operator is reduced and he or she is again up to the task. When the situational demands are reducing, automation levels are lowered, shifting authority and responsibility back to the operator, hence improving operator’s awareness and skills.

4. Design rationale: scenarios, core functions, claims and use cases

Development of systems that use innovative human-machine concepts such as AA, in which there is a critical balance between situated and mutually-influencing benefits (such as workload reduction) and costs (such as reduced situation awareness), warrants the use of the sCE method. We have applied this method to develop and test AA for a task that is conducted on many naval ships, track identification.

4.1 Scenarios

For future naval missions, several high-level scenarios have been developed over the years. For the specific support system of this paper, four scenarios were developed in cooperation with training experts of the Royal Netherlands Navy. All scenarios were intended to bring a reasonable workload on the operator and included various threats or suspicious-looking tracks that contributed to the workload. Two scenarios were developed around more or less traditional air and surface warfare in a high-tension peace-enforcing situation while the other two scenarios were situated against a civilian smuggling background.

4.2 Core functions

When sensors detect an object in the vicinity of the ship, for example another ship or aircraft, a track is built as a representation of the object within the CMS. Tracks that are initially unknown must be classified (is it an airliner, a fighter, or a helicopter?) and subsequently identified (is it a neutral platform, one of our own forces, or possibly a hostile one?) . Track identification can to a large extent be handled automatically by our prototype CMS. When the number of unknown tracks becomes too high, AA increases its support level and handles the 'easy' tracks (i.e., neutral airliners and commercial vessels). The complex tracks (for example, tracks with strange behavior or suspicious characteristics) are still handled by the human operator. If things get really busy, most or all tracks are identified by the system, warning the user only of the very important, complex, or hostile tracks.

The main advantage of AA is improved workload levels for the human operator, potentially increasing overall human-machine performance. Possible drawbacks of AA are lowered SA, increasingly complex systems, mode errors and automation surprises. The perils related to SA have been described by Endsley, who called them demons of SA (Endsley 2003). To tackle these demons, we have used them explicitly in the development of our design rationale. For each demon, we identified a corresponding core function, such as 'Prevent out-of-the-loop problems'. We completed these SA-based functions with other core functions, such as the increase of performance and the prevention of some cognitive biases, making a total of 18 core functions (see Table 1).

Table 1: *Core functions for our system.*

Prevent out-of-the-loop problems	Maintain situational awareness
Prevent skill degradation	Ensure Transparency
Prevent mode errors or automation surprises	Prevent undesirable system behavior
Prevent misplaced salience	Prevent complexity creep
Prevent/limit increases of system demands	Prevent errant mental model of the situation
Prevent unwanted modes of operation	Prevent tunnel vision
Extend human memory capacity	Accomodate workload, stress & fatigue
Prevent data overload	Prevent cognitive lockup
Prevent change blindness	Improve performance

4.3 Claims

For the core functions of Table 1, AA features were derived with corresponding claims on the expected effect to the performance of the tasks in the scenarios. These claims are formulated in terms of operational consequences, which can be empirically evaluated with end-users (e.g., via standard methods for measuring human performance, effort and learning). **Table 2** shows an example for a core function, to which two claims are associated.

Table 2: *An example of a core functions with two associated claims.*

FUNCTION 1. Prevent out-of-the-loop problems
CLAIM 1. *Feature*: When the task load decreases (e.g. few tracks to handle), a lower level of automation is triggered. *Result*: The user does (almost) everything and handles more tracks, so that (s)he is sufficiently engaged in the current operation (e.g., adequate eye movements and medium arousal level), detects relevant objects in time (e.g., adequate identification performance) and is not involved in unrelated and irrelevant activities (e.g., mainly task-related behavior). CLAIM 2. *Feature*: When task load is high (e.g. more pending warnings, and higher volume and complexity of tracks), the system takes over the least important tracks and leaves the user the more important and potentially more threatening ones. *Result*: The user handles the critical tracks adequately and in time (e.g., identification performance), so that (s)he is sufficiently engaged in the current operation (e.g., adequate eye movements and medium arousal level), detects relevant objects in time (e.g., adequate identification performance) and keeps adequate knowledge of the most critical situational changes (e.g., adequate situation reports).

4.4. Requirements Baseline

Based on the core functions and corresponding features of the envisioned system, we specified the requirements (i.e. what the system is supposed to do), taking notice of the following characteristics of a good requirement specification (IEEE, 1993):

- **Unambiguity** Every requirement should only have one interpretation. It is best if each characteristic of the final system will be described using a single unique term.
- **Completeness** Inclusion of all significant requirements for the main function of the system, including the user interface.
- **Verifiability** A requirement should be verifiable; i.e. one should be able to check that the system meets the requirements.
- **Consistency** A requirement specification is consistent if no set of individual requirements described in it conflict.

- **Modifiability** A requirement specification should be modifiable so that changes to the requirements can be made easily, completely, and consistently without introducing inconsistency.
- **Traceability** A requirement specification should be traceable so that the origin of each of its requirements is clear and it should facilitate the referencing of each requirement in future development or enhanced documentation.

Each requirement can be annotated with an indication on the importance of that requirement to implement the functionality of the use case. The MoSCoW list is often used to indicate the importance. MoSCoW stands for **M**UST have this, **S**HOULD have this if at all possible, **C**OULD have this if it does not affect anything else, and **W**ON'T have this time but WOULD like in the future.

Table 3. *Explanation of the use case fields (based on Cockburn, 2001)*

[UC_Nr] [UC_name]	*Number used to link requirements to use case.*	*Example:* UseCase 3 Increasing SA after decreasing LOA
Goal	*What is achieved by carrying out the use case.*	Limit out-of-the-loop problems
Actor	*Main human (or possibly machine) actors.*	Team member of Command and Control Centre
Precondition	*Contains the state of the system or user just before using the functionality.*	AA is at the medium or high level; User has a limited view of tracks as some are handled by the system, limiting his situational awareness to ‘dangerous’ tracks.
Post condition	*What is achieved using the functionality, describes the state of the system or user.*	AA is set at a lower level More tracks will be handled by the user from now on, increasing his or her overall situational awareness.
Trigger	*Defines the event (e.g., time, alarm, …) when a user needs the functionality or how the system knows that the function needs to be carried out.*	Amount of work (pending tracks, tracks requiring user attention) is below a preset threshold level.
Main Success Scenario	*A top-to-bottom description of an easy to understand and fairly typical scenario in which the actor’s goal is delivered.*	1. After decrease of automation level, more tracks of multiple categories will be handled by the user 2. In doing so, the user quickly gets good situational awareness.
Alternative Scenario	*Other ways to succeed, and the handling of the most important failures*	……
Satisfies claim	*List of claim-numbers that link to this use case*	Claim 1, Claim 25
Satisfies requirement	*List of requirement-numbers that link to this use case*	Requirement 13

4.5 Use cases

Based on the scenarios, core functions, claims and requirements of the previous subsections, we specified a number of use cases according to a fixed format. Table 3 shows the structure that was used, extending the description of Cockburn (2001). It should be noted that the specification of the requirements with its rationale (i.e., scenarios, functions, claims and use cases) is an iterative process. During the process, specifications of all elements are refined, and harmonized (e.g., use case refinements may lead to requirement refinements).

To facilitate the transfer of the AA design specification from our research team to the system development group of the RNlN, the specifictions were worked out into an information, a dynamic and a functional model, including state transition and class diagrams (Object-Oriented Analysis of Shlaer and Mellor, 1992). Parts of the functional model were described in pseudo code.

5. Refinement processes

5.1 Review

The established core functions, claims and use cases, were reviewed by domain experts (both operational and technical) from the Royal Netherlands Navy. Based on their comments, requirements were added or updated. However, considering the system is a research and development platform and not a realistic CMS, a thorough review by domain experts was not appropriate at this time. For real space or military systems, thorough reviewing by domain experts forms a major component in the refinement of the design rationale and requirements baseline, due to the criticality of the software and the problems of proper testing in realistic situations.

5.2 Evaluation

A CMS prototype was built to evaluate the adaptive identification design (De Greef et al., 2006; 2007). This prototype proved to be sufficiently realistic to conduct a series of experiments to evaluate the claims and refine the requirements baseline. Besides identification, the prototype consisted of additional functionality, such as classification of tracks, weapon deployment, and navigation, to provide the complex environment required for proper testing of the adaptive identification module. These other modules have not been designed according to the sCE method, but although some interaction exists between these modules and the claims related to the identification task, we feel this did not effect our evaluation.

5.2.1 Participants

The participants were four warfare officers (WO) and four warfare officer assistants (AWO) of the Royal Netherlands Navy with several years of operational experience using naval combat systems on naval ships. At the time of the trials the subjects were either involved in the design and development of new Combat Management Systems for the RNLN or engaged in the training of officers at the operational school of the RNLN. The participants were available for two days and participated in teams of two, one WO and one AWO.

Figure 2: *Naval domain experts participating in the evaluation of the prototype.*

5.2.2 Tasks & Apparatus

The participants worked with a workstation called the Basic-T attached to a simulated combat management system (see Figure 2). The simulated ship resembles a modern RNlN air defense and command frigate in its sensor and weapon suite; the CMS was newly designed and built to encompass new functionality and adaptivity.

The four scenarios of section 4.1, bringing about relatively high workload situations, were used for the evaluation. The participants were given mission goals and were instructed to defend the frigate against any threats. In all cases the primary mission goal was to build a recognized maritime picture (RMP) of the surroundings of the ship and to detect and identify all platforms present. Building the RMP amounted to monitoring the operational space around the ship, classifying and identifying contacts. Decision making involved moving the ship and its associated helicopter and possibly neutralizing hostile entities. As the sensor reach of a modern naval ship extends to a large region around the ship, this represented a full-time job. In addition, the participants were responsible for the short-term navigation of the ship, steering it toward whatever course was appropriate under the circumstances and avoiding collisions with other ships.

5.2.3 Experimental design

The automation mode was either adaptive or fixed. System support in the fixed mode was comparable to the lowest level in the adaptive mode. It is important to remember that in this way even in the fixed mode the system was still giving advice and drawing attention to tracks that merited such attention and as such offered more assistance than current combat management systems would do. The major difference was the fact that in the adaptive mode the automation would scale up and the system would acquire more responsibility in handling tracks autonomously when the user's workload increased. It would scale back again as soon as the workload decreased.

Seven dependent variables were measured to test the claims:

- The *subjective workload* ratings as rated during each scenario on a one dimensional rating scale from one to five, one meaning heavy underload and boredom, three a comfortable and sustainable workload and five an overload of the operator (Zijlstra, 1993).
- The expert was closely observing the participant and estimated the *workload*, the *situation awareness*, and the *quality* and *timeliness* of the actions of the participants every two minutes using the same one-dimensional rating scale. In these cases, a one indicated a bad performance, a five an excellent performance and a three a sufficient performance.
- The *performance* in terms of tracks handled and reaction time on signals of the machine was measured.
- Lastly the *communication* served as the seventh dependent variable.

5.2.4 User experience

The experiments have shown an increase in overall human–system performance. Tracks were handled quicker, and on average much less pending tracks remained. Situation awareness of the participants in the adaptive and fixed condition was similar, as were the observed workload levels. The latter was somewhat surprising, as a decrease was expected. However, the time that became available because the automation took over work was spent on extra communication, and presumably also on other more reflective tasks related to sense making. Opinions were divided on the matter of transparency and automation surprises. Some participants said the system seemed to read his mind, because the automation was set higher at exactly the right moment. Others, however, said that automation seemed random at times.

5.2.5 Refinement

The evaluation was the first time domain experts used the system. Apart from implementation features and bugs, several problems and omissions within the prototype and the underlying claims and uses cases became evident. Two examples will be presented.

During the evaluation some observations were made with respect to system modes and automation surprises. Sometimes user decisions were overruled by the system even though the user had already inspected the system view and had discarded its conclusion in favor of his own. The sequence of system action and user action seems to be the deciding issue: if a user action occurs before system action, the user's decision should be safeguarded from unwarranted changes by the system. Apparently, the claims related to system modes were not appropriately met, and a refinement of the requirements baseline with its rational was required.

A second observation was the wish of users to be able to inspect tracks that were handled by the system autonomously during periods of high workload as soon as things were calming down and automation levels were lowered. A claim was added related to the out-of-the-loop function (see table 4) and a new corresponding use case was constructed.

Table 4*: New claim for the core function related to out-of-the-loop prevention.*

CORE FUNCTION 1. Prevent out-of-the-loop problems
CLAIM 3. *Feature*: When the automation level is lowered, the user is made aware of tracks that have been handled by the system. *Result*: Tracks that were handled by the system at high automation levels are labeled as such, so that the operator can inspect them after the level of automation has lowered (e.g., user behavior) to improve his or her momentary knowledge of the situation (e.g., adequate situation reports).

6. Conclusions & Discussion

Future naval missions set high operational, human factors, and technical demands for a system that enhances human-machine teams' capabilities to cope with complex and potentially hazardous situations. This paper presented a situated Cognitive Engineering (sCE) methodology for an integrated analyses of these three types of demands, the derivation and maintenance of a requirements baseline with its rationale, and the refinement and validation processes of the requirements (including human-in-the-loop evaluations of simulation-based prototypes). It follows an iterative human-centered development process corresponding to recent human-factors engineering methods and standards (e.g. ISO 13407 "Human-centered design processes for interactive systems"), aiming at an incremental development of advanced technology.

Application of this method provided a coherent and concise compilation of design knowledge for an Adaptive Automation (AA) module for track identification on naval ships: a theoretical and empirical founded requirements baseline with its design rationale consisting of core functions, claims, scenarios and use cases. The core functions focused on coping with workload dynamics and maintaining situation awareness; for each core function we specified claims on the operational effects and requirements for the system, which were added to the use case specifications. The evaluation showed that the specified AA-module helps to reduce workload and maintain adequate situation awareness during naval missions. The current requirements baseline can be used *both* to start implementation of an adaptive track-handling support module *and* to extend the AA support to other naval tasks.

The sCE methodology was recently developed and applied for the design of systems for manned space missions (Neerincx et al., 2006). Also in the space domain, it provided a practical, coherent and extendable requirements baseline for crew support that can be incrementally developed and implemented. It is interesting to note that—for a part—similar human factors theories and models were applied in the work domain and support analysis. By explicating the design rationale in a similar way, it is rather easy to identify which support elements can be applied to the different domains and which elements are really domain specific.

References

Adams, B.D., Bruyn, L.E., Houde, S. and Angelopoulos, P. (2003). *Trust in Automated Systems: Literature Review*. DRDC Report No. CR-2003-096. Toronto, Canada: Defence Research and Development.

Clancey, W.J. (2004). Roles for Agent Assistants in Field Science: Understanding Personal Projects and Collaboration. *IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews*, Vol. 34, No. 2, May 2004.

Cockburn, A. (2001). *Writing Effective Use Cases*. Addison-Wesley, ISBN 0201702258.

Endsley, M.R.. (2000). Theoretical Underpinnings of Situation Awareness. In M.R. Endsley & D.J. Garland (eds). Situation Awareness Analysis and Measurement. LEA, Mahwah, NJ, USA.

Endsley, M. R., Bolte, B., & Jones, D. G. (2003) Designing for Situation Awareness: An approach to human-centered design. London: Taylor & Francis.

Grant,T., van Fenema, P., van Veen, M. & Neerincx, M.A. (2007). On regarding 21st century C2 systems and their users as fallible ePartners. *12th International Command and Control Research and Technology Symposium*. Newport, Rhode Island, June 19-21, 2007.

De Greef, T.E., Arciszewski, H.F.R.; Lindenberg, J., Van Delft, J.H. (2007). Adaptive Automation Evaluated; Report Number TBO-DV 2007 A610; TNO Defence Security and Safety, Soesterberg, the Netherlands.

De Greef, T.E., Arciszewski, H.F.R., Van Delft, J.H. (2006). Adaptive Automation using an object-orientated task model. Proceedings MAST conference, 4-6 september 2006 Acropolis Convenion & Exhibition Centre, Nice, France.

Hoc, J.-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human–Computer Studies,* 54(4), 509–540.

Hollnagel, E., and Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man–Machine Studies*, 18, 583–600.

IEEE (1993). *IEEE Recommended Practice for Software Requirements Specifications*. IEEE Std 830-1993.

Klein, G.. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press,

Schneider-Hufschmidt, M., Kühme, T., & Malinowski, U. (1993). Adaptive user interfaces : principles and practices (*Human factors in information technology*, 10 ed.). Amsterdam : North-Holland.

Maanen, P.P. van, Lindenberg, J., and Neerincx, M.A. (2005). Integrating human factors and artificial intelligence in the development of human-machine cooperation. In H.R. Arabnia, and R. Joshua (Eds), *Proceedings of the 2005 International Conference on Artificial Intelligence (ICAI'05)* (pp.10-16). Las Vegas, NV: CSREA Press.

McCrickard, D.S. and Chewar, C.M. (2006). Designing Attention-Centric Notification Systems: Five HCI Challenges. In: J. Chris Forsythe, Michael L. Bernard, and Timothy E. Goldsmith (Eds.). *Cognitive Systems: Human Cognitive Models in Systems Design*, pp. 67-89. Lawrence Earlbaum.

Mohammed, S. & Dumville, B.C.. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behaviour,* 22(2), pp. 89-106.

Naikar, N & Sanderson, P.M. (2001). Evaluating Design Proposals for Complex Systems with Work Domain Analysis, *Human Factors*, 43(4), 529-542.

Neerincx, M.A.. (2003) Cognitive task load design: model, methods and examples. In: E. Hollnagel (ed.), *Handbook of Cognitive Task Design*. Chapter 13. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 283-305.

Neerincx, M.A. & Lindenberg, J. (forthcoming). Situated cognitive engineering for complex task environments. In: Schraagen, J.M.C., Militello, L., Ormerod, T., & Lipshitz, R. (Eds). *Natural Decision Making & Macrocognition*. Aldershot, UK: Ashgate Publishing Limited.

Neerincx, M.A. (2007). Modelling Cognitive and Affective Load for the Design of Human-Machine Collaboration. In: D. Harris (Ed.). *Engineering Psychology and Cognitive Ergonomics, HCII 2007*, LNAI 4562, pp. 568–574. Berlin Heidelberg: Springer-Verlag. ISBN 978-3-540-73330-0

Neerincx, M.A., Lindenberg, J., Smets, N., Grant, T., Bos, A., Olmedo Soler, A,. Brauer, U., Wolff, M. (2006). Cognitive Engineering for Long Duration Missions: Human-Machine Collaboration on the Moon and Mars. *SMC-IT 2006: 2nd IEEE International Conference on Space Mission Challenges for Information Technology*, pp. 40-46. Los Alamitos, California: IEEE Conference Publising Services.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman, and S. W. Draper (Eds) *User-Centered System Design: New perspectives on human–computer interaction*. (pp.31-62). Hillsdale, NJ: Lawrence Erlbaum Associates.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39, 230-253.

Picard, R.W. (1997). *Affective computing*. MIT Press, Cambridge, MA.

Rasmussen, J. (1986). *Information processing and human–machine interaction: an approach to cognitive engineering.* Amsterdam, Elsevier.

Rosson, M. B. and Carroll, J. M. (2001). *Usability engineering: Scenario-based development of human-computer interaction*. San Francisco, CA: Morgan Kaufman.

Rouse, W.B. (1988). Adaptive aiding for human/computer control. *Human Factors* 30, pp. 431–443.

Satyanarayanan, M. (2001) Pervasive computing: vision and challenges. *IEEE Pers Commun* 8(4):10–17.

Scerbo, M.W. (1996). Theoretical perspectives on adaptive automation. In: Parasuraman, R. and Mouloua, M., Editors, 1996. *Automation and human*

performance: Theory and applications, Lawrence Erlbaum Assoc, Hillsdale, NJ, pp. 37–63.
Scerbo, M.W. (2001). Stress, workload and boredom in vigilance: a problem and an answer. In Hancock, P.A. & Desmond, P.A. (ed.), *Stress, Workload and Fatigue*. Mahwah, New Jersey, Lawrence Erlbaum Associates.
Schmorrow, D., Stanney, K.M., Wilson, G. and Young, P. (2005). Augmented cognition in human-system interaction. In: G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed.). New York: John Wiley.
Shlaer, S. and Mellor, S.J. (1992). *Object Lifecycles: Modeling the World in States*. New Jersey: Prentice Hall.
US ARMY PEOSTRI (2007). http://www.peostri.army.mil/PRODUCTS/SBA/ visited December 29, 2007.
Vicente, KJ (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer Based Work*, Lawrence Erlbaum Associates, London.
Weick, K. (1995). *Sense-making in Organisations*. Sage, Thousand Oaks, CA, USA.
Zijlstra, F.R.H. (1993). Efficiency in work behavior. A design approach for modern tools. (PhD thesis, Delft University of Technology). Delft, The Netherlands: Delft University Press.

Intuitive and analytic support for decision making

Marvin S. Cohen – Ewart de Visser – Elan Freedy – Amos Freedy – Melanie LeGoullon – Gershon Weltman

Perceptronics Solutions
1001 19th St N
Suite 1500
Arlington, Virginia 22209
USA

mcohen@percsolutions.com
edevisse@gmu.edu
elanf@percsolutions.com
afreedy@percsolutions.com
melaniel@percsolutions.com
gweltman@speakeasy.net

ABSTRACT. *TGDAS merges cognitive findings about real world decision making with normative decision modelling and analysis. TGDAS (1) captures decision makers' ideas about a decision problem in the natural form of scenarios, (2) captures uncertainty and identifies tacit knowledge about key differentiating factors by comparing scenarios across team members, (3) transforms the scenarios into causal models (influence diagrams), and (4) uses results to evaluate options for action, data collection, and collaboration. Finally, it translates results back into the original scenario-based format. Decision making and collaboration processes are guided by metrics derived from matching historical templates and an emergent full-fledged decision analytic model.*

KEY WORDS: *Decision Analysis, Naturalistic Decision Making, Decision Support, Collaboration, Scenarios*

1. Introduction

Collaborative decision-making that integrates information, judgment, reasoning, and conclusions by individuals separated in time and space – and often across organizations, nationalities, and cultures – is a major challenge of today's military, government, and business operations. Increasingly powerful computational and communication technologies, along with advances in research on both individual and group decision processes, afford opportunities that have not yet been fully realized. The decision aid development to be described here addresses two obstacles to supporting distributed collaborative decision making under conditions of uncertainty:

- Analytical processes required to build logically coherent decision models tend to clash with the more intuitive cognitive strategies that experienced decision makers spontaneously adopt.
- Group decision support systems tend either to structure the collaborative process or to produce a structured product, but in the absence of a specialized human facilitator do not accomplish both.

In this section, we will briefly sketch research and theory bearing on these two issues. Due to space limitations, we will focus on the first. We will summarize the "case" for different approaches by indicating the benefits shown by or suggested for those approaches. These summaries provide evaluative guideposts for the exploration of an integrative decision aiding strategy and the development of technology necessary to make the strategy work. In subsequent sections, we describe the aiding strategy and supporting technology embedded in the Tactical Group Decision Analysis System (TGDAS) and compare its functionality to the guideposts.

1.1 Decision Analysis

Decision analysis is the practical implementation of ideas developed by economists, logicians, and statisticians under the rubric of *decision theory*. Traditionally, decision analysis has been associated with a set of related but distinct modeling tools for tasks such as choice with uncertain outcomes, uncertain inference, evidence-based updating of beliefs, multi-criteria choice, and negotiation (Raiffa, 1968; Keeney & Raiffa, 1976). Preconfigured models based on such tools have been incorporated into decision support systems for direct use by decision makers or their staff in specific problem domains (e.g., Brown & Cohen, 1982). While general purpose software exists to support all the tasks, a trained analyst (or group facilitator) typically chooses among the tools in order to model a client's problem. The most exciting work in decision analysis over the past two decades has been the development of a more powerful modelling technology using Bayesian causal networks and influence diagrams (Howard & Matheson, 1989). Influence

Diagrams, which incorporate Bayes nets, provide a single integrated model for uncertain inference, choice under uncertainty, and tradeoffs among multiple objectives (Clemen & Reilly, 2001). They provide a more compact representation of relationships among variables of different types (e.g., uncertain events, decisions, and objectives) and are solved by increasingly efficient algorithms (e.g., Pearl, 1989; Jensen, 2001). Up to now, however, software implementations of influence diagrams have taken the traditional forms of preconfigured inflexible models or general purpose software used by decision analytic consultants.

Decision analytic modelling is often justified on two grounds: It ensures that reasoning and choice based on judgment conform to a set of compelling *a priori* principles (Savage, 1972),and it helps human decision makers avoid systematic errors or biases associated with intuitive decision making in novel situations (Gilovich, Griffin, & Kahneman, 2002; von Winterfeldt, 1999). Model-building activities, such as breaking a problem down into component factors, assessing probabilities of events and importance weights on objectives, and testing sensitivity of recommended courses of action to assumptions and quantitative parameters (Clemen & Reilly, 2001), lie at the analytical extreme of an analytical-intuitive continuum (Hammond, 2000). Benefits mentioned by practitioners (e.g., Brown, 2005) may be summarized as follows:

Case 1 Decision analysis provides a language for representing decision problems in a logically complete and consistent way. Decision analytic models help ensure that all relevant factors have been accounted for, that inconsistencies among judgments are identified and resolved, that interdependencies among factors are not overlooked, that tradeoffs are recognized and quantified, that assumptions are made explicit, that disparate sources and types of information and expertise are factored into conclusions, and that the logical implications of all relevant data, judgments, and assumptions are understood. It provides a transparent process whose longer term benefits include more effective justification of decisions and the improvement of decision making skills.

Recent work by behavioural economists (Binmore, 2007) and psychologists (Gigerenzer, 2000) suggests that experienced decision makers often learn, consciously or unconsciously, to compensate for decision biases associated with naive intuition: Some of the errors found in laboratory tests vanish if tested after sufficient exposure to conditions in which errors have significant consequences. These findings are consistent with the perspective of naturalistic decision making research.

1.2 Naturalistic Decision Making

Research on *naturalistic decision making* (NDM) studies the methods proficient decision makers actually use to think about real problems, especially ones that are relatively unstructured and time-constrained (Zsambok & Klein, 1997). Rather than imposing an a priori concept of how decisions ought to be made, naturalistic decision aids support methods that experienced practitioners have found successful (Lipshitz & Cohen, 2005; Cohen, 1993). These include recognizing patterns (Klein, 1993), using simple heuristics (Gigerenzer, 2000), constructing stories to explain ambiguous evidence (Pennington & Hastie, 1993; Cohen, Freeman, & Wolf, 1996), and inventing scenarios to mentally test courses of action (Klein, 1993; Schwartz, 1996; Cohen & Thompson, 2001). Virtually all range toward the more intuitive end of the analytical-intuitive continuum. NDM researchers have developed and tested training and decision aids designed to help bridge the gap between more and less experienced decision makers in skills of these kinds (Crandall and Getchell-Reiter, 1993; Cohen & Freeman, 1997).

The expected benefits of naturalistic decision making may be summarized as follows:

Case 2 Naturalistic strategies represent *trained* intuition. They embody implicit knowledge, procedures, and insights that have been learned over years of experience (and sometimes over multiple generations) in a domain. As a result they may be difficult to articulate explicitly, hence, they must often be elicited indirectly, e.g., by observation of real-world tasks, retrospective critical incident interviews, story telling, and realistic simulations. Naturalistic strategies have face validity for experienced decision makers and generate outputs in readily understood formats. In addition, they bypass the combinatorial explosion of possibilities associated with formal models, and do not demand precise numerical assessments that stretch the bounds of decision maker confidence.

Several researchers have speculated on synergies that might result from a synthesis of formal and naturalistic approaches. Recommendations include identifying and recommending a mix of strategies that best suits the particular context and phase of decision making (Bryant, Webb, & McCann, 2003), using analytical models of relatively novel tasks to monitor for and mitigate biases in intuitive decision making (Cohen, 1987, 1993a), and using intuitive methods to elicit inputs and to display and explain outputs of analytical models (Davis, Kulick, & Egner, 2005).

2.0 Overview of TGDAS

2.1 Integration of Decision Analytic and Naturalistic Decision Making

TGDAS starts the decision making process with scenarios, leverages the scenarios to create a model and evaluate courses of action, and returns to scenarios to display model implications. Figure 1 provides an overview of the integration of intuitive (naturalistic) and analytic decision making in TGDAS. TGDAS has five top-level modules or workspaces, which can be worked on sequentially and also viewed simultaneously: (1) *Develop COAs* captures decision makers' ideas about a decision problem in the intuitive form of scenarios and supports the development of alternative courses of action. (2) *Compare Scenarios* stimulates consideration of alternative assumptions and uses points of difference among scenarios to identify tacit knowledge about uncertainty and key differentiating factors. (3) *Create Model* helps users transform the scenarios into causal network models, or influence diagrams, which can be used to calculate probabilities of events and figures of merit (expected utilities) for courses of action. (4) *Evaluate COAs* summarizes expected utilities across candidate courses of action and displays a risk profile (the probabilities and utilities of specific possible outcomes) for each COA. (5) *Strengthen COA* uses the model to generate and evaluate significant scenarios, to identify critical branch points (which may be either opportunities or threats) within those scenarios conditional on different courses of action, and to prioritize options for information collection and course of action improvement.

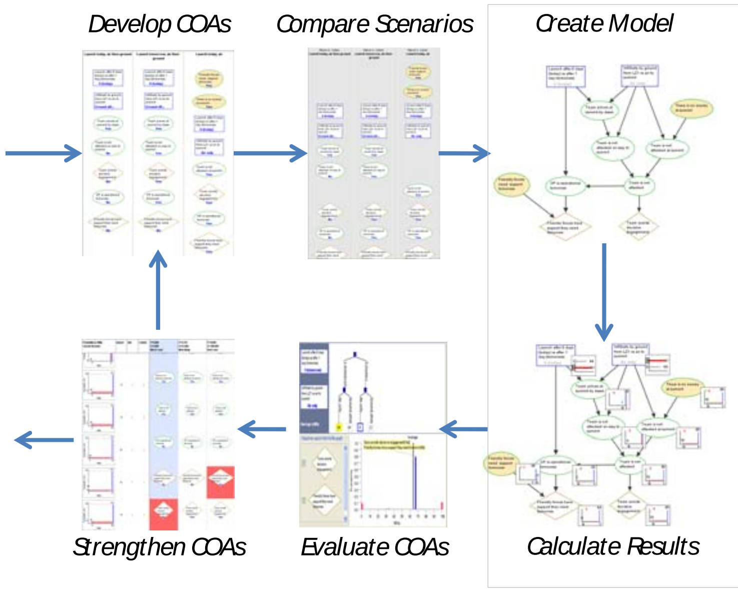


Figure 1. *Five top-level TGDAS modules.*

2.3 Critical Thinking with Scenarios

We define a *scenario* as an interesting combination of assumptions about an uncertain situation. What we call a *COA scenario* is a projected unfolding of events that follows from placing a COA in that setting. Like *stories* in the cognitive research discussed earlier, a COA scenario is a temporally and causally organized representation of how events and actions might unfold.

TGDAS encourages decision makers to make significant assumptions explicit and "test" a candidate COA in multiple COA scenarios, when time allows. It stimulates decision makers to generate COAs that have better outcomes in problematic scenarios and to look for COAs that are robust across a range of alternative possibilities.

In a hypothetical Special Forces example, the decision maker must decide on the timing and method of an insertion. The objective is to establish an observation post to support an on-going friendly conventional engagement. Due to delays in helicopter transport, the usual method of flying part way, then hiking the rest of the way, will result in arrival after dawn and possible exposure to enemy troops. Flying directly to the site may expose its location and risks an ambush if enemy are already there. Waiting a day means conventional forces will lack support.

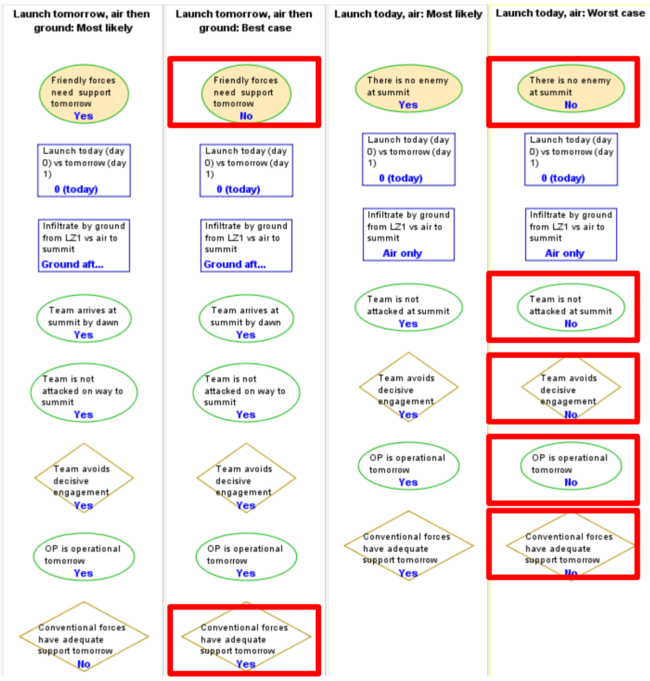


Figure 2. *COA scenarios project outcomes of two courses of action under favorable and unfavorable assumptions for each.*

Planners in this scenario might consider assumptions such as the accuracy of time-distance calculations suggesting that the summit cannot be reached by dawn by ground; that ground travel by day will result in a decisive engagement; that landing at a hot LZ will result in a decisive engagement; that friendly conventional forces really need support tomorrow; or that enemy are not in fact present at the site of the intended observation post. The decision maker might well select the last two assumptions as the most crucial. Figure 2 shows COA scenarios that vary these assumptions and permit the user to examine each of the two adjusted COAs (tomorrow by air then ground *vs.* today by air) under favourable and unfavourable conditions. Deviations from the most likely scenario for each COA are highlighted.

A scenario has three kinds of factors: Ovals are *influences* that help determine whether *objectives*, represented by diamonds, are achieved by *decisions*, which are represented by rectangles. Filled in ovals are situation assumptions, while blank ovals are causal outcomes of other factors. The TGDAS display in Figure 2, which highlights differences due to assumptions, makes it clear that each COA depends on assumptions for successful achievement of objectives: If friendly forces do not need OP support tomorrow, postponing the insertion avoids the negative outcome that led to its rejection. Moreover, if the enemy is on top of the mountain, the favoured COA is in big trouble.

2.4 Streamlined Decision Modelling

Figure 3 is a model of the causal relationships responsible for generating scenarios like those in Figure 2.

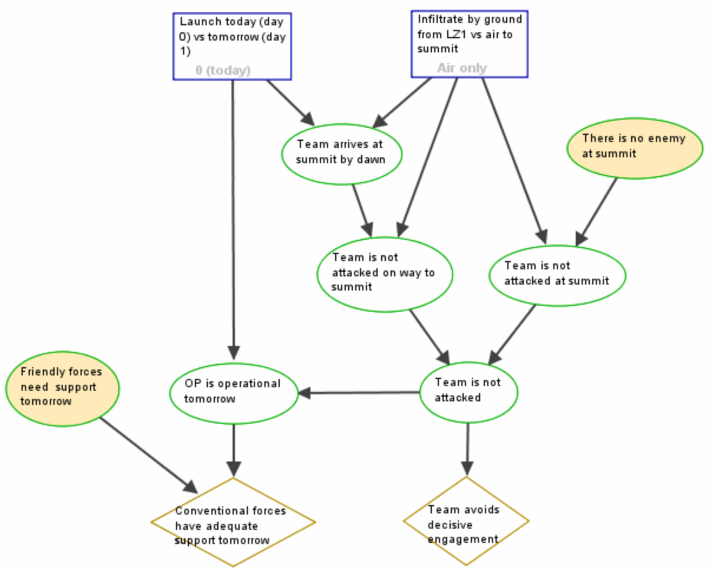


Figure 3. *Decision model.*

Moving from scenarios to a decision model is a step from qualitative reasoning about possibilities and brainstorming about COAs to quantification of uncertainty and goals and numerical evaluation of COAs. It is not that one approach neglects uncertainty and the other does not. They deal with uncertainty in different ways, which are complementary. In this example, because the COAs have different advantages and disadvantages, it may be worthwhile to look more closely at the relevant tradeoffs.

Modelling includes three activities: (1) Selecting relevant factors (or variables), (2) specifying their qualitative causal relationships, and (3) estimating numerical parameters for those relationships. These processes are streamlined in TGDAS by starting with scenarios, and by providing special graphical input devices, highly general canonical operators, and flexible templates.

Starting with Scenarios. Scenario construction jumpstarts step 1 of modelling by supplying an initial set of factors, including all the basic components of a decision model. The number of states associated with a factor corresponds to the qualitatively distinct scenarios it generates. Direct causal relationships between factors are represented by directed arcs (arrows) in the TGDAS Decision Model, as shown in Figure 3. Scenarios help jumpstart step 2 of the modelling process as well by constraining these causal relationships. For example, situation factors that occur before decisions are not causally influenced by the decisions; influences that appear after an objective do not influence that objective. TGDAS can use distinctions among factor types and factor locations to generate an initial set of arrows for users to review and modify.

Canonical causal relationships. The number of quantitative assessments required by an influence diagram is a major challenge to usability, because it tends to grow exponentially with the number of factors in the model. A dramatic reduction in the number of assessments is achieved by recognizing a set of characteristic (or canonical) causal relationships; users assess a small number of parameters of the chosen canonical model and the system automatically generates the full conditional probability tree. Pearl (1989) specified algorithms for probabilistic versions of Boolean operators, called Noisy-Or and Noisy-And gates. In Noisy-And, for example, users assess the *probability* that the truth of each parent factor is necessary (or, in the case of Noisy-Or, sufficient) for the truth of the child factor. (Thus, only one assessment is required for each parent factor.) Another kind of uncertainty is accommodated by allowing for unmodeled but suspected parent factors, yielding so-called Leaky Noisy-Or and Leady Noisy-And. Subsequent work generalized these operators to multi-state variables and to other types of relationships (Noisy-Max; Diez & Druzdzel, 2005). TGDAs has implemented the latter, with some further generalizations (e.g., allowing users to specify ordinal relationships among states of a factor in different ways for different canonical relationships in which it is a parent). These canonical relationships both simplify and reduce the number of assessments and clarify their meaning.

TGDAS users can directly specify conditional probabilities for a child node, or they can choose from a menu of canonical operators, including generalized Noisy-And and Noisy-Or, the average of parent probabilities, and the product of parent probabilities. (Others are in development.) Figure 4 shows an assessment window for a relationship between five binary parent nodes and a binary child node, which would ordinarily require 31 (2^5-1) direct assessments. TGDAS displays the full conditional probability tree in the scrollable panel on the right. The user, however, has selected *Condition1 AND condition2…* operator (i.e., Leaky Noisy-And) from the drop down menu at the top. As a result, the user only assesses a single number for each of the five parent factors. The small tree to the left of the full tree is an example of one such assessment. TGDAS uses these five assessments to generate probabilities for the large tree in the right pane for the user to view in real time.

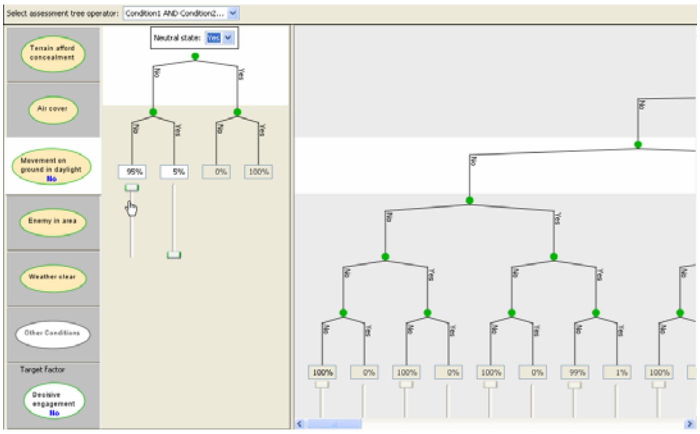


Figure 4. *Assessment window using canonical causal relationships.*

Templates. Although TGDAS provides a general purpose decision analytic modelling capability, its primary purpose is to support relatively non-technical planners, or members of a planning staff, in relatively time constrained environments. Typical users cannot be expected to build models entirely from scratch. For this reason, TGDAs can store and display a library of prepared templates customized to a specific problem domain. Templates may be prepared ahead of time by subject matter experts or stored on the spot by decision makers for later re-use. TGDAS automatically retrieves a subset of the stored templates a for display based on the match of their associated attributes to the characteristics of the current problem as specified and updated by the user. They are displayed in a hierarchical system of folders that is completely customizable and searchable by users.

Decision analytic models are typically handcrafted for each new problem. Such models can be saved, but a complete model, full of details pertaining to the specific problem for which it was developed, is not likely to be re-used without extensive revision. We have tried to design TGDAS templates to be as general and flexible as possible, to overcome this bottleneck.

TGDAS templates can be any size, ranging from a single factor to a network of many factors. Right clicking on a single factor in the Template library displays all multi-factor templates in which it appears. Multi-factor templates include causal relationships and quantitative parameters (i.e., prior and conditional probabilities and importance weights) learned over time. Single factor templates include prior probabilities. Users can drag and drop any number of templates into a scenario or model, then modify them in any way they like, deleting some elements, adding others, linking to other factors, combining with other templates, and so on.

Two additional features of TGDAS templates significantly magnify the effects of this flexibility: First, factors in templates may have "slots" marked by curly brackets enclosing an indexed category name, e.g., {ENEMY1} or {TIME1}. When templates are added to the model, users can fill in the slots with names of specific instances. All slots with the same category name and index are interpreted as referring to the same instance and are filled in automatically. Conversely, the same template can be added to the model multiple times if the user edits the category name or its index, e.g., {Madhi Army} or{ENEMY2}, to distinguish the instances. This allows general- purpose templates to be stored with quantitative parameters corresponding to prior probabilities for the categories in question, and for more accurate probabilities to be acquired over time for specific sub-categories or instances.

Second, templates are automatically self-aggregating. That is, if two or more templates are placed in the model workspace with a common factor and the same slot fillers, they automatically unify on that factor (i.e. it only appears once in the model). For example, two templates automatically merge to create a significant portion of the decision model in Figure 3, based on the shared element, *Team is not attacked*.. Subject to minimal consistency constraints (e.g., avoiding directed cycles), the factor is linked in the model to the same factors to which it is related in both templates. Quantitative parameters from the two templates are always preserved in a probabilistically consistent way. For example, if the same factor functions as a child node in two templates, and in both cases it is assessed by the same canonical relationship (e.g., both templates used to create the model in Figure 3 use Noisy-Or for the shared factor, *{FRIENDLY1} is not attacked*), quantitative assessments of the causal impact of individual parent factors remain exactly the same when both parents are present. If two templates use different canonical relationships, both sets of quantitative assessments are saved; the user can choose between them or add an intermediate variable to capture both relationships. Assessments from templates that share both a parent and a child node are usually averaged.

These features in combination allow TGDAS users to rapidly construct new models from prepared elements, even in novel circumstances.

2.5 COA Evaluation (Cases 1 and 8)

Expected Utility. Figure 5 shows the results of using the decision model (Figure 3) to calculate the mathematical implications of the assessments. The COA

Evaluation module shows all COAs, represented as combinations of actions, with their aggregate evaluation score, or expected utility. Expected Utility here is the percentage of the maximum attainable value a particular COA would on average achieve, where the maximum is complete success on all objectives. In this example (Figure 8), inserting by air today scores 82%. It seems that the advantages of this option far outweigh the disadvantages; the second best COA, inserting by air then ground tomorrow, is worth 20% less.

TGDAS shares the perspective of most decision analysts, who do not equate correct decision making with immediate acceptance of the option that appears to have the highest expected utility. TGDAS is intended to help users explore the strengths and weaknesses of different COAs, and the model itself, in greater depth than a single number can reveal. One possibility is that the apparently optimal action may involve too much risk. In addition, there may still be time to collect more information, resolve disagreements, revise the plan to make it more robust, and dig deeper into the decision model – based on a judgment that one or more of these options may cause Expected Utility scores to shift enough to affect the choice of a COA. TGDAS helps users weigh the benefits against the costs of further information collection, discussion, option generation or analysis.

Risk. The same expected utility (say, 70) might reflect outcomes that reliably cluster near 70 or outcomes that range from, say, 5 to 95. The bar chart in the lower part of the COA Evaluation screen (Figure 5) is a risk profile that shows the outcomes that might actually occur for a selected COA, and which taken together produce the average. Horizontal location of a bar represents utility of an outcome, while the bar's height is its probability. Users who are averse to risk may prefer an option with a lower expected utility if it is more likely to deliver an outcome above the decision maker's "security" level.

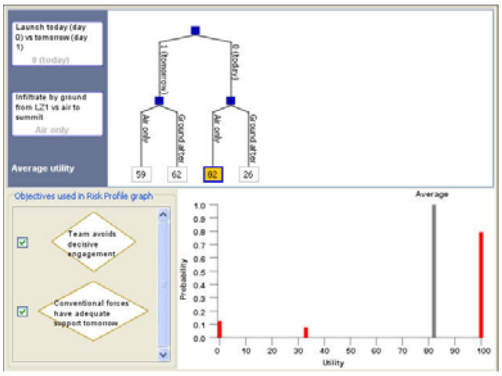


Figure 5. *COA evaluation.*

What if analysis. Once results are calculated, the decision model becomes a tool for exploration of alternative COAs and scenarios. In Figure 6, the user has chosen to display the probability and utility (or degree of preference) of each state of a factor by means of a small "decision graphlet" appearing next to the factor. Right clicking on a graphlet produces a magnified version, illustrated here for the assumption, *There is no enemy at summit.*

The vertical axis represents probability on a 0 to 100% scale; so a higher bar means a more likely state. The horizontal axis of a graphlet represents expected utility conditional on discovering the relevant state to be the case, on a 0 to 100 scale, so that lines further to the right represent more preferred states. In general, one prefers to see a graph with tall bars far to the right, especially for objectives.

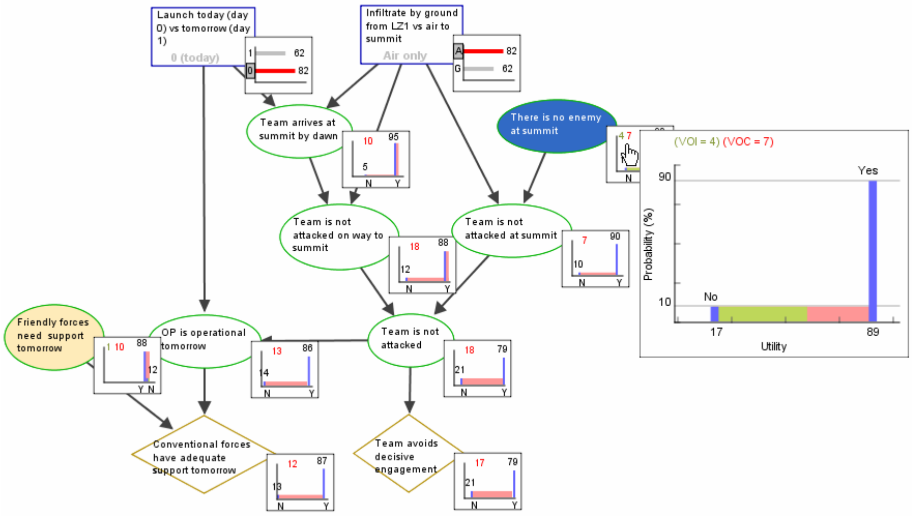


Figure 6. *Decision model with display of value of information and control*

A similar right click menu enables the user to explore the impact of different COAs and other factors on objectives. TGDAS (unlike other decision analytic tools) distinguishes two sorts of hypothetical questions, because *learning* that the state of a factor is, say, "yes" has different implications from *making* the state of factor "yes." (Explanatory inferences based on learning the state of a factor are not valid when the decision maker is the cause of the state.) Results of such exploration can suggest new options, to collect more information or to shape the battlefield.

Information and Control. In Figure 6, the user has chosen a TGDAS display option that summarizes the results of what-if analysis for each individual factor. Shaded rectangular areas represent the extra utility that would be gained by obtaining information about the true state of the node (green rectangles) or controlling the true state of the node (green plus red rectangles). A green area extending to the right of a vertical line is the utility of learning that the node is in the state represented by that line and, based on that knowledge, changing the COA to adapt to that state. The red area extending to the right of a vertical line is the utility gained by *substituting* the most preferred state of the node for the state represented by the line. Note that the geometry of the rectangular shape matches the mathematics of what it represents: Each area is the product of the probability that the node is in a particular state and the change in utility due to adopting a better COA for that state ort changing the state to a better one.

3. Discussion and Conclusion

Using TGDAS, team members create multiple scenarios that differ in the actions, situational factors, or mission outcomes they contain. TGDAS builds formally correct decision models indirectly rather than by frontal assault, via a *scenario matrix* that compares scenarios and identifies significant branch points, and by means of pre-stored editable *templates* that supply variables and relationships matching the scenario branch points in the relevant type of mission and situation. Model-based analyses, in turn, enable users to focus on variables that have the most potential impact on decisions and outcomes. The model helps regulate a process of critical thinking and discussion (possibly asynchronous) about scenarios, in which team members share information and ideas about critical divergences; clarify, challenge, and defend their respective positions; and make hidden assumptions explicit. Formal models emerge from stories and in turn guide story evaluation, provide insight into the respective strengths and weaknesses of different scenarios, and allow a more robust situation understanding and plan to emerge.

References

Bryant, D. J., Webb, R. D., & McCann, C. (2003). Synthesizing two approaches to decision making in command and control. *Canadian Military Journal*(Spring), 29-34.

Clemen, R. T., & Reilly, T. (2001). *Making Hard Decisions with Decision Tools*. Pacific Grove, CA: Duxbury.

Cohen, M. (1993). The bottom line: naturalistic decision aiding. In G. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds.), *Decision Making In Action: Models and Methods*. Norwood, NJ: Ablex.

Cohen, M., Freeman, J.T., & Wolf, S. (1996). Meta-recognition in time stressed decision making: Recognizing, critiquing, and correcting. *Human Factors, 38*(2), 206-219.

Davis, P. K., Kulick, J., & Egner, M. (2005). *Implications of modern decision science for military decision-support systems*. Santa Monica, CA: The RAND Corporation.

Diez, F., & Druzdzel, M. (2005). *Canonical probabilistic models for knowledge engineering*. Madrid: CISIAS, UNED.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Howard, R. A., & Matheson, J. (1989). Influence Diagrams. In R. A. Howard & J. E. Matheson (Eds.), *The Principles and Applications of Decision Analysis* (Vol. II, pp. 719-62). Menlo Park, CA: Strategic Decisions Group.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. New York: Springer.

Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In J. O. G.A. Klein, R. Calderwood, & C.E. Zsambok (Ed.), *Decision Making in Action: Models and Methods* (pp. 138-47). Norwood, NJ: Ablex.

Lipshitz, R., & Cohen, M. S. (2005). Warrants for prescription: Analytically and empirically based approaches to improving decision making. *Human Factors, 47*(1), 121-130.

Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge UK: Cambridge University Press.

Pearl, J. (1989). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo CA: Morgan Kaufmann Publishers.

Pennington, N., & Hastie, R. (1993). A theory of explanation-based decision making. In G. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds.), *Decision Making in Action: Models and Methods* (pp. 188-201). Norwood, NJ: Ablex.

Savage, L. J. (1972). *The foundations of statistics.* New York, NY: Dover Publications.

von Winterfeldt, D. (1999). On the relevance of behavioral decision research for decision analysis. In J. Shanteau, B. A. Mellers & D. A. Schum (Eds.), *Decision science and technology: Reflections on the contributions of Ward Edwards*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Zsambok, C. E., & Klein, G. A. (1997). *Naturalistic Decision Making.* Mahwah, NJ: Erlbaum.

Analysis of Air Traffic Controllers Decisions

David Annebicque, Igor Crévits, Thierry Poulain, Serge Debernard

LAMIH UMR CNRS 8530
University of Valenciennes, UVHC, Le Mont Houy
F 59313, Valenciennes, CEDEX 9 ; France
{david.annebicque, igor.crevits, thierry.poulain, serge.debernard}@univ-valenciennes.fr

ABSTRACT. *This paper presents an approach based on a multiple criteria decision making methodology (MCDM) to analyse the decisions of Air Traffic Controllers. This study will allow modeling some tools able to assist the controllers in their tasks. Currently the platform AMANDA assists the two controllers of one sector of control. This platform was very pleasant welcome, and we wish to extend these principles to the adjacent sectors, and include tools to help the cooperation between controllers on adjacent sectors. This approach is composed of three points. First, it is necessary to model the decision-making process of controllers. The second point is application of the MCDM which will guide all the study, and finally a repertory grid technique will be applied in order to support the operational aspect of MCDM.*

RÉSUMÉ. *Ce papier présente une démarche basée sur une Analyse MultiCritère d'Aide à la Décision (MMCAD) pour analyser les décisions des contrôleurs du trafic aériens. Cette étude permettra de modéliser des outils capables d'assister les contrôleurs dans leurs tâches. Actuellement la plateforme AMANDA assiste les deux contrôleurs d'un seul secteur de contrôle. Cette plateforme a été bien accueillie par les contrôleurs, et nous souhaitons étendre les principes aux secteurs adjacents, et introduire des outils pour aider la coopération entre ces contrôleurs adjacents. Cette démarche se décompose en 3 points. Premièrement, il est nécessaire de modéliser le processus décisionnel des contrôleurs. Le deuxième point est l'application de la MMCAD, qui guidera toute l'étude, enfin une technique d'analyse de grille sera appliquée dans le but de supporter l'aspect opérationnel de la MMCAD.*

KEY WORDS: *Decision Analysis, Multiple Criteria Decision Making (MCDM), Repertory Grid, Human-Machine Cooperation, Situation Awareness, Air Traffic Control (ATC).*

MOTS-CLÉS: *Analyse des décisions, Méthodologie MultiCritère d'Aide à la Décision (MMCAD), Analyse de grille, Coopération Homme-Machine, Conscience de la situation, Contrôle de trafic aérien.*

1. Introduction

The DGAC (French acronym for General Direction of Civil Aviation) foresees that in the next 10 to 20 years the air traffic will double or even triple. This increase of traffic will be impossible to assume with the current control methods. Indeed from a mental point of view the number of aircraft and information to manage will be considerable and operators risk to be overloaded at certain times of the day. It becomes necessary to assist controllers in their work, by offering them new tools and new ways of working that will allow them to assume this increase.

The LAMIH has been working with the DGAC for many years in this optical. The laboratory has developed several platform with a common philosophy which is to keep the operator at the centre of the loop, and thus to develop cooperative systems. The objectives are to extend the principles developed in the last platform, and for this it is necessary to understand how the planning controllers (PC) work and how the PC of adjacent sectors cooperate.

This paper begins with a presentation of Air Traffic Control (ATC), with its problematic of traffic increase. The second part presents the project AMANDA (Automated machine MAN Delegation of Action), in its current version, and the objectives of the new version. The third part presents the approach which is put in place to understand and analyse the activities of the PC. This approach is divided into three parts: the decision-making process modelling, application of a methodology multiple criteria decision making (MCDM) to support the collection of information and then the repertory grid as operational aspects of the knowledge acquisition. Finally a last section presents a first decision model based on the expertise of a “decision engineer”.

2. Management of en-route air traffic

2.1. *Organisation of Air Traffic Control*

The ATC is organized in 3 layers: “Airport control”, “Approach and terminal control” and “en-route control”. This latter layer manages flights passing through in the airspace between departure airport and the approach control of the destination airport. The objective of en-route ATC is to guarantee the safety of aircraft and their passengers. To do this the controllers must take care that aircraft remain separate by a minimum separation distance (5NM in horizontal plan, and 1000 ft in vertical level), while ensuring that they also respect the economic constraints of time and fuel consumption.

To simplify the management and the supervision of traffic, airspace is divided in geographical sector and in level of 1000 feet. A sector is permanently supervised by two controllers: a Planning Controller (PC) and an Executive Controller (EC). The PC coordinates the movement of aircraft between his sector and the adjacent sectors.

This coordination consists in a negotiation of entrance levels and exit levels. The PC takes care too, to regulate the workload of EC. The EC is in charge of sector supervision, namely to supervise that the aircraft respect the flight plans, and to maintain the safety distances. If the EC detects a possibility of crossing under this safety distance, he/she must do all is possible to restore the separation distances and avoid the conflict. Generally it is necessary to reroute one of the aircraft, and then to take back this aircraft in its original trajectory when the separation is guarantee. This action is called conflict resolution.

2.2. *Motivation of the study*

Some statistics can quickly demonstrate the problem of air traffic control. In 25 years (1977 to 2002) the traffic transiting in the French airspace has increased of 250% (DGAC). The Air traffic is today higher than 2.500.000 aircraft per year that gives on average 7.000 aircraft per day. In a sector like Bordeaux for instance, the controllers must manage 20 to 25 aircraft per hour, this is the reasonable limit of workload for the controllers. The DGAC foresees that in 10 to 20 years these statistics go double even triple. The controllers risk thus to have some difficulty managing this increase with actual tools (radar view, strip, telephone…) and risk to be overloaded to certain moment of the day, and this to the detriment of the security. Reduce sectors is now impossible, because the conflicts resolutions need a minimal geographical area.

A total automation of the ATC is impossible too, outside psychological consequence that this would induce to the passengers, the techniques to realise this automation imply an entirely instrumentation of aircraft, that is not economically conceivable. Currently to avoid this overload of controllers, who could not maintain an optimal security level, different solutions are adopted. For example, the planning of flights and the regulation of the departure of airports, or the coordination between sectors. These solutions allow reducing the complexity of air conflict even to avoid that these conflicts had really happened.

The question is approached in terms of assistance to the controllers. Tools which help to improve the regulation of the workload of controllers are proposed. It is imperative that these tools come within perfectly the control tasks and the work of controllers (as a pair, as individually), to produce a beneficial effect.

3. Project AMANDA

It is in this perspective that the project AMANDA (Debernard *et al.*, 2002; Guiost *et al.*, 2004), as well as other projects developed in the laboratory (Crévits *et al.*, 1993; Lemoine *et al.*, 1996), takes its place. These projects have always a same philosophy, which is to keep Human, operator, in the control loop. These projects do not research to fully automate the management of ATC, which would result in loss of competences for the operators, as well as a loss of situation awareness (SA) (Endsley 96; Endsley *et al.*, 1999).

3.1. ***AMANDA V2***

AMANDA V2 assists controllers (PC and EC) of one sector, in giving some tools which are able to allow a delegation of task (Guiost *et al.*, 2004), but also some tools which permit to share rapidly a same representation of airspace, and conflicts, and thus to maintain a common SA.

3.1.1. *STAR*

AMANDA integrates a tool of trajectory calculation and of assistance to the resolution of air conflict, called STAR (French acronym for Tactical help system of Resolution). STAR works in cooperation with the controller. The controller detects a conflict (STAR does not perform the task of conflict detection); he/she has the possibility to use STAR to help his/her to resolve the conflict. To do this the controller indicates the strategy (called directive) that he/she desires apply to resolve the conflict. A directive or strategy is like, for example, "AFR1542 PASS_BEHIND KLM1080". STAR takes into account this directive in order to propose a solution. To do this STAR calculates the whole of trajectories which response to the directive, without, of course, creating new conflict. STAR proposes then ONE trajectory to the controller (after a choice in function of some criteria like number of deviation…). The controller can examine the solution proposed by STAR. If the solution is satisfactory, the controller can delegate the effectuation that means the sending of instructions to aircraft. In this case STAR has in charge to communicate instructions (change of heading, FL…) directly to the aircraft. The controller is thus discharged of the effectuation and communication with pilots.

3.1.2. *Common Work Space*

The Common Work Space (CWS) (Bentley *et al.*, 1992; Pacaux-Lemoine *et al.*, 2002) is an essential notion introduced with AMANDA. This space allows a sharing of information between all agents (human, like controllers and artificial like STAR). Each agent can introduce new information in this CWS according to its competencies (know-how), and in accordance to its role (authority) in the process. All the agents can take this information into account in order to carry out their tasks, or to control and check those of the other agents.

This CWS allows mainly maintaining a common situation awareness between the two controllers, to share their representation of the problems (here in sense of air conflict or loss of separation) to supervise and/or to resolve. The controllers have the responsibility to maintain up to date this space, in order to, on the one hand to preserve a coherent “picture” of the situation and airspace, and on the other hand to inform the platform, and mainly STAR, with the conflicts that they detect.

3.2. *Experimental results*

The principles presented have been tested experimentally with qualified controllers (Guiost *et al.*, 2007). Four pairs of controllers have tested AMANDA V2. Three scenarios of traffic have been designed to test three experimental situations differentiated by the level of assistance provided. The scenarios simulated a realistic traffic but twice more loaded than into reality.

From a general point of view, the general principle of providing assistance allowing a regulation of workload has been recognized relevant by controllers. In the situation where STAR and CWS assisted controllers, 93% of clusters expected have been created. For 75% of these clusters a directive or a differed order (DO) have been selected and 63% of these directives or DO have been delegated to STAR. In terms of workload, the tools available have allowed controllers to manage without any difficulty the traffic load.

The experimentations have emphasized that the tools given, have favoured the anticipation of controllers, improving thus the accomplishment of control objectives. However this anticipation has been increased by the absence of simulation of adjacent sectors. Indeed, the PC was liberate of the management of coordinations with the adjacent sectors, and has an entirely liberty to change the level of entry or exit of aircraft. This excess of anticipation has allowed the PC to act on traffic and aircraft in order to reduce the number of conflicts.

The module STAR has proved unsuited to the practice of the EC. Indeed, the calculation methods used to provide a trajectory avoiding the aircraft at the meadows of the standard separation and returning to the original trajectory in the shortest. The controllers were then disconcerted by the efficiency of STAR. In addition, taking into account the unstable aircraft (changing flight level) by STAR was not optimal, as is the concept of "interfering aircraft” (aircraft that the system considers necessary to take into account to solve the conflict and in many cases an unstable aircraft). The controllers do not seem to have this notion of interference, for them an aircraft is in the conflict or it is not.

3.3. *AMANDA V3*

The objectives of this new study are (Annebicque *et al.*, 2008): the integration of adjacent sectors and improvement of trajectory calculation, STAR. The integration of adjacent sectors consists of an extension of CWS principles to the cooperation between PC of adjacent sectors. This new CWS will:

- Facilitate the negotiations between sectors; by allowing quickly visualizing the flight concerned by negotiations. This way, the workload, the time necessary, and the risk of ambiguity are reduced.

- Allow sharing between sectors: changes in the trajectory of aircraft, this should help to reduce uncertainty about the positions and conditions of entry for flights in a sector.

Concerning the module of calculation STAR, it is too much "efficacious" compared to the methods and habits of controllers. Indeed, the calculation methods use mathematical methods to provide the new trajectory answering to the directive, and allowing resolving the conflict. That gives thus "perfect" trajectories. This tool does not include additional factors introduced by controllers such as a safety margin above the minimal separation distance (15NM), a deviation rate (heading) comfortable (<30 °), an anticipation of unstable aircraft.

3.4. *Approach*

The study is divided into three phases. The first phase focuses on the analysis and the structuring of the decision-making process. First of all, an analysis of the decisions of PC in phases of coordination with the adjacent sectors is required. But these decisions must be put in coherence with the decisions of PC in the intern management of his/her sector. They must be also put in coherence with the intern management of the sector by the EC. This phase will conduct in a description of a coherent decision-making process. This point is developed in details in section 4.1.

The second phase is methodological. It aims to structure each decision of decision-making process. A general methodological framework must be researched to promote the coherence of each decision considering their links with the decision-making process. Several participants contribute to the decision; each one according to his/her own value system. The methodological framework must also allow structuring the exchanges between the different participants in the decision. It should also help to identify, to represent and to manipulate the different value systems of the participants. This phase is described in section 4.2.

The third phase is classic in the field of decision-support, it is the modelling phase. This phase aims to identify and to structure the elements allowing designing some tools to aid the decision makers. It is therefore necessary to collect the decisions elements handled by the controllers. It is important to note that the controllers are not the only holders of these elements. Staffs of Air Navigation responsible for the training of controllers have a favourable position in this phase. It is the same with “decision engineer”, designers of tools present in AMANDA. However air traffic controllers are the only ones who can make validation judgments of model (through the results they produce). This phase is developed in section 4.3.

4. Structuring of problem

4.1. *Decision-making process*

The decisions of control are in line with a continuum. At the most complete level, they consist of to change the trajectory of the aircraft by applying adjustments to flights parameters of aircraft in order to resolve, operationally, a conflicting situation. The EC has in charge this operational level, and he/she can cooperate with STAR (axis 3 on the figure 1).

Previously, these operational decisions have been prepared by the PC who has information before EC. The PC may already identify a conflict situation and inform the EC at the good time. This latter will integrate this new situation in the management of his/her traffic. He/she will specify the preparation, and the "pre detection" of PC to be able to operationalize later. The EC occupies a central position in the tactical level (figure 1, axis 2) in collaboration with the PC. The CWS constitutes a cooperation help between the two controllers.

Finally, at the sector level (figure 1, axis 1), the PC is the first to have available information about flights which preparing to pass in the sector. The PC has a strategic vision of potential conflicting situations. The CWS enables him/her to explain this vision and to share it in order to the EC exploit these information to manage the sector. In the context of this strategic management, PC may contact adjacent sectors in order to, for example, change flights levels of entry of aircraft to avoid a conflict in his/her sector and thus reduce preventively the workload of EC. The CWS is therefore quite naturally an area of strategic management between PC, coherent with the tactical management by the synthetic vision which it presents.

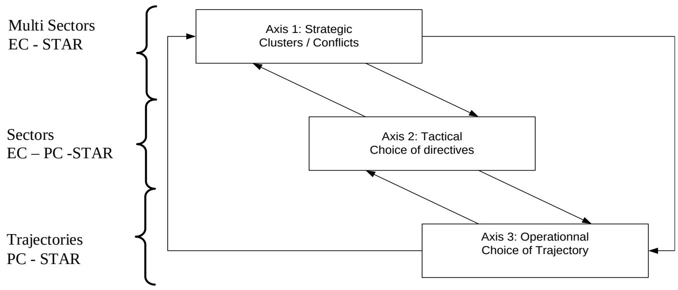


Figure 1. *Synthesis diagram of three axes of the study, and the links between them.*

These three axes are studied as independently as possible with the aim of obtain refined results and focused results on a specific problem, and therefore provide the opportunity to go into details each level. But the axes are interconnected; indeed choice a trajectory without having problems is somewhat surprising. It is thus quite logically, that appeared influence between axis 1 and 2 and between axis 2 and 3. The existence of operational decisions quickly appears plausible in the current state of our thinking. These quickly decisions correspond on direct link between the axis 1 and the axis 3. This possibility will be studied.

4.2. Multiple criteria methodology.

The job of an air traffic controller is characterized by the research for a compromise between different systems of value. This is typically the concept of managing flows aircraft. Thereby, the controllers act on the traffic by ensuring optimal security, while trying to reduce delays and the consumption of fuel. ATC is by nature multiple criteria. It is quite unrealistic to summarize the actions taken by the controllers in a single goal, which would be safety, the cost or time. In addition, the actions of the controllers constitute the terminal part of the management of control situations. They are therefore the result of decisions taken previously by controllers. Consequently, it seems appropriate to address the design of aid with the point of view of the methodology of Multiple Criteria Decision Making (MCDM).

The MCDM methodology (Roy 96) replaces the concept of decision as resulting from the wider concept of the decision-making process wherein several participants can play a role in their own interests. For that reason, the study of decision-making problem is itself accentuated.

The MCDM methodology proceeds in four levels (figure 2). The first level is to clearly define the potential actions. The potential actions are all possibilities (real or fictitious) on the basis of which the decision is made. The criteria (level 2) are the factors (witness of the decision) which characterize the potential actions for decide. Preferences (level 3) are a set of rules by which the potential actions are put in relation across criteria. Finally, the level 4 is the establishment of a recommendation; it is the operational level of the methodology, the implementation.

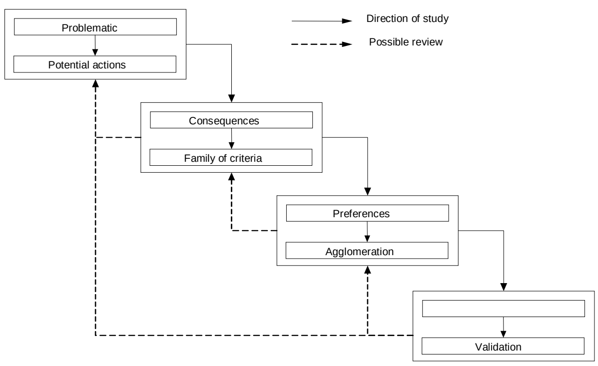


Figure 2. *Synthesis diagram of the Multiple criteria Decision Making methodology (MCDM).*

The study of three axes independently will therefore lead to conduct three MCDM, and to define three problematic; obtain three families of criteria… But the recommendations (level 4) will be most certainly more general. For example during cooperation between PC, strategic level, the PC can be lead to justify his/her requests, operational level. In any case, it will result of these three studies only one cooperative system, a single platform. This platform will be composed of different decisions, different tools responding and corresponding to each of the recommendations, but they will be grouped within a single environment, CWS.

Human-Machine Cooperation aspect is the unifying thread of this study. This aspect takes place essentially in the level 4, the recommendation. The main objective is to understand the steps and the use that the controllers do of adjacent sectors, their manner of cooperate… Human-Machine Cooperation aspect can thus be considered as a synthesis of MCDM.

4.3. Repertory grid

Repertory grid is a methodology developed by an American psychologist, Georges Kelly (Kelly 55), in order to study the psychological construct in pathological case (schizophrenia…). This method will allow comparing « elements » (different event, actions, states or entities). To do that, the method will « force » the subject to ask him/her and thus establish a list of « constructs », as exhaustive as possible. The constructs are divided in two groups: the similarities and contrasts. Each construct (similarities or contrasts) will be then evaluated or weighted in function of the different elements of the grid. The elements, the constructs, and the weighting will represent the « construct map » of the subject.

The standard representation of a repertory grid is a matrix, with in column, the elements and in rows, the constructs. The constructs are divided in two poles and generally the similarities, obtained in first, are on the left of the matrix and the contrasts are on the right (see table 1). At the intersection of each pair element-construct there is a weighting given by the subject and which represents how the subject applies or evaluates a construct in relation to an element.

	ELEMENTS	
Constructs, Similarities	Weight	Constructs, Contrasts

Table 1. *Standard representation of repertory grid.*

To establish the list of constructs, several methods exist; one of the most common is the “triad methods” which consist in taking 3 elements and to ask itself “what two elements have in common that the third has not”. The list of similarities and contrasts is thus obtained in comparing each element with this method, or until the subject has no new construct to propose. The second step is to complete the grid with weightings. It is possible to use 5 weightings (1, 2, 3, 4 and 5) or 9. The most frequently use is a weighting of 5, avoiding a too important dispersion. The principle is the following: the subject uses a 1 if the element is near of the notion of similarities of the construct and 5 if it is rather the contrast side. A 2 (similarity) or 4 (contrast) will be used if the notion, evoked by the construct, is less evident than previously. Finally a weighting 3 exists, if the subject can not, or be not able to evaluate the element, the subject has no preference.

The repertory grid is the operational part of the MCDM. The MCDM only guide and propose several questions, the repertory grid will guide the interviews and the collect of information and responses. The important points of the MCDM are present in the repertory grid. Notably the potential actions appear as elements as the grid, the constructs can be assimilated to the « elementary consequences » which serve to define the family of criteria. The weightings correspond to the different level of preferences: 1 and 5 for a strict preference, 2 and 4 for a low preference and 3 for an indifference (it is not possible to choose), or an incomparability (the constructs have no sense for this element).

The choice is made to divide the problem “detection/resolution of conflict” in three axes. These axes will be studied independently, three MCDM will be applied and hence at least a repertory grid for each MCDM (Scheubrein *et al.*, 2004).

5. A first decision model

5.1. *Preparation of the grid*

During the AMANDA V2 experiments, all data have been recorded. It is thus possible to replay what the controllers have made. Data will be used in order to identify some interesting cases which will be used during interviews. Interesting cases means: situations which are typical on the sector (here Bordeaux) or which have elicited different responses or actions by the controllers. A first grid is presented in table 2. This grid concerns the axis 1 of the decisional making process.

BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884
BCS1080	AFR1657 BCS1080	AFR1657 AEL2789	AFR1657	

Table 2. *Repertory grid proposed to the subject*

These elements (table 2) are composed by the different clusters proposed by controllers during AMANDA V2 experiments and like the MCDM suggested, it is possible to complete these elements with hypothetical (fictitious) situations (that the controllers did not chosen). These hypothetical situations can be "real" (potentially executable by the controller) or "unreal" (impossible choice by controllers) in order to bring a maximum of reaction. This notion of "real fictitious" and "unreal fictitious" is a concept important to the MCDM (Roy 96).

A radar view is proposed in figure 3 to better understand the situation. The first row of the table 2 corresponds to the real conflict. That means that the BAL632 and the KLM1884 pass on the beacon “VELIN” with less of one minute. These two aircraft are not separated by the minimal distance of separation (5NM). The second row of this table corresponds to the additional aircraft that the controllers of Bordeaux have choice to take into account to resolve the conflict. These aircraft are “interfering aircraft”, they are not directly in conflict with the BAL632 or KLM1884, but it is necessary to take these aircraft into account to resolve the conflict.

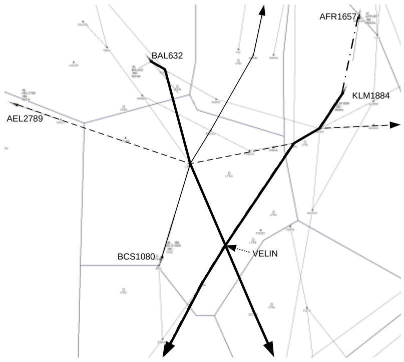


Figure 3. *Schematic view of the conflict.*

5.2. *Application of repertory grid - Interviews*

This first grid is proposed to a « decision engineer », who is one of the designers of the platform and of the experimental situation. He knows relatively well the sector and its configuration, and has a good expertise of the job of controllers.

The first key point is the variety of responses from different pairs of controllers during the experiments. All controllers have detected the conflicts and included the BAL632 and the KLM1884 in the cluster, but what would have mean if these two aircraft are not in the cluster? The likely answer is that in this case the aircraft had been reroute upstream (previous sector), and thus the conflict did not exist. This manoeuvre involves coordination between the two sectors.

The second construct proposed by the subject concerns the presence of the AFR1657. For the subject, the AFR1657 is therefore essential in the conflict because it constrains strongly the BAL632 trajectory. For him the fact that it is not here in some clusters does not mean that controllers do not take into account of this aircraft, but they exclude certain resolutions which can be problematic.

The third constructs proposed is the case of BCS1080, which is an unstable aircraft (changing Flight Level). In other words, it will necessarily come to cross flight level 350, where the conflict is situated. In the same way that AFR1657, BCS1080 will therefore constrain the future trajectory of BAL632. Half of the pairs have added this aircraft, and have decided to climb very early so that it does not interfere with the trajectory of BAL632 (with the exception that the pilot acts rapidly). This requires anticipation for instruction on BCS1080, maybe even upstream of the sector, and therefore coordination. The other pairs felt that the aircraft was not a problem, because it has enough time to climb, and does not interfere with the trajectory of BAL632.

Similarities	BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884	BAL632 KLM1884	Contrasts
	BCS1080	AFR1657 BCS1080	AFR1657 AEL2789	AFR1657		
BAL-KLM Basis conflict	1	1	1	1	1	Absence of BAL-KLM
AFR constraint BAL632	5	1	1	1	5	AFR take into account, not included
BCS constraint BAL	1	1	5	5	5	BCS1080 is another conflict.
AEL in conflict if action on AFR	5	5	1	5	5	AEL no problem

Table 3. *Result of first repertory grid*

Finally the fourth construct proposed is the case of AEL2789. The AEL2789 only included in a single cluster. It is true that its involvement in the conflict is not obvious because it is relatively far from the aircraft (BAL632 and KLM 1884). However, it must be taken into account if the controller chose to reroute the AFR1657 to put behind the BAL632. In this case the AFR1657 will be closer to the AEL2789 and it is necessary to supervise the distance between this two aircraft. For other pairs, for which a deviation on the AFR1657 was not envisaged, the AEL2789 was not a problem. (Table 3 shows the result of this first grid).

5.3. *Analysis of this first grid*

The first point which appears and which is important is the fact that controllers already have a fairly accurate idea of how they will resolve a conflict even before creating the cluster. This "knowledge" of the solution is very decisive for the choice of aircraft to be taken into account. Depending on the strategy already established, the controllers therefore choose only the aircraft that will be a problem in the application of their strategy for resolving the conflict. This is therefore clearly a link between the axis 2 and the axis 1 of the decision-making process (Figure 1).

The second issue concerns the unstable aircraft (here BCS1080). The controllers do not control well these aircraft, and their trajectories. So an important criterion will be the anticipation. Anticipate an unstable aircraft can make possible that this aircraft will be on its new level before it crosses the initial problem (2 pairs of controllers do it). But this anticipation takes time, a reasonable workload and implies generally coordinations. This can become difficult with the increase of traffic.

6. Conclusions

This paper begins with an introduction of Air Traffic Control and presents the problematic, which is the increase in air traffic, and limits of operators who will be unable to maintain an optimum level of safety.

The second part presents the platform AMANDA developed in the laboratory, which has for objectives to help controllers in their tasks, only on one controlling position for the moment. The platform is composed of two main tools: A module for trajectories calculating, as well as delegation of tasks (STAR), and a space of cooperation between the controllers and the tools, called Common Work Space (CWS). Thanks to these tools, the controllers can cooperate more efficiently, and discharge a portion of the activity (the calculation and application of trajectories) to manage new aircraft. These tools have been tested with professional controllers and have obtained encouraging results. These results lead to the new version, AMANDA V3. The objectives of this new version are presented at the end of the second part, and they concerned particularly the introduction of adjacent sectors.

The third part concerns the establishment of an approach to model the new tools of AMANDA V3. This approach is divided into three main points; the first is to model the decision-making process of controllers. The second point is a presentation of the MCDM methodology. This MCDM will guide the study. And the last point, concerns the repertory grid methodology, which will serve of operational support to the MCDM, and will support the interviews and the knowledge acquisition. Finally a last section presents an initial decision model, and an application of the approach to one axis of our decision-making process.

The continuation of this study will be to achieve other grid on three axes, in order to obtain as much information as possible, and to determine a set of criteria and a model of preference.

Acknowledgments

We want to thank the DGAC/SDER/DTI for their financial support. We also thank all the personnel of the ATC, which by its availability and its collaboration made it possible to develop, and to evaluate the various platforms. This work is supported in part by the Nord Pas-de-Calais region and FEDER (European Funds for Regional Development) through the project AUTORIS T31.

References

Annebicque D., Debernard S., Poulain T., Crévits I., "AMANDA V3: Toward a Common Work Space between air Traffic Controllers", In *ACHI 08*, Sainte Luce, Martinique, February, 2008.

Bentley R., Rodden T., Sawyer P., Sommerville I., "An architecture for tailoring cooperative multi-user displays", in *Conference proceedings of 4th Computer – supported cooperative work*, Toronto, Canada, pp. 187-194, 1992.

Crévits I., Debernard S., Vanderhaegen F., Millot P., "Multi level cooperation in air traffic control", *Fourth International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace*, Toulouse, France, September, 1993.

Debernard S., Cathelain S., Crévits I., Poulain P, "AMANDA Project: Delegation of tasks in the air-traffic control domain", *Cooperative Systems design*, IOS PRES, pp. 173-190, January, 2002.

DGAC, Direction générale de l'aviation civile, contrôle des routes du ciel. http://www.aviation-civile.gouv.fr/html/actu_gd/routciel.htm.

Endsley M.R., "Automation and situation awareness", *Automation and human performance: Theory and applications*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 163-181, 1996.

Endsley M.R., Kaber D.B., "Level of automation, effects on performance", *situation awareness and workload in a dynamic control task. Ergonomics*, 42(3), pp. 462-492, 1999.

Guiost B., Debernard S., Poulain T., Millot P., "Supporting Air-Traffic Controllers By Delegating Tasks", In *IEEE-SMC 2004*, The Hague, The Netherlands, pp 164-169, September, 2004.

Guiost B., Debernard S., "Common Work Space or How to Support Cooperative Activities Between Human Operators and Machine: Application to Air Traffic Control", *12th International Conference on Human-Computer Interaction*, 13, Springer, Beijing, P.R. China, July, 2007.

Kelly G.A., *The psychology of personal constructs*, New York: Norton, 1955.

Lemoine M.-P., Debernard S., Crévits I., Millot P., "Cooperation between Humans and Machines: First results of an Experiment with a Multi-Level Cooperative Organisation in Air Traffic Control", *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, 5, Kluwer Academic Publishers, The Netherlands, pp 299-321, 1996.

Pacaux-Lemoine M.-P., Debernard S., « A common work space to support the Air Traffic Control", *Control Engineering Practice*, A journal of IFAC, 10, pp 571-576, 2002.

Roy B., *Multicriteria Methodology for Decision Aiding*, Kluwer. London, 1996.

Scheubrein R., Zionts S., "A problem structuring front end for a multiple criteria decision support system", *Computers & Operations Research*, 33, pp.18-31, 2004.

“Data Mining for HCP” session

02:00 – 03:30 pm

Tuesday, June 10, 2008

Towards a human eye behaviour model by applying Data Mining Techniques on Gaze Information from IEC

Denis Pallez* – Laurent Brisson –Thierry Baccino*****

** I3S Lab (University of Nice - Sophia Antipolis - CNRS)*
930, route des Colles BP. 145
F-06903 Sophia Antipolis, FRANCE

E – Mail : denis.pallez@unice.fr

*** Institut TELECOM, TELECOM Bretagne*
CNRS FRE 3167 LAB-STICC,
Technopôle Brest-Iroise, FRANCE

E – Mail : laurent.brisson@enst-bretagne.fr

**** LPEQ Lab (University of Nice - Sophia Antipolis – EA 1189)*
24, Av des Diables Bleus
F-06357 Nice Cedex 4 – FRANCE

E – Mail : thierry.baccino@unice.fr

ABSRACT. In this paper, we firstly present what is interactive evolutionary computation (IEC) and rapidly how we have combined this artificial intelligence technique with an eye-tracker for visual optimization. Next, in order to correctly parameterize our application, we present results from applying data mining techniques on gaze information coming from experiments conducted on about 80 human individuals.

RÉSUMÉ. Dans cet article, nous présentons tout d'abord ce que représentent les algorithmes évolutionnaires interactifs et ensuite comment nous avons combiné cette technique d'intelligence artificielle avec un oculomètre pour de l'optimisation visuelle. Ensuite, dans le but de paramétrer correctement notre application informatique, nous présentons les résultats que nous avons obtenus en appliquant des techniques de fouilles de données sur les informations oculaires provenant d'une expérience conduite sur environ 80 sujets humain.

KEY WORDS : interactive evolutionary computation, data mining, eye-tracker.

MOT-CLÉS : algorithmes évolutionnaires interactifs, fouille de données, oculomètre.

1. Introduction

(Jaimes, Gatica-Perez et al. 2007) presents Human-Centered Computing (HCC) as an emergent field resulting from a convergence of multiple disciplines such as computer science, sociology, psychology, cognitive science… All these disciplines are concerned with understanding humans and with the design of computational devices and interfaces. However, this field is not just about the interaction, the interface or the design process but it is concerned with knowledge, people, technology and everything that ties them together. As said by the authors, Interactive Evolutionary Computation (IEC) more recently known as *human computation* (Ahn, Ginosar et al. 2006) and presented in section 2 is in the scope of HCC in the sense that humans have a central position. "Although HCC and human computation approach computing from two different perspectives, they both try to maximize the synergy between human abilities and computing resources. Work in human computation can therefore be of significant importance to HCC. On one hand, data collected through human computation systems can be valuable for developing machine-learning models. On the other hand, it can help us to better understand human behaviour and abilities, again of direct use in HCC algorithm development and system design."

In this article, we address the problem of understanding human behaviour in a particular context: the human being has to optimize a certain problem; ocular behaviour is collected with the help of an eye-tracker and analyzed with data mining techniques. We present results from this studies.

Firstly, we present in the next section interactive evolutionary computation and related works. In section 3, we present our E-TEA algorithm and the Java application that have already been proposed in order to minimize user's fatigue during interactive evaluation. In section 4, we present an experiment that we have conducted with our application in order to understand behavior of human eyes movement. Then, in section 5, results and behavior models obtained by applying data mining techniques on gaze information tracked during experiments are presented and discussed (section 6). Finally, we finish by presenting future experiments and some future works.

2. What is IEC?

Interactive Evolutionary Computation (IEC) is an optimization technique based on evolutionary computation such as genetic algorithm, genetic programming, evolution strategy, or evolutionary programming. Evolutionary computation consider several candidate solutions to a problem called the population. Thanks to an iterative progress, this population is computationally evolved by using mechanisms inspired by biological evolution such as reproduction, mutation, recombination, natural selection or survival of the fittest (Wikipedia 2007) according to the Darwin's theory. In classical evolutionary computation, a selection operator is often a program or a mathematical expression called the *fitness function* that expresses the quality of a candidate solution. So, Interactive Evolutionary Computation is used when it is hard or impossible to formalize efficiently this function where it is therefore replaced by a human user. A large survey of more than 250 papers can be obtained in (Takagi

2001), but the generally accepted first work on IEC is Dawkins (Dawkins 1986), who studied the evolution of creatures called "biomorphs" by selecting them manually. A very good example to better understand the interest of IEC could be photofit building (Takagi and Kishi 1999). In that case, there is no mathematical function which could specify how much a photofit is interesting; only the witness can subjectively tell whether proposed photofits are similar or not to the person he had seen before.

Subsequently, much work was done in the area of computer graphics: for instance using IEC for optimizing lighting conditions for a given impression (Aoki and Takagi 1997), applied to fashion design (Kim and Cho 2000), or transforming drawing sketches into 3D models represented by superquadric functions and implicit surfaces, and evolving them by using divergence operators (bending, twisting, shearing, tapering) to modify the input drawing in order to converge to more satisfactory 3D pieces (Nishino, Takagi et al. 2002). We can also mention work in combining human interactions with an artificial ant, applied to non-photorealistic rendering (Semet, O'Reilly et al. 2004). Another use of IEC involves a human patient using a PDA on which an IEC is launched to define best parameter values for cochlear implants (Bourgeois-Republique, Valigiani et al. 2005). First results show that patients using PDAs obtain a better parameterization than previously through lengthy interaction with a doctor. Following the same idea of using other human senses for human interaction, we can also mention the optimization of coffee blends (Herdy 1997) by using evolution strategies.

As mentioned before, IEC is used when a fitness function is difficult and sometimes impossible to formalize. Human-Based Genetic Algorithms (HBGA) go further by allowing evolutionary computation where a good representation of individuals is hard or impossible to find (Cheng and Kosorukoff 2004), for instance they can be used in storytelling or in development of marketing slogans. To prove the usefulness of such techniques, the authors changed the classical One-Max optimization problem into an interactive one by interpreting the individuals (strings of bits – 0 or 1) as colors to be interactively presented and manipulated.

Characteristics of IEC are *inconsistencies* of individuals fitness values given by the user, *slowness* of the evolutionary computation due to the interactivity, and *fatigue* of the user due to the obligation to evaluate manually all the individuals of each generation (Takagi 2001; Semet 2002). For instance, the user is often asked to give a mark to each individual or to select the most promising individuals: it still requires active time consuming participation during the interaction. The number of individuals of a classical IEC is about 20 (the maximum that can be represented on the screen), and about the same for the number of generations.

However, some tricks are used to overcome those limits, e.g., trying to accelerate the convergence of IEC by showing the fitness landscape mapped in 2D or 3D, and by asking the user to determine where the IEC should search for a better optimum (Hayashida and Takagi 2002). Other work tries to predict fitness values of new individuals based on previous subjective evaluation. This can be done either by constructing and approaching the subjective fitness function of the user by using genetic programming (Costelloe and Ryan 2004) or neural networks, or also with Support Vector Machine (Llorà and Sastry 2005; Llorà, Sastry et al. 2006). In the latter case, inconsistent responses can also be detected thanks to graph based modeling.

Nonetheless, previous work is mostly algorithmic-oriented and not really user-oriented, which seems to be the future domain for IEC (Takagi 2001; Parmee 2007). That's why we have presented in (Pallez, Collard et al. 2007) a new technique, totally domain independent called E-TEA (Eye-Tracking Evolutionary Algorithm), to minimize this fatigue by combining an IEC and an untraditional input device. This device allows capturing user's gaze (where the user is looking on a monitor). This is possible by using eye-tracking systems such as Tobii™ which are totally non-intrusive for users. Thus, we ensure there is no need for explicit user action (choosing and clicking the most promising individual, valuating all the solutions etc.) during the evaluation process of the IEC; he just has to watch various solutions on the screen and to tell when he has finished evaluating/looking. The E-TEA algorithm then has to determine automatically which solution is better amongst presented solutions by combining gaze parameters obtained by a Tobii™. This is the work we address, applying data mining techniques on data collected during an experiment we have conducted (cf. §4).

3. The Eye-Tracking Evolutionary Algorithm (E-Tea)

3.1 What is an eye-tracking system?

An eye-tracking system consists of following the eye's motions while a user watches a screen on which something is presented. It pinpoints in real time the position where the eye is looking, with the help of one video camera focusing on a reflected infrared ray sent to the user's cornea. This device coupled with a computer regularly samples the space position of the eye and the pupil diameter. This latter parameter lets us know the cognitive intensity of the user: the more the user is concentrated on looking at something, the smaller the diameter is (Just and Carpenter 1993). Nowadays, eye-tracking systems are very useful because they can analyze in real time what a user is focused on without any effort and in a completely non-restrictive manner. In fact, the user does not know he is being observed by a machine. With such equipment, one can finally capture when, how much time, and with which cognitive intensity a screen area is looked at.

3.2 How to use an eye-tracker in IEC for minimizing user's fatigue?

3.2.1 The E-TEA Algorithm

A new evolutionary algorithm called Eye-Tracking Evolutionary Algorithm (E-TEA) has been proposed in (Pallez, Collard et al. 2007). It is based on a classical evolutionary algorithm eventually using breeding, selection, mutation and so on in order to evolve computationally a population of candidate solutions to a problem:

1. generate initial population randomly;
2. present the population to the user;
3. let the user watch the candidate solutions;

4. compute how much time, how many times and with which cognitive intensity the presented solutions are looked at thanks to an eye-tracker;
5. combine previously obtained parameters and compute a fitness value or a rank for each solution;
6. select the most promising solutions thanks to the computed fitness value or rank
7. make crossover and mutation
8. return to step 2 until no further good solutions are found

Thus, the user just has to watch the screen and says when he has finished watching/evaluating. There is no need for the user to mark each solution, nor to explicitly choose the best or the most promising one. This will save considerable time and the user will be capable to evaluate more solutions; consequently there will be more evaluated generations. We estimate we can double the number of evaluated screens. The main difficulty is to determine how to combine different parameters captured by the eye-tracker (step 5 of the algorithm) in order to define a computable fitness or rank.

The presented algorithm is not domain-dependent. However, we need to choose a specific domain to make experiments.

3.2.2 Application to the Interactive One-Max Optimization Problem

Our optimization problem is borrowed from (Cheng and Kosorukoff 2004) where the One-Max problem is considered as an interactive optimization problem in order to compare Interactive Genetic Algorithm (IGA) and Human-Based Genetic Algorithm (HBGA). Recall that the classical One-Max optimization problem consists in maximizing the number of 1s in a string of bits (0 or 1) only in using evolving operators (selection, mutation, crossover…). It is the simplest optimization problem and it is used here in order to parameterize our system. Basically, it consists in choosing the clearest color amongst presented colors on a screen.

3.2.3 Our application

We developed an application in Java 1.6 based on the Evolutionary Computation in Java library (ECJ)[1]. Solutions are represented by a string of 24 bits, 8 bits each for red, green and blue. As we capture eye motion, the screen presents only 8 zones (one solution per zone) and no individual in the center of the screen. We avoid presenting solutions in the center because eyes are naturally attracted to the center. Also, if the user wants to compare two solutions that are diametrically opposite, eyes are obliged to cross the center. Consequently, the number of transitions for the center will increase considerably and will disrupt the estimated fitness of the solution which could be in the center. Moreover, when the application is launched, we present a screen composed of a cross in the center in order to captivate the user's gaze in the center where no candidate solutions will be presented. When the gaze is concentrated on the cross, the next screen composed of colors is presented. But, just before this screen of colors is presented to the user, a reference's value of the pupil diameter is computed and stored.

[1] http://www.cs.gmu.edu/~eclab/projects/ecj/

When the user estimates he has finished watching solutions, we ask him to press the keyboard's space bar. When done, we detect whether the user was watching a solution. If it is the case, the solution is marked as "selected".

The issue in this algorithm is to compute either fitness or rank value for each color (that is to say for each candidate solution) from gaze information. That is why we have conducted an experiment from which we tend to determine user's behaviors.

4. Experiments

4.1 Conditions

First of all, we have previously presented that the main goal of our research is to combine an evolutionary algorithm with an eye-tracker with the aim that an end-user and a computer collaboratively and rapidly converge to a solution satisfying the user. In fact, the user visually evaluates solutions of a problem and the computer tries to interpret user's interest for each solution. Next, the computer has to produce new solutions taking into account previous evaluation results; this is the task of evolutionary computation by using crossover, mutation and so on… By this way, we hope the new solutions will be better than previous ones. However, in this article we chose to randomly design and to randomly present solutions in order to have a better sampling of the solutions space.

When a new subject (experimenter) wants to participate, we ask him to read the following instructions: "The experiment is made up of a set of tries. Each try will proceed in two phases (Phase 1 and Phase 2). The experiment begins by the calibration of the device (the eye-tracker). All over the experiment, we recommend not to move the head. During the calibration, a blue circle is presented; fix it. Phase 1 named 'cross fixation': A white cross is presented in the center of the screen. Fix this cross to go next screen (when correctly fixed a red rectangle will surround the cross). Phase 2 named 'evaluation': Several colored squares will be presented simultaneously. Detect *color* that seems to be *lighter*. Once you think you have finished, press the space bar without looking at it to go next screen (next try)."

It is important to know that colors are randomly selected and set on the screen (two colors may be exactly the same or visually the same). The experiment is finished when the subject is tired (no constraint was given to subjects). Moreover, there are no dependencies between 2 screens: in fact, the evolutionary computation algorithm is not used. Nevertheless, it will be used in the future when the application will be correctly parameterized. In the following, we present the different data used in this experiment.

4.2 Data

4.2.1 Raw data coming from the eye-tracker

Data obtained from the eye-tracker (Tobii™ 1750) each 20 millisecond for each eye are the following:

- Timestamps of data in seconds and milliseconds;
- Eye position (x and y) related the current calibration;
- Eye position (x and y);

- Distance between eye and camera of the eye-tracker;
- Pupil size in millimeter;
- Validity of eye: that is whether the eye was capture or not by the eye-tracker.

In order to simplify, we only consider the gaze position represented by center of gravity of both eyes and computed from eyes positions.

4.2.2 Computed data

When the subject pressed the space bar indicating that he had finished visual evaluation, our application computes and store some data in files before showing next screen. Raw data are filtered in order to delete gaze positions that are called "jerk".

4.2.2.1 Fixations

As the eye-tracker capture eyes position each 20 milliseconds, we need to extract some semantic information from this gaze information: what is interesting for the subject? To answer this question, we need to compute *fixations*; that is to say: what did the subject fix during movement of his eyes? According to psychologists, a fixation last between 100 and 300 milliseconds. So fixations are computed from filtered raw data. For each fixation computed, we know the following:

- Coordinates (x,y) of subject's gaze;
- Duration in microsecond;
- Colored square corresponding to the fixation. If no colored square is attached to a fixation, it is not consider as a fixation.

In raw data, the eye-tracker has given the pupil diameter and we know that it is correlated with the subject's concentration; however, we do not know how. That is why we have computed several data relating to this pupil diameter. As a fixation last at least 100 ms, a fixation is made of 5 measures at least; and we know for each of them the size of the pupil diameter. So, the following data are stored for each fixation related to the size of the pupil diameter:

- The mean;
- The size at the beginning and at the end of the fixation;
- The value of the reference pupil that corresponds of the pupil diameter when focusing on the white cross and just before presenting the colored squares (cf. Phase1 and Phase2);
- The maximum variation of the size;
- The sum of variation of the size.

Once all fixations are computed from filtered raw data, new data for each candidate solution (colored square) are computed. Unfortunately, fixations were not stored in files. In the future, if we need information related to fixations, we have to compute them again.

4.2.2.2 Stored data

There are several fixations for one candidate solution. Thus, we have to compute new data from fixations for each colored square (screen region or candidate solution). So, data that we had really stored are the following:

- subject's number that have participated to the experience;

- screen number evaluated by the subject;
- elements of the color model (in our case, it is Red, Green and Blue values);
- The number (called Trans) of transition towards the square region representing the color;
- The rank (TransRank) of the previous value compared with the other values of the screen;
- The sum of transition's number (TransPop) for all the candidate solutions of the screen;
- The relative transition's number (TransNorm = Trans/TransPop);
- The time (Time) the user has focused on a colored square;
- The rank (TimeRank) of the Time value compared with the other values of the screen;
- The sum of focused time (TimePop) for all the candidate solutions of the screen;
- The mean of the pupil diameter (MeanDP) and its relative rank (MeanDPRank);
- The relative time focused on screen (TimeNorm = Time/TimePop);
- The reference value (RefDP) of the pupil diameter;
- The cognitive pupil diameter (CognitiveDP = MeanDP-RefDP);
- The maximum variation of the pupil diameter and its relative rank (MaxDPVarRank);
- The sum of variation of the pupil diameter and its relative rank (SumDPVarRank);
- A Boolean value (Selected) representing whether the color has been fixed / "selected" just before going to next screen;
- 3 objective distances (M1, M2, MS) and their relative rank;
- Positions of candidate solutions on the screen (between 0 and 1): (x,y) of upper left corner and (x,y) of bottom right corner.

Three distances for an objective fitness have been proposed in (Cheng and Kosorukoff 2004):

[1] $M_1(R,G,B) = R + G + B$

[2] $M_2(R,G,B) = 255\times\sqrt{3} - \sqrt{(255-R)^2 + (255-G)^2 + (255-B)^2}$

[3] $M_S(R,G,B) = \min(R,G,B)$

However, the first distance has been replaced by another one that better respect the human being color model:

[4] $M_1(R,G,B) = 0.299R + 0.587G + 0.114B$

5. Results

During one week and a half, 81 subjects have evaluated 7350 screens composed each by 8 colored squares. In this section we present four steps of the data mining process: analysis of physiological and behavioral data, data preparation, modeling and models evaluation.

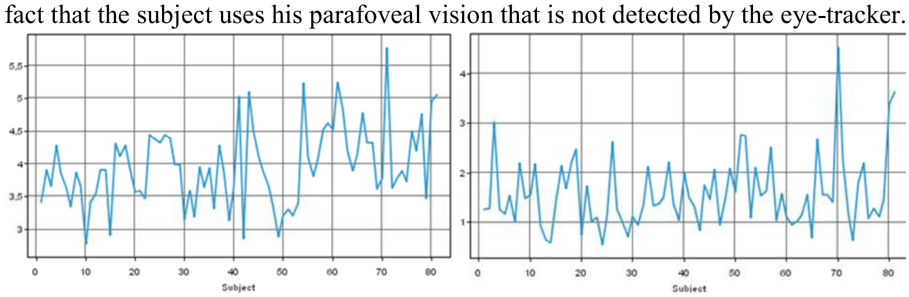
5.1 Data analysis

Three data are important in order to build a predictive model to classify solutions presented on a screen: pupil diameter, time used on colored squares, and number of transitions on these squares.

Figure 1a shows the mean reference value of pupil diameter by subject. The reference value (measured before each screen) helps to determine the cognitive pupil diameter value when the user sees a colored square. The reference value depends on subject's physiology and brightness of the office where experiences take place. We can observe on that figure that the pupil diameter increases after subject number 50: it is explained by the fact that the experiment took place in two different office; the subject number 54 corresponds to a new office. We can see mean reference value of pupil diameter varies between 2.8mm and 5.8mm.

Figure 1b shows the mean time spent for observing one screen by subject. The time varies between 0.5s and 3.5s (if we don't consider some exceptions).

Figure 2 shows minimum, maximum and mean number of transitions on one colored square by subject. We can notice that, for each subject, some colored squares aren't watched. By mean, a subject sees 0 to 2 colored squares. The maximum value varies between 3 to 12 transitions towards a colored square: it is an exceptional behavior which underlies a doubt. This behavior can be explained by the fact that the subject uses his parafoveal vision that is not detected by the eye-tracker.



(a) **(b)**

Figure 1. *Reference value of pupil diameter by subject (a) and mean time used for observing one screen by subject (b)*

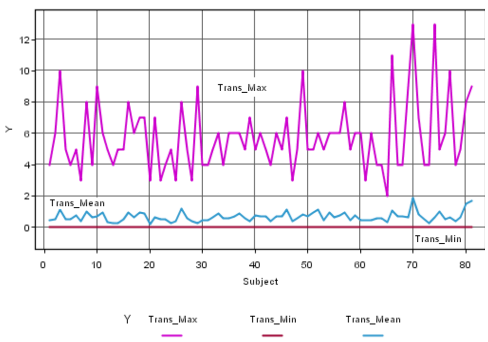
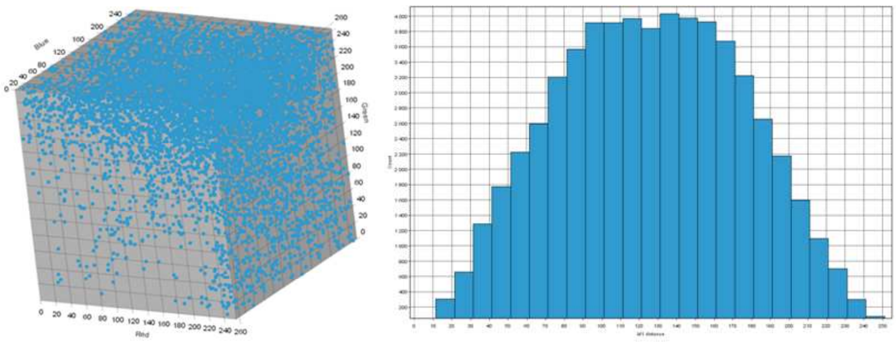


Figure 2. *Min, max and mean number of transitions on one individual by subject*



(a) (b)

Figure 3. *Distribution of selected colors in the RGB model (a) and Number of colored squares according M1 value (b).*

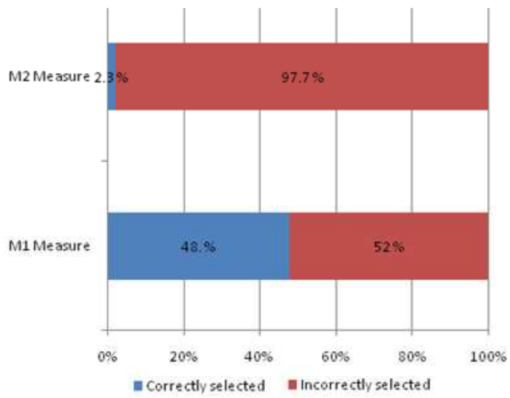


Figure 4. *Distribution of subjects choices according M1 and M2 distances*

Figure 3a presents distribution of user selected colors according the RGB model. We can notice an high density of selected colored squares which maximize R,G and B values. If we consider all of the colors displayed on the screen, there is an uniform distribution.

Figure 4 shows if the subject has correctly chosen the colored square according to M1 and M2 distances. If we consider M2, users rarely made the right choice, while if we consider M1 users made the right choice once on two. Thus, we can consider M1 is the best distance measure in order to build our model.

5.2 Data preparation

5.2.1 Discretization

M1 is the distance we chose to use in order to build our model. **Figure 3**b shows the number of colored squares according M1 value. Since the distribution isn't uniform it is necessary to discretize this distance in order to construct some sets of values whose size is equals. This operation is necessary to avoid learning biases.

We created five sets:

- Darker: M1 ∈ [1,81[
- Dark: M1 ∈ [81,112[
- Undefined: M1∈ [112,141[
- Light: M1∈[141,172[
- Lighter: M1∈ [172,251]

5.2.2 Data selection

In a first step, we create two datasets in order to try different paradigms with relative attributes or rank attributes. First set, called A, contains as predictive attributes *TransNorm, TimeNorm*, *CognitiveDP* and the second set, called B, contains as predictive attributes *TransRank, TimeRank, MeanDPRank, MaxDPVarRank, SumDPVarRank*. The target attribute is the discrete M1 value.

In a second step, we created two sets A' et B' which contains one more attribute: *Selected*. These datasets are available on our web webpage[2].

5.3 Modeling

We use C5.0 algorithm, a widely used and tested decision tree algorithm successor of C4.5. Description of this algorithm is beyond the scope of this paper and the reader should refer to (Quinlan 1993).

We created two sets each one with one square on two. The first set is the learning set and we use 15 subsets for cross validation, while the second set is a validation set for evaluating models.

We generated 4 models with the same parameters for each dataset A, A', B, B'.

5.4 Evaluation

Figure 11 shows results of the models. We notice that the best model is the one using dataset A without *Selected* attribute. Figure 12 shows coincidence matrix for model A in order explain these results. For each row, we can observe distribution of predicted values. For example, 4760 darker colored square are correctly classified however 444 of them were classified as "Ligther". We can notice the algorithm often predicts a colored square as *"Darker"* which implies lots of errors. To solve this problem we tried cost matrix without significant results.

	A	A'	B	B'
Correctly classified	28.81%	28.37%	28.02%	28.19%

Figure 5: Models evaluation

↓ Real / Predicted →	Darker	Dark	Undefined	Light	Lighter
Darker	4760	268	166	288	444
Dark	4359	343	169	348	667
Undefined	3892	320	237	401	998
Light	3235	331	212	413	1739
Lighter	2269	285	231	307	2718

Figure 6: Coincidence matrix for model A

[2] http://perso.enst-bretagne.fr/laurentbrisson/activites-recherche/osef/

6. Discussions

The Eye-Tracking Evolutionary Algorithm is a very simple but very innovative proposition that is at the intersection of two different domains: computer and cognitive sciences. This approach presents many advantages:

- First, it is the first time that an eye-tracker takes a very active part in a computer application. More traditionally, eye-tracking systems are used for analyzing human behavior when looking at an image, a text, a 3D model, a webpage, etc.
- Second, with such a combination we automate interactive evaluation of individuals with no constraints for the user. The only thing he has to do is to watch individuals and to say when he has finished. There is no explicit task imposed on the user, and thus no additional fatigue.
- Next, such material is completely non-intrusive, i.e., the user could forget that he is being observed. Interactive evaluation is as natural as possible.
- Finally, by analyzing the cognitive activity of the user, we can easily detect when the user is tired. "PERCLOS" measure is the most reliable and valid determination of a user's alertness level. PERCLOS is the percentage of eyelid closure over the pupil over time and reflects slow eyelid closures ("droops") rather than blinks. A PERCLOS drowsiness metric was established in a 1994 driving simulator study as the proportion of time in a minute that the eyes are at least 80 percent closed (Wierwille, Ellsworth et al. 1994).

Of course, each new system has its drawbacks, but they are few compared to the advantages:

- The eye-tracker can follow eyes if and only if it has been calibrated to the user. However, this takes only few seconds, and the user just has to focus on concentric moving circles.
- The other small constraint is that the user does not have total freedom of head movement. For instance, he can not look away and then resume evaluating. However, the freedom is large enough (30x16x20 cm) because of the use of two video cameras. If the signal is lost for one eye, the eye-tracker uses the other eye.

7. Conclusion and Future Works

In this article, we have presented a combination of a classical optimization technique represented by Interactive Evolutionary Computation and less classical device (an eye-tracker). Result is an innovative approach to minimize user's fatigue during interactive evaluation of proposed solutions to an optimization problem.

Before proving that this approach is better than others which use a classical device as a mouse, we need to correctly parameterize our application and understand human eye behavior by an experiment: ask people to detect the lightest color amongst 8 presented colors. During this experiment, data were stored and analyzed in order to find models of behavior for human eye movements. We have to improve resulting models in order to better know how to combine ocular data for computing either a fitness value or a rank value for each candidate solution. Once better models will be found, we'll have to prove that our E-TEA algorithm is better than others using a mouse. To do that, we need to conduct another experiment by evolving solutions with the help of evolutionary computation rather than presenting random solutions. Next, we also need to integrate in our application a machine learning module that will be able to predict fitness or rank value for each candidate solution as already mentioned in (Takagi 2001). Finally, with all these modifications, it will be interesting to test our approach in a real world application.

8. Acknowledgments

We would like to thank the Institute of Technology at Nice University (http://www.iut-nice.fr), polytech school of Nice Sophia-Antipolis (http://www.polytech.unice.fr/) which let this research work be possible and persons which participated to experiment.

9. REFERENCES

Ahn, L. v., S. Ginosar, M. Kedia, R. Liu and M. Blum (2006). "Improving Accessibility of the Web with a Computer Game." CHI.

Aoki, K. and H. Takagi (1997). 3-D CG lighting with an interactive GA. 1st International Conference Conventional and Knowledge-Based Intelligent Electronic Systems, Adelaide, Australia.

Bourgeois-Republique, C., G. Valigiani and P. Collet (2005). An Interactive Evolutionary Algorithm for Cochlear Implant Fitting: First Results. SAC05, ACM Symposium on Applied Computing. H. H. e. al. Santa Fe, New Mexico.

Cheng, C. D. and A. Kosorukoff (2004). Interactive One-Max Problem Allows to Compare the Performance of Interactive and Human-Based Genetic Algorithms. Genetic and Evolutionary Computation (GECCO), Seattle, Washington, USA.

Costelloe, D. and C. Ryan (2004). Genetic Programming for Subjective Fitness Function Identification. Genetic Programming. LNCS**:** 259-268.

Dawkins, R. (1986). The Blind Watchmaker.

Hayashida, N. and H. Takagi (2002). "Acceleration of EC convergence with landscape visualization and human intervention." Applied Soft Computing **1**(4): 245.

Herdy, M. (1997). Evolutionary optimization based on subjective selection - Evolving blends of coffee. 5th European Congress on Intelligent Techniques and soft computing

Jaimes, A., D. Gatica-Perez, N. Sebe and T. S. Huang. (2007). "Human-Centered Computing: Toward a Human Revolution." from http://www.computer.org/portal/site/computer/index.jsp?pageID=computer_level1&path=computer/homepage/May07&file=gei.xml&xsl=article.xsl.

Just, M. A. and P. A. Carpenter (1993). "The intensity dimension of thought: Pupillometric indices of sentence processing." Canadian Journal of Experimental Psychology, Special Issue: Reading and language processing **47**(2): 310-339.

Kim, H.-S. and S.-B. Cho (2000). "Application of interactive genetic algorithm to fashion design." Engineering Applications of Artificial Intelligence **13**(6): 635.

Llorà, X. and K. Sastry (2005). Combating User Fatigue in iGAs : Partial Ordering , Support Vector Machines and Synthetic Fitness. GECCO, Washington, D.C., USA.

Llorà, X., K. Sastry and F. Alias (2006). Analyzing Active Interactive Genetic Algorithms using Visual Analytics. Genetic and Evolutionary Computation (GECCO), Seattle, Washington, USA.

Nishino, H., H. Takagi, S. Saga and K. Utsumiya (2002). A virtual modeling system for intuitive 3D shape conceptualization. IEEE International Conference on Systems, Man and Cybernetics.

Pallez, D., P. Collard, T. Baccino and L. Dumercy (2007). Eye-Tracking Evolutionary Algorithm to minimize user fatigue in IEC applied to Interactive One-Max problem. GECCO (Genetic and Evolutionary Computation COnference), University College, London.

Parmee, I. (2007). "User-centric Evolutionary Computation." from http://www.ad-comtech.co.uk/Workshops.htm.

Quinlan, J. R. (1993). C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc.

Semet, Y. (2002). Interactive Evolutionary Computation : a survey of existing theory, University of Illinois.

Semet, Y., U.-M. O'Reilly and F. Durand (2004). An Interactive Artificial Ant Approach to Non-photorealistic Rendering. Genetic and Evolutionary Computation (GECCO), Seattle, Washington, USA.

Takagi, H. (2001). Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation. IEEE.

Takagi, H. and K. Kishi (1999). On-line Knowledge Embedding for an Interactive EC-based Montage System. Third International Conference on Knowledge-Based Intelligent Information Engineeing Systems, Adelaide, Australia.

Wierwille, W., L. Ellsworth, S. Wreggit, R. Fairbanks and C. Kirn (1994). "Research on vehiclebased driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness." National Highway Traffic Safety Administration Final Report: DOT HS 808 247.

Wikipedia. (2007). "Evolutionary computation." from http://en.wikipedia.org/wiki/Evolutionary_computation.

Visualization and Interactive Exploration of Factorial Correspondence Analysis Results on Images

Nguyen-Khang Pham*, ** — Annie Morin* — Patrick Gros*

** IRISA*
Campus de Beaulieu
35042 Rennes cedex

{pnguyenk;amorin:pgros}@irisa.fr

*** Cantho University*
Campus III, 1 Ly Tu Trong Street, Cantho city, Vietnam

pnkhang@cit.ctu.edu.vn

ABSTRACT. *The aim of our investigation is to interactively explore the results of Factorial Correspondence Analysis (FCA) applied on images in order to be able to extract knowledge and to better interpret the obtained results.*

We propose an interactive graphical tool, CAViz, which allows us to view and to extract knowledge from the FCA results on images. Originally, FCA deals with contingency tables and is very often used in Textual Data Analysis (TDA). In Textual Data Analysis, the contingency table crosses words and documents. For adapting FCA to images, the first step is to define the "visual" words in images (similar to words in the texts). These words are constructed from local descriptors (SIFT, Scale Invariant Feature Transform) in images.

CAViz projects clouds of points in factorial planes and allows to view and to extract valuable information such as characterizing words and relevant indicators (representation quality and contribution to the inertia). An application to the Caltech4 base demonstrates the interest of CAViz for the analysis of FCA results.

KEY WORDS: *Factorial correspondence analysis, Visualization, SIFT*

1. Introduction

Data mining (Fayyad *et al.*, 1996) intends to extract useful hidden knowledge from the large datasets in a given application. This usefulness relates to user goal. In other words, only the user can decide whether the resulting knowledge answers his/her goal. Therefore, data mining tools should be highly interactive and user-friendly. The idea here is to increase the human involvement through interactive visualization techniques in data-mining environment.

In the past years, many visual methods have been developed in different domains and used for data exploration and knowledge extraction process (Fayyad *et al.*, 2001, Keim, 2002). Visual methods are used for data selection (pre-processing step) and results displays (post-processing step). Some recent visual data mining methods (Ankerst *et al.*, 2000, Do *et al.*, 2004, Poulet, 2004) try to involve more intensively human factors in the data mining step through visualization. The effective cooperation can bring out some advantages such as the use of domain knowledge during the model construction, the improvement of the confidence and the understanding of the obtained models, when using human pattern recognition capabilities in model exploration and construction.

Our investigation aims to explore the results of Factorial Correspondence Analysis (FCA) (Benzécri, 1973). In textual data analysis, the Bi-Qnomis tool (Kerbaol *et al.*, 2006) developed by M. Kerbaol allows to visualize results and to find relevant topics in a text corpus analyzed by a FCA. The point clouds (terms/words and documents) are projected on a factorial plane. Then for each axis, a list of words whose contributions to inertia are large (called metakeys after M. Kerbaol), generally three times the average contribution by word or/and documents, is created. FCA is based on classical results in matrix theory and the central result is a eigenvector-eigenvalue decomposition of a square matrix. Total inertia on an axis is equal to the corresponding eigenvalue; so the threshold is easy to compute. The metakeys are displayed on the left and the titles of the documents with high contributions are listed on the top of screen. We can click on the title of the document to get immediately the plain text. When examining metakeys and/or documents, an expert can summarize the content of these documents. Bi-Qnomis supports also a visualization of metakeys using a hyperbolic tree.

For adapting FCA to images, we meet some difficulties because there are no real words in images. We use “visual words” instead. Recently, some methods originally developed for textual data analysis such as PLSA (Probabilistic Latent Semantic Analysis) (Hofmann, 1999), LDA (Latent Dirichlet allocation) (Blei *et al.*, 2003) have been applied in image analysis for image classification (Willamowski *et al.*, 2004), topic discovery in the image (Sivic *et al.*, 2005), scene classification (Bosch *et al.*, 2006), and image retrieval (Lienhart *et al.*, 2007). These methods try to find a model of the corpus and to reduce the dimensions. A disadvantage of these methods is the use of an ad hoc model and of the EM algorithm to find a local optimum. In addition, it is difficult to interpret the results of these methods. Most of the works use such methods as black boxes.

This paper will focus on adapting FCA on images and on interpreting the results using relevant indicators through a visualization tool, CAViz, in which the user can interactively explore the results to better understand them. The article is organized as follows: we briefly describe the FCA method in Section 2. Section 3 presents interactive knowledge extraction from the FCA results. In conclusion, we present some perspectives for this work.

2. Adaptation of FCA on images

2.1. *FCA*

FCA is a classical exploratory method for the analysis of contingency tables. It was proposed by J. P. Benzécri (1973) in the linguistic context, *i.e.* textual data analysis. The first study was performed on the tragedies of Racine. FCA on a table crossing words and documents allows answering the following questions: is there any proximity between certain words? Is there any proximity between certain documents? Is there any link between certain words and certain documents? FCA like most factorial method uses a singular value decomposition of a particular matrix and allows viewing of words and documents in a reduced space. This reduced space has a particular propriety where points are projected (words and / or documents) with a maximum inertia. In addition, FCA provides relevant indicators for the interpretation of the axes as the contribution of a word or a document to the inertia of the axis or the representation quality of a word and/or document on an axis (Morin, 2004).

2.2. *Construction of visual words and images representation*

In order to adapt FCA on images, we must represent the image corpus in the form of contingency table. Here images are treated as documents and the “visual words” (to be defined) as terms/words.

The words in the images, called visual words, must be calculated to form a vocabulary of *N* words. Each image will be finally represented by a word histogram. The construction of visual words is processed in two steps: (i) computation of local descriptors for a set of images, (ii) classification (clustering) of obtained descriptors. For each cluster, we have a visual word. There would thus be as many words as clusters obtained at the end of step (ii). The local descriptor computation in an image is also done in two stages: we first detect the interest points in images. These points are either maxima of Laplace of Gaussian (Lindeberg, 1998), or 3D local extrema of the difference of Gaussian (Lowe, 1999), or points detected by a Hessian-Affine detector (Mikolajczyk *et al.*, 2004). Figure 1 shows some interest points detected by a Hessian – Affine detector.

Figure 1. *Interest points detected by Hessian-Affine detector*

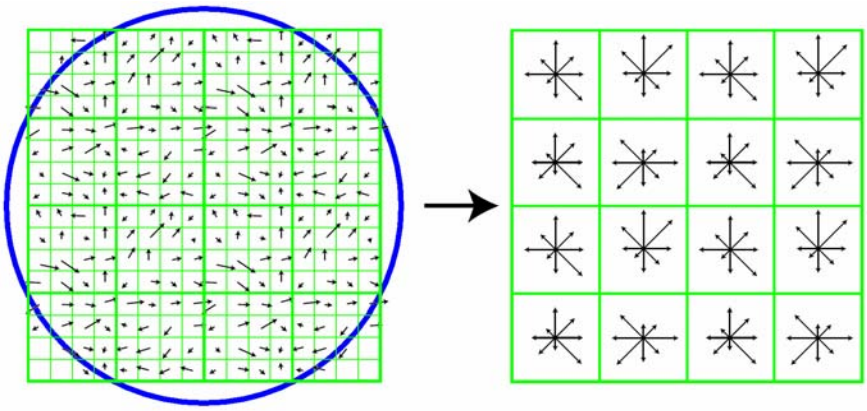


Figure 2. *A SIFT descriptor computed from the region around the interpret point (the circle): the gradient of the image (left) and the descriptor of the interest point (right)*

Then, the descriptor of the interest points is computed on the gray level gradient of the region around the point. The Scale-Invariant Feature Transform descriptor, SIFT (Lowe, 2004), is often preferred. Each SIFT descriptor is a 128-dimensional vector. Figure 2 describes a SIFT descriptor. The second step is to form visual words from the local descriptors computed in the previous step. Most of works perform a *k*-means on descriptors and take the average of each cluster as visual word (Willamowski *et al.*, 2004, Sivic *et al.*, 2005, Bosch *et al.*, 2006). Some visual words computed from Caltech4 dataset are shown in figure 3. After building the visual vocabulary, each descriptor is assigned to its nearest cluster. For this, we compute, in $\mathbf{R}^{128}$, distances from each descriptor to representatives of previously defined clusters. An image is then characterized by the frequency of its descriptors. The image corpus will be represented in the form of a contingency table crossing images and clusters.

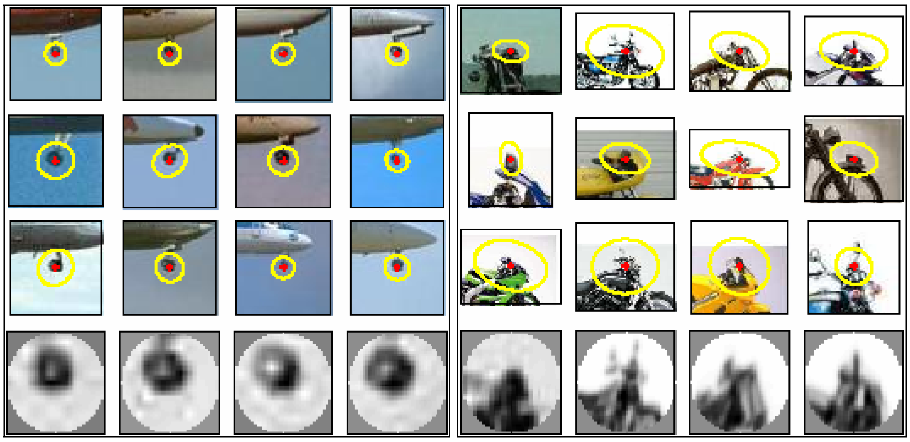


Figure 3. *Some visual words constructed from Caltech4 dataset*

Since these visual words are defined on descriptors which are computed from several categories, it is possible to have images in different categories sharing one word. We call it the *polysemy* of visual words.

3. Interactive exploration of FCA results

3.1. *Projection on factorial plane*

The screen is divided into two parts: point clouds (images and/or visual words) are drawn on the left, and the right part is reserved to display the selected images. A point "image" is displayed as a red square; and a point "word" as a blue square. The user can select an image or an *image group* (*image group* is a group of points standing nearly around each other) by pointing to the image interesting. All the images found in a neighbourhood of radius ***r*** of the interesting image will be displayed on the right of the screen. The points corresponding to the selected images will change their colour (from red to green). The visual words (in the form of an ellipse) will be also drawn on images. Selected images are shown on the right hand side. This gives us immediately a general summary of the content of these images.

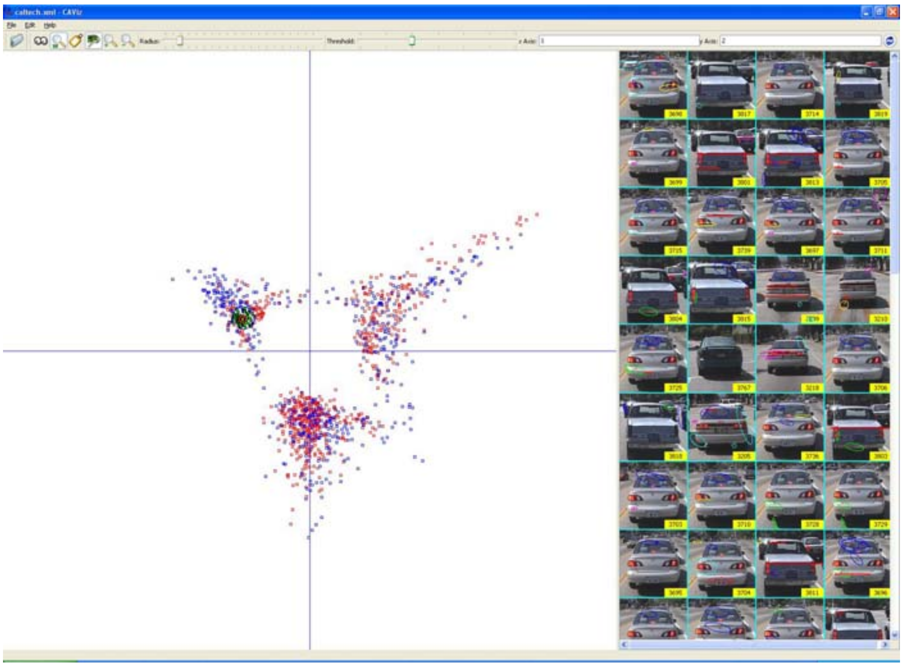


Figure 4. *Projection of Caltech4 base on axes 1 and 2*

To focus only on images and/or interesting words, we display only the images and/or words whose contribution to the inertia is high, usually 2 or 3 times the average contribution. The total inertia on one axis is equal to the eigenvalue associated to this axis. So the threshold is easy to determine. Figure 4 shows the projection of the Caltech4 base (Sivic *et al.*, 2005) on the axes 1 and 2 with the threshold equal to 2 times the average. Mr. Kerbaol calls metakey a group of words whose contribution is very high on an axis. We therefore have 2 metakeys by an axis, a positive one and a negative one. The words belonging to each metakey will be displayed on the associated image. In Figure 5, we have some metakeys and their visual words. Images located at each end of axes display corresponding metakeys. The image on the upper left is superimposed with visual words which are closed to it. It is easy to see that most of these words are found in both the left metakey and the bottom metakey.

Interesting information on an image is interactively extracted by selecting the image. There are two relevant indicators for interpreting FCA results: the quality of display on the one hand and the contribution to the inertia on the other hand. This information helps us in the following tasks. Figure 6 shows an example of the interactive information extraction. The image in this example is well represented by the axes 2 (negative), 3 (negative), 18 (negative), 12 (positive) and highly contributes to the inertia of these axes.

One of advantages of FCA is a double representation of documents and words on a same factorial plane. Words closed to an image well characterize this image. It is easy to visualize words well characterizing an image or an image group by selecting the image and displaying the words closed to this image. The figure 7 shows the words well representing the topic "visage".

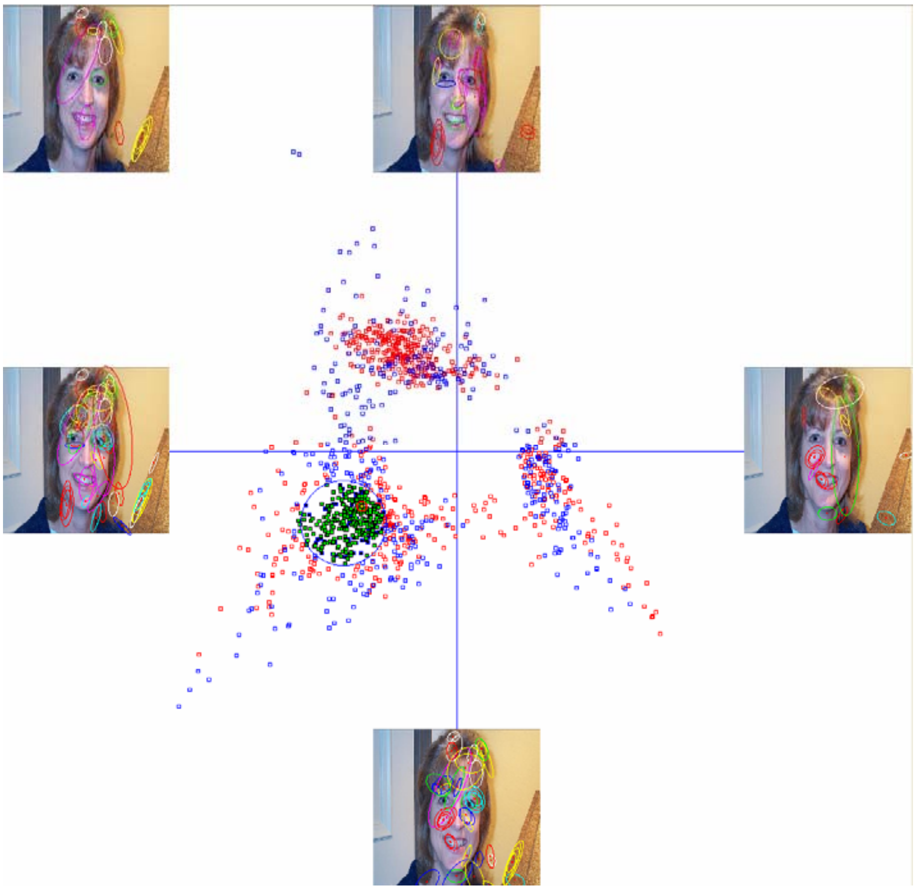


Figure 5. *Visualization of metakeys*

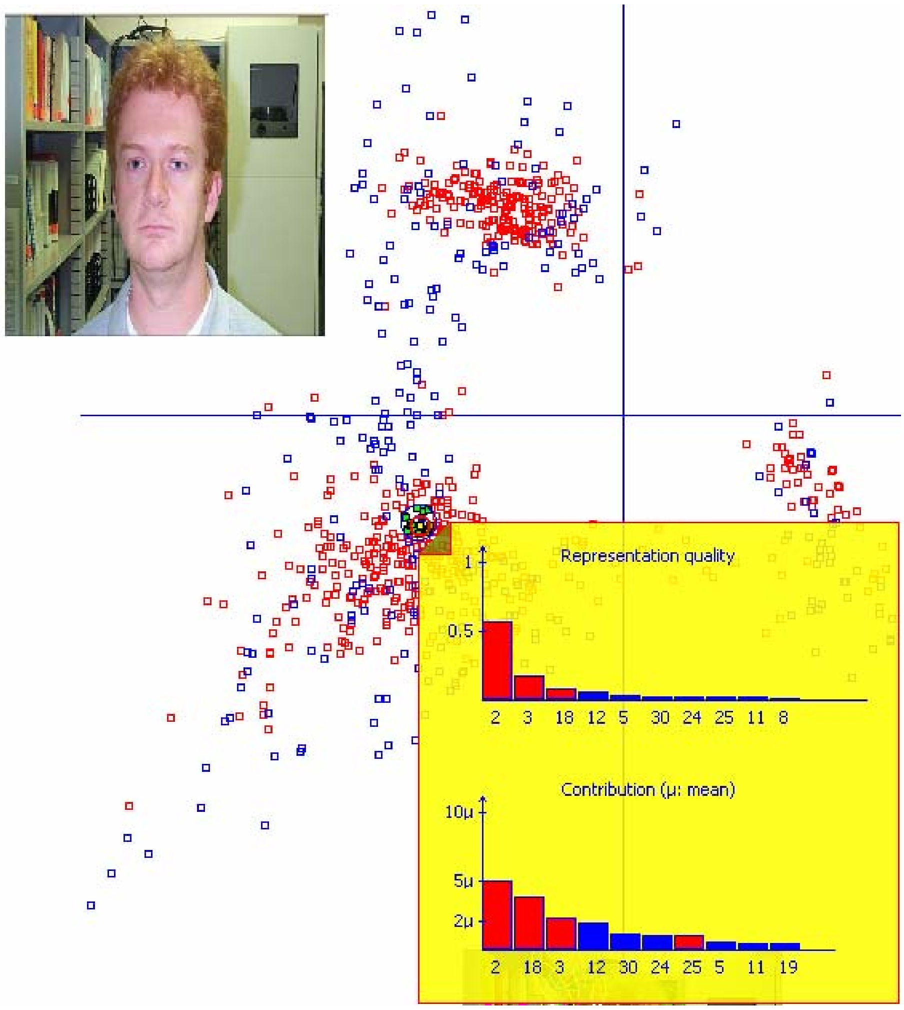


Figure 6. *Image information extraction: histogram of representation quality on axes and histogram of contribution to inertia of axes of an image (red: positive, blue: negative)*

3.2. *Image topic discovery*

After having displayed point clouds on the factorial plane, we have a look on group of points which defines a topic in this factorial plane. However, in order to discover topics, we should look for axes which well represent topics. CAViz allows to display the information about representation quality of axes and to support a dual view which facilitates viewing information about representation quality of an image group.

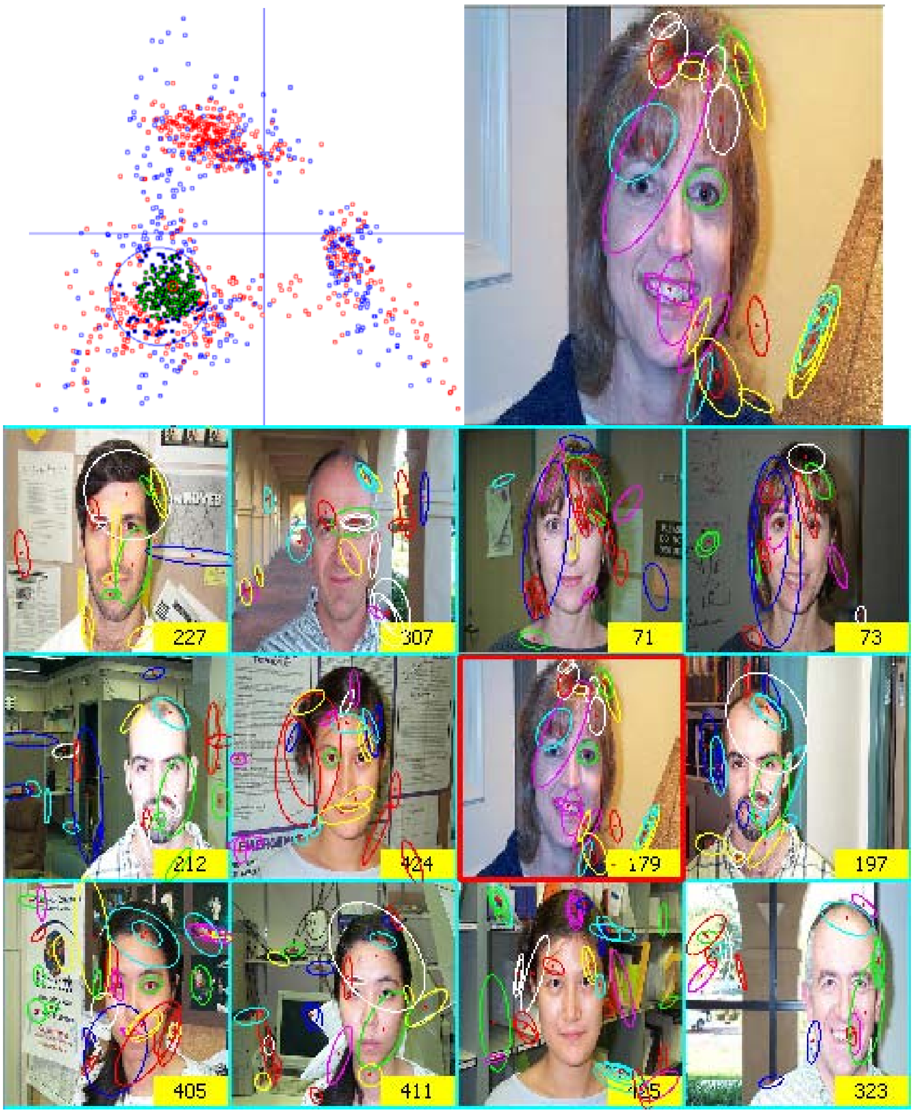


Figure 7. *The visual words characterizing the "face" topic*

3.2.1. *Representation quality of images*

The representation quality of a point ***i*** (image ***i***) on the axis ***j*** is the square cosine of the angle between the axis ***j*** and the vector which joins the centre of the point cloud to the point ***i***. The more the square cosine is close to 1, the more the position of the observed point on the projection is close to the real position of the point in the original space. We use this criterion to look for relevant axes.

The point cloud is projected on the first plane (*i.e.* axes 1 and 2). We extract then the representation quality of image on axes by looking at the first factors of the histogram. This information helps us to find the best plane for an image group.

3.2.2. *Dual views*

CAViz displays the points on the left part and selected images on the right part. When clicking on an image on the right, the corresponding point on the left is selected and whole information is showed. These dual views allow us to select easily interesting images. The human visual perception is a good tool for pattern recognition. We select the similar images on the right hand side and look at their information on the left. If we find that there are the same axes which well represent the selected images, we will take these axes and project the points on them.

3.2.3. *Image topic discovery*

We give here a case study of topic discovery in Caltech4 database (Sivic *et al.*, 2005) drawn from the Caltech101 database (Fergus *et al.*, 2003). This database contains 4090 images divided into 5 categories. Table 1 describes this database. Figure 8 shows images drawn from the Caltech4 database.

Category	Number of images
Faces	435
Motorbikes	800
Airplanes	800
Backgrounds	900
cars (rear)	1155

Table 1. *Description of the Caltech4 database*

First a FCA is applied on images. We then project point clouds onto the first factorial plane and select the group on the left (cars category). The selected images are displayed on the right part of the screen. We start to select images by clicking on the right and by looking at their representation quality on axes. We find that there are images that are well represented by the axes 6 (negative) and 7 (positive) and there are many images well represented by the axes 11 (negative) and 12 (positive) (Figures 9, 10). Projecting point clouds on these axes, we find some high-quality topics which content similar images.

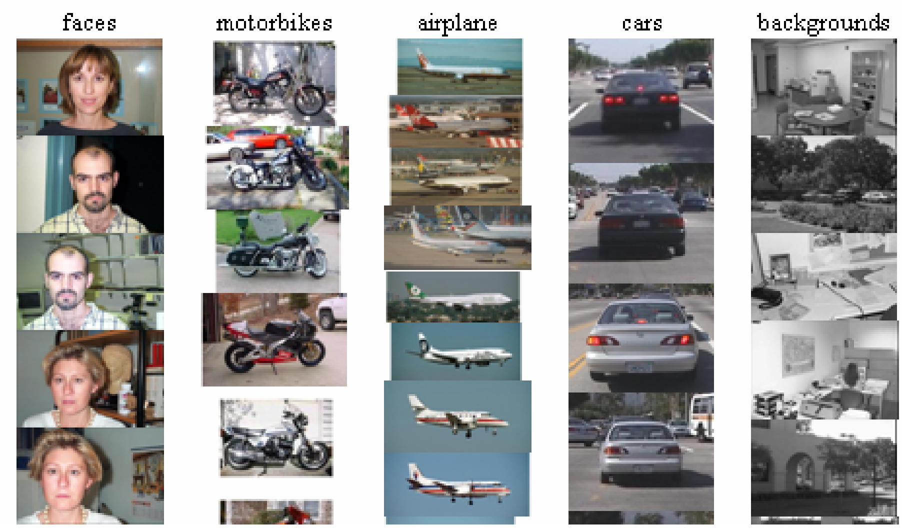


Figure 8. *Images drawn from the Caltech4 database*

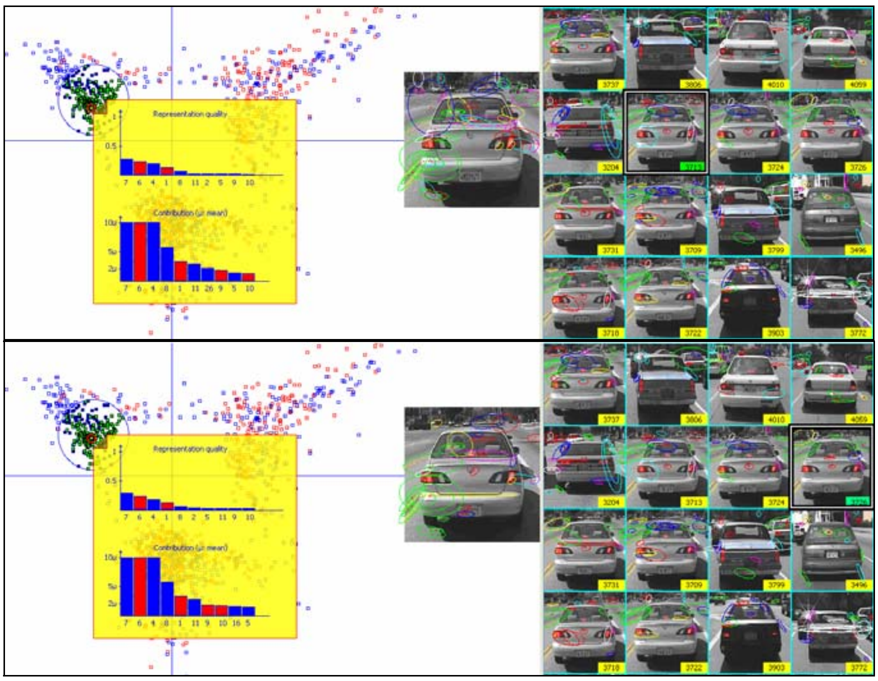


Figure 9. *Images well represented by the two axes 6 and 7*

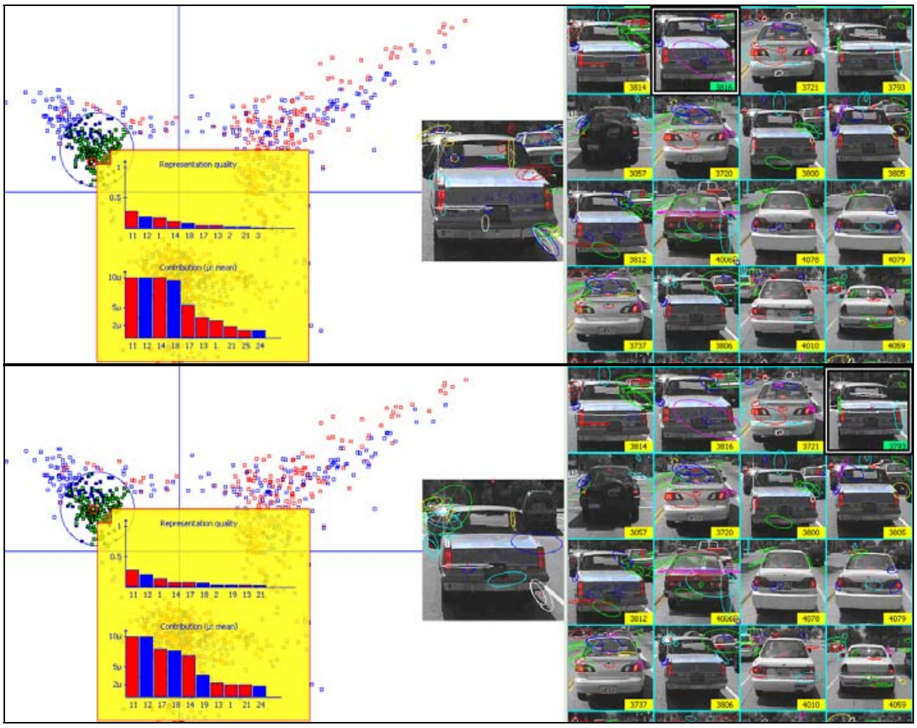


Figure 10. *Images well represented by the two axes 11 and 12*

Projecting on the axes 6 and 7, we find that there is a group of points on the top left (axis 7 - positive and axis 6 - negative). We select this group and look at the images displayed on the right part: these images are very similar (they are white cars, figure 11). In addition, there is also another group on the bottom left. We do the same thing as previously and find another topic (red cars)! Similarly, on the axes 11 and 12, there is another topic for cars (Figure 12).

4. Conclusion and future works

We have presented in this article a graphical tool, called CAViz, which visualize the results of the FCA on images. This tool helps users to extract knowledge and to interpret FCA results. We have also shown an application on the image topic discovery in the Caltech4 database. The obtained topics in this study have proved the simultaneous value of CAViz and of FCA.

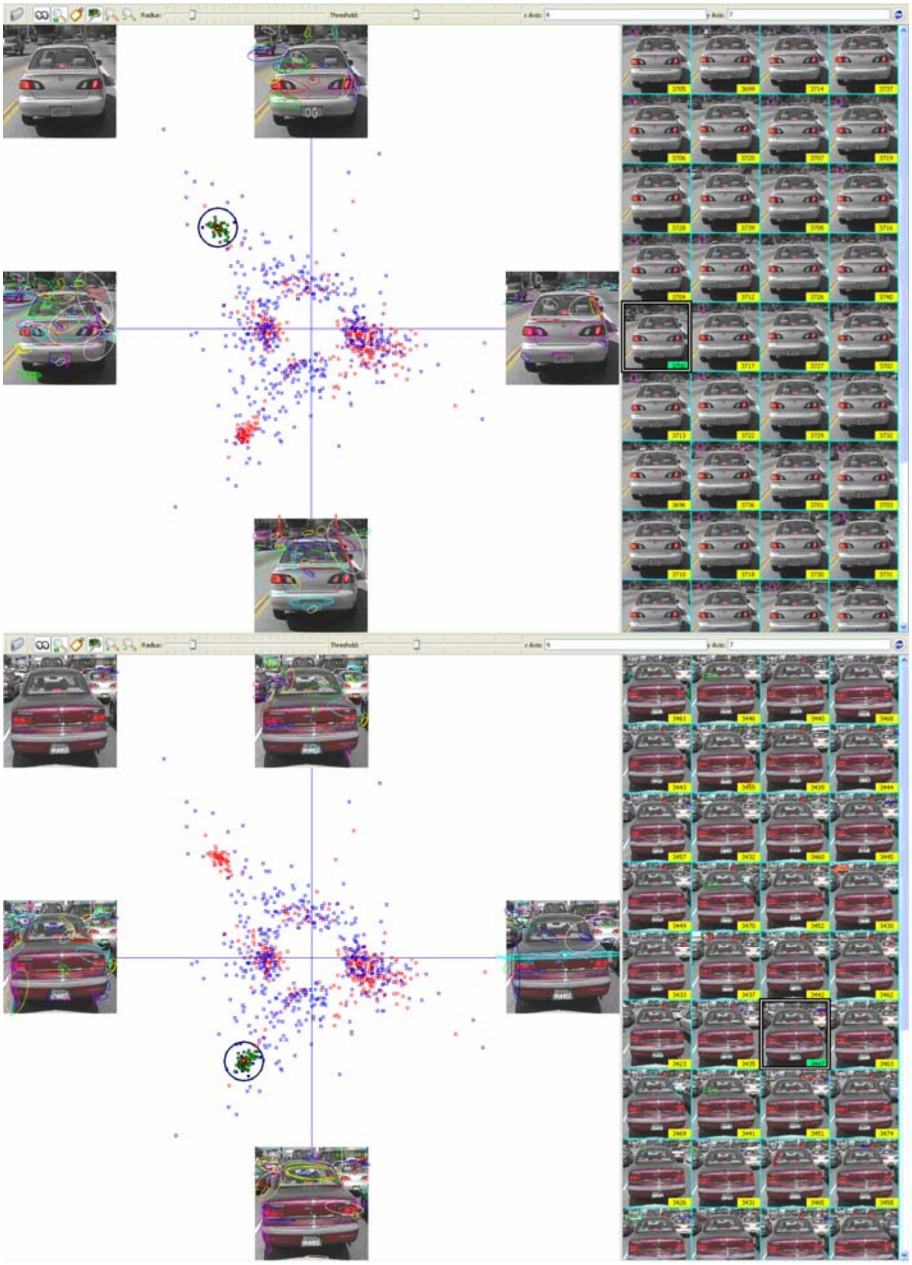


Figure 11. *Topic discovery by projecting points on the best plane: we found two different topics "cars" on the plane 6 – 7*

Figure 12. *Topic discovery by projecting points on the best plane: we found two different topics "cars" on the plane 6 – 7*

Some improvements should be useful such as a quick search of important axes for an image group and flexibility in the selection of images. For this, we can extract information from the group's gravity centre and/or use a rectangle, ellipse or an arbitrary shape instead of a circle. One of the possible developments of this method is the building of a system that would facilitate the exploration, navigation, and image labelling by integrating visual information and other information associated with the images.

Acknowledgements

The authors would like to thank A. Zisserman for Caltech4 database and F. Poulet for useful discussions.

5. Bibliography

Ankerst M., Ester M., Kriegel H-P., “Towards an effective cooperation of the computer and the user for classification”, In *Proceeding of KDD'00, 6th ACM SIGKDD*, 2000, p. 179 – 188.

Benzécri J.-P., *L'analyse des correspondances*, Paris: Dunod, 1973.

Blei, D. M., A. Y. Ng, and M. I. Jordan., “Latent dirichlet allocation”, *Journal of Machine Learning Research*, vol. 3, p. 993 – 1022, 2003.

Bosch A., Zisserman A., Munoz X., “Scene Classification via PLSA”, In *Proceedings of the European Conference on Computer Vision*, 2006, p. 517 – 530.

Do T-N., Poulet F., “Enhancing svm with visualization”, In *Discovery Science 2004*, LNAI-3245, E. Suzuki et S. Arikawa Eds., 2004, p. 183 – 194.

Fayyad U., Piatetsky-Shapiro G., Smyth P., “From data mining to knowledge discovery in databases”, *AI Magazine*, vol. 17, n° 3, 1996, p. 37 – 54.

Fayyad U., Grinstein G., Wierse A., *Information Visualization in Data Mining and Knowledge Discovery*, Morgan Kaufmann Publishers, 2001.

Fergus R., Perona P., Zisserman A., “Object Class Recognition by Unsupervised Scale-Invariant Learning”, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2003, p. 264 – 271.

Hofmann T., “Probabilistic Latent Semantic Analysis”, In *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI’99)*, 1999, p. 289 – 296.

Keim D., “Information Visualization and Visual Data Mining”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, n° 1, 2002, p. 1 – 8.

Kerbaol M., Bansard J. Y., Coatrieux J. L., “An Analysis of IEEE Publications”, *IEEE Engineering in Medicine and Biology Magazine*, vol. 25, n° 2, 2006, p. 6 – 9.

Lienhart R., M. Slaney, “PLSA on Large Scale Image Databases”, In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, 2007, p. 1217 – 1220.

Lindeberg T., “Feature detection with automatic scale selection”, *International Journal of Computer Vision*, vol. 30, n° 2, p. 79-116, 1998.

Lowe D. G., “Object Recognition from Local Scale-Invariant Features”, In *Proceedings of International Conference on Computer Vision*, 1999, p. 1150 – 1157.

Lowe D. G., “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, vol. 60, n° 2, 2004, p. 91 – 110.

Mikolajczyk K., Schmid C., “Scale and affine invariant interest point detectors”, *Proceedings of IJC V*, vol. 60, n° 1, p. 63-86, 2004.

Morin A., “Intensive Use of Correspondence Analysis for Information Retrieval”, In *Proceedings of the 26th International Conference on Information Technology Interfaces*, 2004, p. 255 – 258.

Poulet F., “SVM and Graphical Algorithms: A Cooperative Approach”, In *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM’04)*, 2004, p. 499 – 502.

Sivic J., Russell B. C., Efros A. A., Zisserman A., Freeman W. T., “Discovering Objects and their Localization in Images”, In *Proceedings of the International Conference on Computer Vision*, 2005, p. 370 – 377.

Willamowski J., Arregui D., Csurka G., Dance C., Fan L., “Categorizing Nine Visual Classes Using Local Appearance Descriptors”, In *Proceedings of ICPR 2004 Workshop Learning for Adaptable Visual Systems*, Cambridge, United Kingdom, 2004.

An Animaloid Robot as a Cognitive Stimulator to support Elders with Cognitive Impairments: Preliminary Requirements

Giuseppe Anerdi* — Alberto Greco — Luca Odetti* — Guido Rodriguez*****

** EZ-Lab Research Centre on Gerontechnology, Scuola Superiore Sant'Anna, Pisa, 56127 Italy*

g.anerdi, l.odetti@sssup.it

*** DiSA, Lab. of Psychology and Cognitive Sciences, University of Genoa, 16126 Italy*

greco@unige.it

**** DiSEM, U.O. Neurofisiologia Clinica, Azienda Ospedale-Università San Martino, Genoa, 16132 Italy*

guido@unige.it

ABSTRACT. *We focus on some preliminary, theoretical and practical, requirements for an agent-based system for cognitive interaction with older people. We stress (a) the acceptability of an artificial agent by elders; (b) a user model, which includes user knowledge, expectations, goals; (c) a representation system, (d) an interaction system, including physical and verbal interaction. We describe in some detail a preliminary study aimed at assessing the acceptability of interaction of cognitively impaired elders with an artificial companion by studying their reactive behaviour during a simple experimental session. First results show that affective aspects of interaction with an artificial companion are not affected by negative feelings towards technology, but that positive attitudes are required in order to achieve awareness of its usefulness for a cognitive interaction. We also give some suggestions about method for the development of a working cognitive interaction system. The agent should build its user model - including relevant knowledge, expectations, and goals - by interactive learning, and operate jointly with a situation awareness engine.*

KEY WORDS: *animaloid, cognitive interaction, elders, gerontechnology.*

1. Introduction

The age-related physiological fall of cognitive abilities becomes a more and more challenging problem in our society; cognitive impairments affecting older people are often combined with the fact that many elders are, under various circumstances, living alone. Technology should not replace natural human interaction, but a possible help in some situations may come from interaction with artificial agents.

One natural solution, which older people have been using for centuries, is to live with an animal: this may improve the quality of life, obviously because it relieves loneliness but also because it gives elders something to do, to attend, to physically interact with, a cognitive challenge, and also enables them to gain a safety feeling because an animal does not focus the person's physical or mental disabilities. Pet-therapy is frequently used for the care of people affected by cognitive disabilities, with good results, e.g. improving social behaviours (Kongable *et al.*, 1990). On the other hand, for safety or hygienic reasons real animals are often not allowed in residential houses, or they are not otherwise liked. Moreover, their intelligence may not be sophisticated enough for cognitive challenges required in helping elder people. Thus the hypothesis of replacing real animals with animal-shaped objects is becoming object of a serious investigation (Nakajima *et al.*, 2001).

Animal-shaped toys have been evolved into animal-shaped robots, or animaloids, and researchers have started investigating how robotic companions can be used with elderly people, not only as “emotional activators”, but also for addressing interactive and communication functions, so acting as a support for their cognitive difficulties. For example, the interaction with the AIBO robotic dog on four elders with dementia resulted in improved communication patterns (Yonemitsu *et al.*, 2002). In a different comparative study, nine older women from a nursing home, with moderate to severe dementia, received two interactive sessions of 10 minutes each, with a plush cat and with the robotic cat NeCoRo; the emotional effects of such interactions on the patients were also studied (Libin *et al.*, 2004a). The same authors introduced also the new concepts of "robotic psychology" and "robot therapy", focusing on “interactive stimulation robots” (Libin *et al.*, 2004b). A unified assessment tool, named the Person-Robot Complex Interaction Scale (PRCIS) was also defined. In Japan, at the National Institute of Advanced Industrial Science and Technology (AIST), Takanori Shibata et al. have developed a robotic baby seal, named Paro. They carried out several experiments involving Paro and different groups of subjects, from children to elderly, to persons with cognitive disabilities (Wada *et al.*, 2005). At MIT Media Lab a group of researchers are developing a Teddy Bear-like robot, to investigate on the recognition of affective contents of touch in human-animaloid interaction (Stiehl *et al.*, 2005).

A most advanced result to be pursued, but still less achieved, is to make robotic companions act as true cognitive stimulators, by fully exploiting their interactivity and their (although limited) processing capabilities. In this paper, we focus on some preliminary, theoretical and practical, requirements for such a project.

In our view, an agent-based system for cognitive interaction with older people should be based at least on studies about:

(a) the **acceptability** of an artificial agent by elders;

(b) a **user model**, which includes user knowledge, user expectations, user goals;

(c) a **representation** or ontological system, essentially able to implement at least categorical functions;

(d) an **interaction** system, including physical and verbal interaction.

In order to better define acceptability, it is useful to distinguish three aspects of human interaction with artefacts:

affective, including emotional responses elicited by the artefact like fear, attraction, pleasantness, etc., or something referable to an immediate sense of a possible and meaningful integration into the person's own life space;

cognitive, including not much the knowledge about how to use it (which would be the usual cognitive aspect of attitude), but rather the use of cognitive resources such as memory or attention in tasks concerning a joint operation involving both the person and the artefact;

functional, including the perception of the potential usefulness and the possible control of the artefact in practical situations.

For example, an artefact may be acceptable on the affective grounds if it looks attractive and not frightening, on the cognitive grounds if it supports person's cognitive processes, and on the functional grounds if a person can clearly see what she can do with it.

We already undertook the first step, trying to assess the acceptability of interaction with an artificial companion by studying the reactive behaviour of older people with limited cognitive impairment during a simple experimental session. We are going to describe this study in some detail (for a full treatment, see Odetti *et al.*, 2007), and then to give some suggestions about the other points made above.

The pilot study described in this paper was conducted jointly by the Gerontechnology Group of the ARTS Lab at the Scuola Superiore Sant'Anna, and by the DISEM (Department of Endocrinological Science and Medicine) of the University of Genoa, between September 2006 and January 2007, to preliminarily evaluate how acceptable robot-mediated pet-therapy is for older people with moderate cognitive impairments (MCI or other kinds of dementia diseases in early stage).

The study was intended to gather some basic preliminary user-centered information. It is worth to point out that, although our preliminary survey was based on a group of people with clear cognitive disorders, it is not intended to be restricted to a clinical domain but it is aimed at investigating somewhat broad requirements, that can hold with elderly people in general. We think, however, that initially considering interaction with an artificial agent in patients with some clearly diagnosed cognitive impairment can force us to set such requirements more strictly even for less severe, not openly diagnosed, cognitive diseases "normally" found in senescence.

2. Method

2.1. *Participants*

Participants were 24 patients of the DISEM dementia evaluation service, in early stages of dementia of different types; their MMSE score averaged 27 ± 3 points. Informed consent was signed by each patient. Participants' average age was 76.6 (s=6.23); 12 of them were males and 12 were females. They were affected by the following disorders: 11 probable Alzheimer's disease (AD) (NINCDS-ADRDA criteria), 7 amnestic Mild Cognitive Impairment (aMCI) (Petersen *et al.*, 2001), 1 vascular dementia, 1 fronto-temporal dementia, 1 Parkinson dementia, 1 multi system atrophy and 2 subjective memory deficit.

2.2. *Scenario*

The experimental sessions took place in the room of medical examination, at the end of standard visits; some patients were at their first visit, others were returning patients. The examiner was sitting at the desk, with a laptop computer, and the patient was sitting in front of him/her. During the "standard" visit, the robot was not visible by the patient, who knew that she/he would have been involved in a very simple experiment at the end of examination. Nothing suggested the presence of unusual technology in the scenario. The informed consent briefly talked about a safe experiment, without suggesting that a robotic dog was going to enter the room. This was fundamental, since it was important to measure the reaction to a completely unexpected event like the entrance of an animaloid in the scene.

2.3. *Materials*

The experimental equipment was composed of one Sony AIBO ERS-7 robotic dog, one laptop computer (normally used by the examiner to take notes during the visit), and a simple client-server software, through which the examiner could easily control the execution flow of the experiment.

Simple robot motion tasks and reactive behaviours were implemented using URBI (Universal Real-Time Behaviour Interface, a client-server framework for high-level control of robotic platforms; Baillie, 2005).

The visual client application consisted essentially in a remote control to trigger the transitions between successive steps in the experiment. It was developed using C++ and QT4, and connected to the URBI server running on AIBO through a dedicated library.

2.4. *Experimental setup*

Sessions with patients were articulated into two main phases: first, on a command received from the remote controller managed by the examiner, the AIBO robot walked along a pre-defined path close to the patient and sat down, roughly in front of him/her. Then, it wagged its tail and barked in a “friendly” way. It was obviously fundamental that the dog was perceived as not aggressive.

With AIBO walking in, the examiner observed the patient reaction. If the patient showed fear or repulsion, the examiner tried to reassure her/him. Then, the examiner asked a couple of questions to the patient: “Does it reminds you of something? What does it look like to you?”, and “Do you like it?”. The answer to the second question was used by the examiner to decide whether the session could go on. In case of a negative answer, AIBO was sent away by the examiner. The same if the patient was showing fear or repulsion, and the examiner did not succeed in encouraging him/her. In case of a positive or a neutral answer, the examiner started the second phase of the session, picking up the AIBO dog from the floor and putting it on the desk, between him and the patient.

With a command from the remote control, the examiner let AIBO activate two different reactive behaviours. Then, he started petting the AIBO’s back while talking to the patient. AIBO reacted to this action by wagging its tail and moving its head and mouth, to suggest a feeling of “happiness”.

Then, the examiner asked the patient whether he/she wanted to try and do the same. If the answer was positive, the examiner invited the patient to gently pet the dog’s back, and observed his/her reaction to AIBO’s behaviour.

Then he asked the patient “Would you enjoy having such a robot at home?”, going more in depth in case of positive answer: “what could you do with it? Do you think that you could play together? Do you think that it could be useful to you in any way? How?”. At this point, during this discussion, the examiner suggested: “this dog could help you in reminding today’s date. Do you want to try?”. If Yes, the examiner would tell the patient how to let AIBO say the current date: “please, touch [the sensor on] its head for 2 seconds”. After this interaction was completed, the session ended with a few more questions, aimed at evaluating the patient’s feelings while interacting with the AIBO dog.

The following questions were asked during the sessions: Q1: Are you afraid of it? Q2: Does it remind you of a dog? Q3: Do you like it? Q4: Do you think it could be useful to you? Q5: Do you feel any attraction towards electronics goods? Q6: Would you like to give orders to it?

Furthermore, a general question about which electronic goods were present at the patients' home was asked.

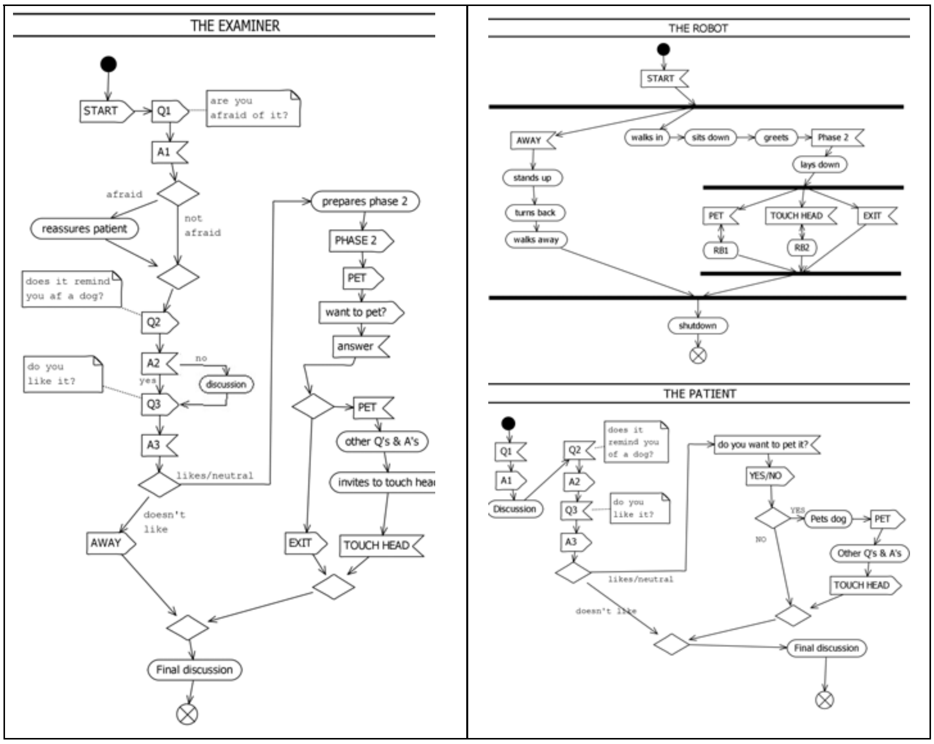


Figure 1. *UML activity diagram*

3. Results

Table 1 summarizes the patients' answers in terms of frequency, average score, and normalized average score in the [0..1] range for questions Q1, Q2, Q3, Q4, Q5, and Q6.

From the analysis of the scores, we can easily see that the AIBO is generally perceived by the pilot group as harmless, friendly, cute. During the study, only 2

	0 (NO)	1	2	3	4	5 (VERY)	Avg. (norm.)
Q1	23	0	0	1	0	0	0.13 (.03)
Q2	1	1	0	2	6	14	4.21 (.84)
Q3	1	1	1	1	1	19	4.38 (.88)
Q4	9	0	2	4	2	7	2.46 (.49)
Q5	9	2	2	4	2	5	2.13 (.43)
Q6	(NO) 12					(YES) 12	0.50

Table 1. *Scores for answers to questions 1-6: frequency, average and normalized average*

patients answered negatively to Q1 or Q3, causing the examiner to send the robot away. On the other hand, it is rather evident that the participants to the study tended not to perceive its potential usefulness, or at least, not as clearly as they recognized it as cute and friendly (avg. score 2.06 for Q4). Only one half thought that they would like to give orders to the dog. Answers to Q4 and Q6 actually show a strong relationship with Q5 scores, as shown by Table 2.

	Q1	Q2	Q3	Q4	Q6
Corr(Qi, A5)	-0.22	0.43	0.32	0.67	0.65

Table 2. *Correlation between answers to Q1,2,3,4,6 and answers to Q5*

Dividing the group into two subgroups SG0 (including patients who answered with a 0 score to Q5) and SG+ (Q5 score ≥ 1) the resulting scores (Table 3) are significantly different. They clearly show that the patients who did not completely reject technology tended to accept the idea that a companion robot like AIBO could also be useful in their daily life.

Question		Avg. score		
		All $\overline{Q5} = .43$	SG+ $\overline{Q5} = .68$	SG0 $\overline{Q5} = 0$
Affective interaction	Q1	0.03	0.00	0.07
	Q2	0.84	0.92	0.71
	Q3	0.88	0.91	0.82
Functional interaction	Q4	0.49	0.68	0.18
	Q6	0.50	0.73	0.11

Table 3. *Subjects with a 0 score on Q5 have a very low score also on Q4 and Q6, while scores on Q1,2 and 3 remain substantially uniform*

Significant difference in answers to Q4 and Q6 also appears when the group is divided in two subgroups basing on gender (Table 4).

Question		Avg. score		
		All	M	F
Affective interaction	Q1	0.03	0.00	0.05
	Q2	0.84	0.91	0.76
	Q3	0.88	0.85	0.91
Attitude to technology	Q5	0.43	0.38	0.47
Functional interaction	Q4	0.49	0.40	0.60
	Q6	0.50	0.38	0.64

Table 4. *Normalized average scores, by gender*

Finally, as shown in Table 5, no relationship seems to exist between answers to questions from 1 to 6 and MMSE score, while a negative correlation is suggested with education level.

	Corr(Qi, S)	Corr(Qi, MMSE)
Q1	-0.18	-0.06
Q2	-0.05	0.02
Q3	-0.49	0.25
Q4	-0.55	0.35
Q5	-0.51	0.34
Q6	-0.45	-0.02

Table 5. *Correlation values between question scores, years of scholarity and MMSE scores*

4. Discussion

A preliminary analysis of the acceptability of animaloid companion robots by older people with limited cognitive impairment has shown that different attitudes to technology (rejection vs. a generically non negative attitude) do not affect significantly the **affective** components of interaction between an AIBO robot and older people with early/pre-clinical dementia (Tables 3, 4). On the contrary, the ability to imagine that such an artefact can also be useful, and that it is possible to have a **functional** interaction with it, seems to depend strongly on the non negative feeling for technology.

Differences between women and men did not seem very relevant in answers to questions dealing with the affective component of human-robot interaction, while women seem to be slightly more attracted than men by electronic goods and technology. This is a somewhat unexpected result, which is coherent with answers to question Q4 and Q6: attention to functional aspects of interaction, including usefulness and control of the artefact, seems to be stronger in women.

Finally, the negative correlation between question scores and level of education could suggest that less educated people accept the robot and its potential usefulness more easily than people with a higher level of education.

The animaloid-like appearance seems not to be a significant advantage because some of the subjects are not able to recognize it (as we could expect due to the cognitive problems). Few emotional links are also created. This is probably due to the already limited cognitive capacities of the subjects. These are very preliminary results which need to be investigated more in details with the animaloid robot autonomously behaving and during the evolution of the cognitive disability from the very onset.

A fully interacting system for older people, acting in human centered environment as a companion in their daily life, capable of recognising and monitoring their actions, would require complex perceptual, knowledge representation, and communication subsystems. Some practical compromises about sensorimotor aspects then should be put in place. Perceptual constraints can be overcome by mature technologies, e.g. by using scanning systems for object recognition, smart tags, or by making communication rely on voice recognition/production. Our preliminary study about interaction shows that in some conditions this can be acceptable.

Our study did not yet consider **cognitive** aspects of interaction. From this point of view, various challenging problems must be solved in order to make a working model. Here we can only suggest some ways to set up next steps, obviously not deriving from above described results. The most difficult problems from the cognitive perspective are in knowledge representation, notably in how to feed the system with really relevant knowledge. We think that a good way to achieve this result would be through interactive learning, not explicit top-down teaching. The agent should be actively searching for information in order to build and update its user model. User model should include user goals, expectations, and actions (routines firstly, but also non routines); the agent should have its own expectations and problem detection would rely on the comparison between both expectation patterns.

The above considerations are leading toward the new, emerging paradigm of Ubiquitous Robotics, strictly related to the application domain of Ambient Assisted Living, that arises from a shift of the focus from the more “traditional” concept of Ubiquitous/Pervasive Computing (i.e. from information) to matter and physicality. Networked, ubiquitous robotic systems conveying data *and* physical actions in intelligent environments, can provoke a profound and pervasive impact at different scales: global, local, “personal”, external and internal and so forth. The role of a relational artifact, like an animaloid companion robot, embedded in an Ubiquitous

Robotics system, would vary greatly, depending on the user's functional and cognitive profile.

When considering the full range of conditions related to ageing, three large groups of potential older users of Ubiquitous Robotics systems can be identified: people with physical or psychological reductions but without definite cognitive deficits; older people affected by minor cognitive deficits; older people with moderate to severe dementia. In the first group, the robot could interact with a digital personal planner, to remind the user of her/his daily schedule. With a user of the second group, the robot could provide the user with cognitive feedback and support during the execution of activities of daily living. Finally, with patients affected by moderate to severe cognitive impairment, the communication functionalities would drastically lose meaningfulness, because such patients would be unable to understand the meaning of the messages themselves.

Cognitive aspects of a model aimed at interacting in real time with impaired people especially include resource-related processes, like working memory and attention. One possible development of research is studying how the above described targeted knowledge representation could work in conjunction with a situation awareness monitoring engine. Situation awareness (Endsley, 2004) is a concept that has been proposed especially in studies about aviation, emergencies or other contexts where environmental knowledge is critical and information flow may be overwhelming. This term refers to the continuous extraction of environmental information, the comprehension of its meaning, its integration with previous knowledge to form a consistent mental representation, and the projection of that representation in anticipating events in the near future. Older people often may find themselves in contexts where attention to situation and error-free decision can be critical. In order to build a working monitoring engine, a shared representation system should be developed; for example, each event could be coded as having attributes like "what", "when", and "where" (Livnat et al., 2005).

Overall, our preliminary work shows that interaction of older people with artificial agents can work in some definite conditions. In particular, it shows that merely encouraging affective aspects of interaction or a mere acquaintance with a robotic companion may not be sufficient. Also an explicit belief of its usefulness, and of its functional operation and ways of control should be promoted. Our long-term aim is to design and develop cognitive ubiquitous robotics, including a full treatment of cognitive aspects, for older people both subject to "normal" cognitive ageing and to "mild" cognitive impairments.

The full team of authors that participated in the preliminary experiment included Luca Odetti, Paolo Dario, Silvestro Micera (EZ-Lab), and Maria Paola Barbieri, Debora Mazzei, Elisa Rizza (DiSEM).

5. References

Baillie J.-C. “URBI: towards a universal robotic low-level programming language”, *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems,(IROS 2005)*, 2005, p. 820-825.

Endsley M.R. “Situation awareness: progress and directions”, in S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory, measurement and application*, Aldershot, UK, Ashgate Publishing, 2004, p. 317-341.

Kongable L. G., Stolley J. M, Buckwalter K. C. “Pet therapy for Alzheimer's patients: a survey”, *J. Long Term Care Adm.*, 18(3), 1990, p. 17-21.

Libin A. V., Cohen-Mansfield J., “Therapeutic robocat for nursing home residents with dementia: preliminary inquiry”, *American J. of Alzheimer’s Disease and other Dementias*, 19 (2), 2004a, p. 111-116.

Libin A. V., Libin E. V., “Person-robot interactions from the robopsychologists' point of view: The robotic psychology and robotherapy approach”, *Proc. of the IEEE*, 92 (11), 2004b, p. 1789-1803.

Livnat Y., Agutter J., Moon S. , Foresti S., “Visual Correlation for Situational Awareness ”, *Infovis*, IEEE Computer Society, 2005.

Nakajima K. , Nakamura K., Yonemitsu S. , Oikawa D., Ito A., Fujimoto T., Nambu A., Tamura T. Y., “Animal-shaped toys as therapeutic tools for patients with severe dementia”, *Proc. 23rd Int’l Conf. of the IEEE Engineering in Medicine and Biology Society*, 42001, p. 3796- 3798.

Odetti L. , Anerdi G. , Barbieri M.P. , Mazzei D. , Rizza E. , Dario P., Rodriguez G., Micera S. , “Preliminary experiments on the acceptability of animaloid companion robots by older people with early dementia”, *EMBC07, 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, August 23-26, 2007.

Petersen R. C. , Stevens J. C. , Ganguli M. , Tangalos E. G. , Cummings J. L. , DeKosky S. T. , “Practice Parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). Report of the quality standards subcommittee of the American Academy of Neurology”, *Neurology*, 562001, p. 1133-1142.

Stiehl W. D., Lieberman J., Breazeal C., Basel L. , Lalla L., Wolf M. , “Design of a Therapeutic Robotic Companion for Relational, Affective Touch”, *14th IEEE Int’l Workshop on Robots and Human Interactive Communication (RO-MAN 2005)*, Nashville, TN, USA, 2005, p. 408-415.

Wada K., Shibata T., Saito T., Tanie K., “Effects of Three Months Robot Assisted Activity to Depression of Elderly People who stay at a Health Service Facility for the Aged”, *Proc. 2005 IEEE Int’l Conf. on Robotics and Automation (ICRA 2005)*, Barcelona, Spain, April 2005, p. 2785-2790.

Yonemitsu S. , Higashi Y., Fujimoto T., Tamura T., “Research for practical use of rehabilitation support equipment for severe dementia”, *Gerontechnology*, 2(1), 2002, p. 91.

"HCP-based Applications" session

04:00 – 05:30 pm

Tuesday, June 10, 2008

Context-based learning for support systems

Juliette Brézillon – Patrick Brézillon

LIP6, 104 avenue du Président Kennedy,
75016 Paris, France
{Juliette, Patrick}.Brezillon@lip6..fr

ABSTRACT. *Often, decision-making processes are embedded in "official" procedures to address focuses in any case. However, procedures often lead to sub-optimal solutions for any specific decision-making, and actors are obliged to develop practices to address the specificity of the context in which a decision is made. This opposing procedure versus practices is well known in different domains (prescribed and effective tasks, logic of functioning versus logic of use, etc.). We have shown what the differences were in a context-based formalism called Contextual Graphs. In this paper, we discuss the possibility of an explicit use of "good" and "bad" practices for the training of human actors, thanks to an incremental acquisition of knowledge and the learning of new practices by the system. We discuss these aspects in the framework of a real-world application in road safety (modelling of drivers' behaviours).*

RÉSUMÉ. *Souvent les processus de prise décision sont inscrites dans des procédures officielles qui sont supposées correspondre à un focus en toute situation. Cependant, de telles procédures ne conduisent qu'à des solutions sous-optimales dans la plupart des prises de décision particulières, et les acteurs doivent alors développer des pratiques pour s'adapter aux spécificités de chaque contexte particulier d'une prise de décision. Cette dichotomie procédures-pratiques se retrouve dans de nombreux domaines (tâches prescrites et effectives, logique de fonctionnement et logique d'utilisation, etc.). Le formalisme des Graphes contextuels nous a permis d'expliciter la forme que prenait concrètement une telle dichotomie. Ce papier se concentre sur l'utilisation explicite des « bonnes » et « mauvaises » pratiques qui y sont représentées grâce à une acquisition incrémentale de connaissances et simultanément l'apprentissage de nouvelles pratiques par le système. Ces aspects sont discutés dans le cadre d'une application en sécurité routière (la modélisation du comportement du conducteur).*

KEY WORDS:*Context modeling, Contextual graphs, procedures and practice,s driver modelling.*

MOTS-CLÉS:*Modélisation du contexte, Graphes contextuels, procédures et pratiques, modélisation du conducteur.*

1. Introduction

Decision makers face many heterogeneous contextual cues. Some of these cues are always relevant (time period, unpredicted event, etc.) but others are only used in some cases (number of lines on the road, position on the line, etc.). Thus, in the driving domain, actors must deal with a set of heterogeneous and incomplete information on the driving situation to make their decisions. As a consequence, a variety of strategies are observed for solving driving situation: from an actor to another one, but also for the same actor at different instants and/or different contexts. Thus, it is not obvious to get a comprehensive view of the mental representations at work in the subject's brain in many human tasks, especially because it is better to store advantages and disadvantages rather than the complete decision.

As a consequence, actors plan their action in real time rather than relying on official procedures for two main reasons. Firstly, the procedure is never perfectly adapted to the situation at hand and can lead to improper actions or sub-optimal solving strategies. Secondly, the actor can miss some important facts or notice them too late to adequately solve the problem. Nevertheless, actors consider procedures as useful guidelines to be tailored for each particular focus. Thus, each actor transforms the procedure in a practice to address a focus in a specific context, and one observes almost as many practices as actors for a given procedure. This is a general way to reach the efficiency that decision makers intended when designing the task (Pomerol, 2001).

For capturing such an experience-based knowledge, Intelligent Assistant Systems must be able to deal with the contextual dimension of the knowledge in order to manage practices and not only procedures like first expert systems did. The key is that the decision-making process relies on a practical reasoning that depends on contextual knowledge and information. Practical reasoning is not a logical and theoretical reasoning for which the action leads to a conclusion. Practical reasoning has more a status of inductive probabilistic reasoning: the conclusion cannot be detached (i.e. take a meaning) from the premises. Modeling actors' reasoning is a difficult task because a number of contextual elements are used. These pieces of knowledge, which are not necessarily expressed, result in more or less proceduralized actions that are compiled in comprehensive knowledge about actions.

Hereafter, the paper is organized in the following way. The next section discusses the relationships between context and decision making. After, we introduce an example of a real-world application on which we work, namely the modelling of car drivers arriving at a crossroad. In the following section, we discuss the relationships between learning and context-based decision making. First, we show that Contextual Graphs lead to the creation of bases of experiences for intelligent systems. Another important aspect concerns the possibility to represent good practices as well as bad practices in a contextual graph. We finally present the perspectives open by a context-based approach of decision making.

2. Context and decision making

2.1. *A conceptual framework for modelling context*

We cannot speak of context out of its context. Context surrounds a focus (e.g. the decision making process or the task at hand) and gives meaning to items related to the focus. Thus, on the one hand, context guides the focus of attention, i.e. the subset of common ground that is pertinent to the current task. Indeed, context acts more on the relationships between the items in the focus than on items themselves, modifying their extension and surface. On the other hand, the focus allows identifying the relevant elements to consider in the context. It specifies what must be contextual knowledge and external knowledge in the context at a given step. For example, a focus on the driving task mobilizes contextual knowledge such as the fact of knowing the meaning of the traffic signs, the fact to have learned how to drive, etc., i.e. knowledge that could eventually be used when the focus evolves. Some knowledge from driver's personal context could also be considered such as a previous experience in the driving task. For example, this corresponds to the choice of a specific method at a given step of a task. For solving a driving situation, a driver has several solutions, e.g. several behaviours for crossing an intersection. Indeed, some contextual elements are considered explicitly, say for the selection of the behaviour and thus can be considered as a part of the way in which the problem is solved at the considered step.

In reference to a focus, Brézillon and Pomerol (1999) consider context as the sum of two types of knowledge. First, there is the part of the context that is relevant at this step of the decision making, and the part that is not relevant. The latter part is called **external knowledge**. External knowledge appears in different sources, such as the knowledge known by the decision maker but lets implicit with respect to the current focus, the knowledge unknown to the decision maker (out of his competence), the contextual knowledge of other actors in a team, etc. The former part is called **contextual knowledge**, and obviously depends on the decision maker and on the decision at hand. Here, the focus acts as a discriminating factor between the external and contextual knowledge. However, the frontier between external and contextual knowledge is porous and moves with the progress of the focus.

A sub-set of the contextual knowledge is proceduralized for addressing specifically the current focus. We call it the **proceduralized context**. The proceduralized context is a sub-set of contextual knowledge that is invoked, assembled, organized, structured and situated according to the given focus and is common to the various people involved in decision making. A proceduralized context is quite similar, in the spirit, to the chunk of knowledge discussed in SOAR (Schank, 1982; Laird et al., 1987), and, in its building, to Clancey's view (1992) on diagnosis as the building of a situation-specific model. A proceduralized context is like a local model that accounts for a precise goal in a specific situation (at a given step). In a distinction reminiscent of cognitive ergonomics (Leplat and Hoc, 1983), we could say that the contextual knowledge is useful in identifying the activity whereas the proceduralized context is relevant to characterize the task at hand (i.e. concerned by the activity).

An important issue is the passage of elements from contextual knowledge to a proceduralized context. This proceduralization process, which depends on the focus on a task, is task-oriented just as the know-how, and is often triggered by an event or primed by the recognition of a pattern. This proceduralization process provides a consistent explanatory framework to anticipate the results of a decision or an action. This consistency is obtained by reasoning about causes and consequences and particularly their relationships in a given situation. Thus, we can separate the reasoning between diagnosing the real context and, anticipating the follow up (Pomerol, 1997). The second step needs a conscious reasoning about causes and consequences.

Brézillon and Brézillon (2007) discuss a second type of proceduralization, namely the instantiation of contextual elements. This means that the contextual knowledge or background knowledge needs some further specifications to perfectly fit the decision making at hand. The precision and specification brought to the contextual knowledge is also a part of the proceduralization process that leads from the contextual knowledge to the proceduralized context. For each instantiation of a contextual element, a particular action will be executed. There are as many actions as different instantiations. However, once the corresponding action is executed, the instantiation does not matter anymore and the contextual element leaves the proceduralized context and goes back in the contextual knowledge. For example, arriving to a crossroad, a driver looks at the traffic light. If it is the green signal, then the driver will decide to cross. The instantiation of the contextual element "traffic light" (green signal) has guided the decision making process and then the decision is made. The colour of the traffic light does not matter after the decision is made.

2.2. *A context-based representation formalism*

Contextual graphs (Brézillon, 2005) propose a representation of this combination of diagnosis and actions. (A contextual graph represents problem solving.) Contextual nodes represent diagnosis. When a contextual node is encountered, an element of the situation is analyzed (and the value of the contextual element, its instantiation, is taken into account). Thus, contextual graphs allow a wide category of diagnosis/action representations for a given problem solving. Contextual graphs are acyclic due to the time-directed representation and guarantees algorithm termination. Each contextual graph (and any sub-graphs in it) has exactly one root and one end node because the decision making process starts in a state of affairs and ends in another state of affairs (not necessarily with a unique solution for all the paths) and the branches express only different context-dependent ways to achieve this goal. This gives a general structure of a spindle to contextual graphs. A path represents a practice developed by an actor, and there are as many paths as practices known by the system.

Elements of a contextual graph are: actions, contextual elements, sub-graphs, activities and parallel action groupings (Brézillon, 2005). An action is the building block of contextual graphs. A contextual element is a pair of nodes, a contextual node and a recombination node; a contextual node has one input and N outputs (branches) corresponding to the N instantiations of the contextual element. The recombination node is [N, 1] and represents the moment at which the instantiation of the contextual element does matter anymore and the paths of all the branches starting at the contextual node are identical. Sub-graphs are themselves contextual graphs. They are mainly used for obtaining different displays of the contextual graph by aggregation and expansion like in Sowa's conceptual graphs (2000).

An activity is a particular sub-graph that is identified by actors because appearing in a same way in different problem solving. An activity is defined in terms of actor, situation, task and a set of actions. More precisely, an activity is a sequence of actions executed, in a given situation, for achieving a particular task that is to be accomplished by a given actor. In the decision making area, actors identify an activity as a recurring structure in problem solving. This recurring sub-structure is a complex action in the spirit of the notion of scheme given in cognitive ergonomics (Vergnaud, 1985) where schemes are intended for completing sub-goals. Each scheme organizes an activity around an object and can call other schemes to complete specific sub-goals. A scheme can be specified by a name, a goal and a contextual graph representing a decision-making that allows achieving its goal in a context-sensitive way. Both contextual graphs and schemes allow the representation of an actors' activity and all the variants (procedures and practices), the integration of automatic learning and adaptation in a system, a clear representation of the context in actors' reasoning, and the organization of the actors' activity itself.

3. An example: Car-drivers' behaviours at a T-intersection

Elements of a contextual graph are: actions, contextual elements, sub-graphs, activities and parallel action groupings (Brézillon, 2005). An action is the building block of contextual graphs. A contextual element is a pair of nodes, a contextual node and a recombination node; a contextual node has one input and N outputs (branches) corresponding to the N instantiations of the contextual element. The recombination node is [N, 1] and represents the moment at which the instantiation of the contextual element does matter anymore and the paths of all the branches starting at the contextual node are identical. Sub-graphs are themselves contextual graphs. They are mainly used for obtaining different displays of the contextual graph by aggregation and expansion like in Sowa's conceptual graphs (2000).

Consider the following intersection with a T-configuration (Figure 1). The situation concerns two cars, a black car (called here car-A with driver-A) going straight ahead and a white car (called here car-B with driver-B) at a "Yield (a stop) sign" sign. The study is according to the viewpoint of driver-A (and not from the "external" viewpoint of an observer), and thus with his partial viewpoint. Driver-A is supposed to stay on the same road, and arriving at the T-intersection, he knows that he has priority over any car coming from the right road. Thus, the procedure given by the Highway Code is summed up by a unique action: "You have priority, you can cross the T-intersection".

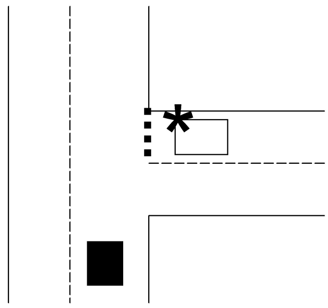
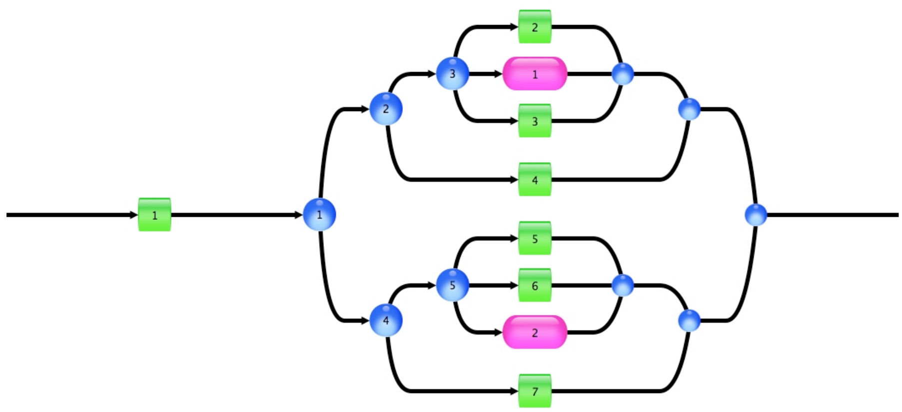


Figure 1. *A T-intersection situation*

Now, driver-A can be any individual, some will look anyway on the right with the strong mental representation of "I must let priority to cars coming from right", even if this procedure is inhibited by the fact that driver-B has a "give way" sign. We represent the different types of driver-A's behaviours in Figures 2 and 3.



(a) A1: Arrive to a T-intersection with the right-of-way
(b) C 1: Look at right?
(c) Yes C2: Is there a car at right?
(d) Yes C3: Pay attention to Car-B?
(e) No A2: Cross the T-intersection
(f) Yes 1: Activity « Car-B Management »

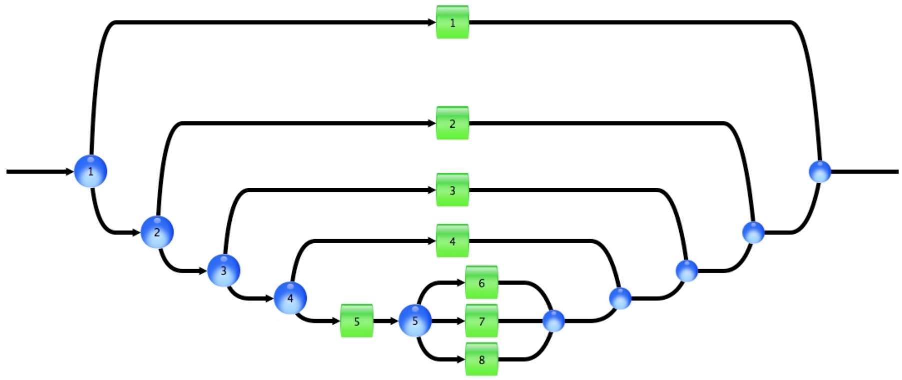
(g) Stop A3: Brake strongly
(h) No A4 : Cross the T-crossroad
(i) No C4 : Expectation?
(j) Unpredicted event C5: Decision?
(k) Stop A5: Brake strongly
(l) Keep right-of-way A6: Cross the T-intersection
(m) Take care 2: Activity « Car-B Management »
(n) Correct A7: Cross the T-intersection

Figure 2. *Contextual graph of the T-intersection situation solving*

Figure 2 presents different possible behaviours of driver-A. Normal behaviours conclude on the procedure “Cross the T-intersection” (e.g. lines e and h but not line l) but there are some abnormal behaviour. For example, line (g) driver-A brakes when he has the priority. One reason could be that driver-A does not pay enough attention to the situation, and a strong automatism to “Allow right-of-way to cars coming from the right” is responsible of this decision. Abnormal behaviours can correspond to “bad behaviours” of Driver-A depending on the context. For example, maintaining our priority (line e) when car-B is moving could lead to a critical situation. It is interesting also to note that a decision can be “good” or “bad” according to the context. For example, braking strongly line (g) can be dangerous because the car behind may have a different interpretation (I follow car-A that would cross), when braking strongly line (k) could be a safe decision.

The two pink ovals in Figure 2 represent the activity “Car-B management” and is detailed in Figure 3. Its activation corresponds to a good behaviour of driver-A. Figure 3 presents different interpretations of driver-A on the driver-B’s behaviour. Note that it is the explanation given by driver-A on what may happen to the car-B. There are different reasons for a movement of car-B, normal reasons (e.g. the car-B is just arriving to the intersection and its driver is braking to stop at the level of the road sign “Yield sign”) and abnormal reasons (and thus critical situations) like driver-B believes he has time to act before the arrival of car-A.

The different practices described in Figures 2 and 3 could be enriched by introducing additional contextual elements concerning more specifically the driver: “Do you think that Driver-B has seen you?”, “Are you in a hurry?”, “Do you take usually risks?”, “Do you try to impress the other?”, “How do you behave if Driver-B is a woman?”, etc. This could be done incrementally, the system acquiring new contextual elements and learning new practices. A contextual graph is thus a real base of experience on a specific topic and all the known ways to address this topic. The possibility to represent good practices as well as bad practices is the key point for humans training after.



C1: Status of car-B?
Stopped A1: Cross the T-intersection
Moving C2: Type of motion?
Reducing speed
A2: Cross the T-intersection
Accelerating

C3: Is braking sufficient?
Yes A3: Brake, stop and let car-B decide
No C4: Where is going car-B?
on my direction
A4: Reduce the effect of the collision
on the opposite direction
A5: Car-B will cross my road
C5: Overtaking?
Before Car-B A6: Overtake on the left
No possibility A7: Reduce collision
After Car-B A8: Overtake behind Car-B

Figure 3. *Activity "Car-B management" of driver-A in the Figure 2.*

4. Learning and context-based decision making

4.1. *"Experience bases" for intelligent assistant systems*

Storing "experience-based situations" in contextual graphs leads to the development of "corporate memories," each corporate memory corresponding to the collection of all the practices developed for a situation solving. The next step is the design and development of an intelligent assistant system that would be able to use such a structured knowledge base to provide the actor with an experience-based support in contrast to early expert systems using flat knowledge bases built from the explicit knowledge or heuristics independently.

By the uniform representation of elements of reasoning and of contexts, contextual graphs offer a way to learn the way that these elements (reasoning and contexts) are assembled in a practice. Based on such a context-based representation, an intelligent assistant system will address experience building more directly rather than a simple knowledge building. This is why we call such systems Context-based Intelligent Assistant Systems (CIASs) and this is now our concern in a design process in road safety, the objective being to develop a CIAS for supporting a continuous training of car drivers, especially apprentices but also more advanced, in the development of their experience by contextualization of the Highway Code learned in driving school. For such CIASs, contextual graphs allow to develop “experience-oriented knowledge bases.”

The knowledge base of the CIAS is developed in two steps: (1) by a short elicitation of knowledge from operators and a use of reports, books and related matter, and (2) an incremental enrichment of contextual graphs by interaction with actors. This approach has the advantage of rapidly providing a mock-up that will be improved progressively. This is a kind of incremental acquisition of new knowledge pieces where knowledge is acquired when needed and in its context of use, and, jointly, of learning new strategies when actor’s reasoning relies on a new practice (i.e. a new path in the contextual graph). Such systems present a smart solution to one of the main weaknesses of prior systems, namely the lack of ability to evolve.

Anticipatory capability is enhanced in a representation by contextual graphs because a CIAS is able to develop a reasoning more directly related to the real situation, not in a mechanical way like with procedure. Thus, the support to an actor concerns elements of reasoning and of contexts and how all these elements are assembled in a practice. An anticipatory system uses knowledge about future states to decide what action to make at the moment. It should be able to predict what will probably happen and alert the driver for the occurrence of a crucial or time-critical event and its consequences. An anticipatory capability assumes a simulation component in the system. Simulation is also a way for a system to provide explanation by contextualizing its reasoning.

An efficient system supports actors by dealing with practices and not only procedures. This supposes that such systems can use a context-based representation of the knowledge and of the reasoning for problem solving. As a side effect, CIASs are able to rationalize solutions then to the actors. They can explain at different levels of detail the rationale behind the solution as operators are generally interested in the solution through the contextual cues. Because contextual graphs organize the knowledge at a high level, the intelligent assistant system expresses its knowledge and reasoning in a form directly understandable by the actors.

4.2. *System learning*

Contextual graphs bring a solution to Henninger's claim (1992) that "you won't know what is really needed until you're in the design process," because contextual graphs include a natural process for incrementally acquire missing knowledge and jointly learn new practices. Incremental knowledge acquisition and practice learning intervene in contexts in which the system fails, that is, the contextual graph does not include the right practice. Incremental knowledge acquisition plays an important role in two situations:

- When the knowledge is missing in the current context, the user adds new knowledge through explanations that enable the contextual knowledge to be proceduralized.
- A chunk of knowledge may be used incorrectly by the system because a link to the current context is missing. Here, incremental knowledge acquisition must focus on the refinement of the proceduralized context, and explanation can support that refinement.

Gathering and using knowledge in the context of use greatly simplifies knowledge acquisition because the knowledge provided by experts is always in a specific context and is essentially a justification of the expert's judgment in that context.

The main idea of these approaches is that experts provide their knowledge in a specific context and that knowledge can only be recorded with its context, that is, the system acquired the new knowledge and the context in which this new knowledge is required. Thus, knowledge is not generalized when it is acquired. It is fundamental to record the context in which the knowledge is acquired and can be used. In other term, the expert provides a triple {problem, context, solution} when for early expert systems the knowledge engineer generalized by a pair {problem, solution} in a more general context that could be different from those in expert's mind.

The learning process is an assimilation process because it corresponds generally to a process of refinement of contexts. An accommodation process would suppose that the whole structure of the contextual graph could be reorganized in another way. The interest to learn what is a bad practice presents different aspects. First, the system can identify automatically how an actor behaves, which (contextual) elements have been misunderstood by the actor, evaluate the criticity of his behaviour, and propose a way to change his behaviour by finding scenarios in which the misunderstood contextual elements will play a key role. In that sense, the system and the human will cooperate effectively like a joint cognitive system (Woods et al., 1990; Brézillon and Pomerol, 1997a).

In the example given Figure 2, the system knows:

(g) Stop A3: Brake strongly

However, driver-A may only reduce the speed of car-A. The system benefits of this failure to learn a new practice and acquire the missing contextual knowledge not considered up to now. A reason provides by driver-A could be:

C Is there another car behind me?
Yes Brake to reduce speed and alert the other driver of a problem
No Brake strongly

Indeed, the contextual element "another car behind" was initially supposed to have the value "No" and thus not considered explicitly.

4.3. *Human learning through good and bad practices*

By its mechanism of incremental knowledge acquisition and practice learning, Contextual Graphs allow to collect all the ways to solve a problem. This is the policy that follow in different real-world applications like the applications for incident solving in subway (Pomerol et al., 2002; Brézillon et al., 2003) or the information retrieval (Brézillon, 2005). In our current application for training of car drivers, we explore a new use of Contextual Graphs by considering the correct practices provided by the Highway Code (behaviours of good drivers), but also the bad practices executed by novices, bad drivers or drivers under drug influence (Brézillon et al., 2006). The interest of storing bad practices in contextual graphs is to allow a CIAS to identify online the current behaviour of a driver and anticipate the consequences of the scenario chosen by the driver. Rich data are available concerning road accidents in a number of situations (for modelling bad practices). We are currently identifying drivers' behaviours based on previous works (GADGET, 1999) and a questionnaire. The following step will be to couple the contextual graph of drivers' behaviours (i.e. their decision making process) with the corresponding situation space in order to identify scenarios of good and bad behaviours. For the training, the system will propose a scenario with a critical driving situation learnt from bad drivers and interact with the driver-player to help him to make the right decision in the pre-critical situation in which the driver has yet the option to retrieve a normal situation. In more general terms, this leads to a game in which the system knows the different ways leading to a good decision as well as the ways leading to a known bad decision. In the latter case, the CIAS will have the means to anticipate critical situation, and to explain mistakes of the apprentice, generally a misinterpretation of contextual cues.

Other variants will be considered later. We study the representation of a given situation solved by all the drivers in a unique contextual graph to focus on all the practices developed by drivers independently of their experience. Another approach is to represent each driver's view (and all the practices imagined by the driver on a situation solving), according to his experience, familiarity with driving, etc. Then a contextual graph corresponds to a problem solving and an actor. All these contextual graphs can be then classified from the total novice to the highly experimented driver. (We are now checking this approach on a simple decision making "Buy a subway ticket" solved by different persons more or less knowledgeable with the subway in Paris.) The two approaches could lead to two views on driver's behaviour, a personal view on his personal evolution in his contextualization of the theoretical knowledge, and a collaborative view of the driver interacting with other drivers.

5. Conclusion

We propose contextual graphs for a uniform representation of elements of reasoning and contextual elements at the same level. This is different from the view of Campbell and Goodman (1988) for example (and Hendrix, 1975, before semantic networks) that consider context as a way to partition a graph. Moreover, context in our formalism intervenes more at the levels of the links between elements of reasoning than elements themselves. Contextual elements being organized in contextual graphs in the spirit of “nest of dolls”, we have not a hierarchy of context because a given contextual element is itself contextualized and can appear encompassed in different other contextual elements. Rather, a contextual element is a factor of knowledge activation.

A contextual graph is a kind of “micro corporate memory” that provides a knowledge base that is more “experience oriented” than “goal oriented.” Contextual graphs are the experience-based representation of the knowledge and reasoning need to intelligent systems. Relying on contextual knowledge and the possibility to acquire automatically most of the contextual information, an intelligent system is able to: (a) identify user’s intention, (b) simulate (in accelerated time) the execution of the user’s decision to anticipate consequences, (c) compare theoretical and user behaviours, (d) alert the user either for a wrong decision (by lack of the right context) or for a discrepancy in planned and effective outcomes of the decision.

There are several problems yet open from a theoretical point of view as well as a practical one.

First, we point out that contextual elements considered in a contextual graph constitute a heterogeneous population that is difficult to represent in a hierarchy or an ontology. A contextual element can concern the actor (e.g. I prefer a secure solution to a risky one) but does not belong directly to the domain of application. The representation of a set of such (heterogeneous) contextual elements is a challenge.

Second, a contextual element may itself be a “chunk of contextual knowledge” including more basic contextual elements. For example, a contextual element such as “Must I respect the yellow traffic light?” may cover (sub) contextual elements such as “I am in a hurry”, “I have room and time to cross”, etc. The challenge here concerns a modelling of a contextual element at a finer granularity and, maybe, by extension a modelling of parallel action groups.

Third, an activity is a sub-graph identified by actors as a recurring structure appearing in several contextual graphs. The introduction of activities relieves the representation by contextual graphs by introducing a sub contextual graph and leading to a network of contextual graphs, i.e. of problem solving. However, the most interesting observation here is the fact that the notion of activity allows simplified interaction among actors, one actor giving the name of the activity and the other actor developing the activity. For example, “Turn right” is an activity that a car driver translate in “Put indicator on left”, “Look behind if the following car is not too close”, “Brake to reduce speed”, “Look at pedestrians crossing the other road”, etc.

Four, a proceduralized context is perceived as a chunk of contextual knowledge leading to the choice of an action to execute. However, the contextual elements intervening in this proceduralized context, their instantiations and their relationships will remain available. This leads to the possibility of generating rich explanations, and even of new types of explanations like the way in which a contextual graph growths from the initial procedure to the last practice introduced.

Context plays a role in many types of reasoning, and notably in decision making. Making explicit context in the representation of a decision making process allows to integrate incremental knowledge acquisition and practice learning as part of the process of the decision making itself. Moreover, Contextual Graphs offering the possibility to represent good practices as well as bad practices allows to couple a tool for learning all the ways for a situation solving (good and bad practices) and a tool for identifying any behavior and propose a rational way to improve our behavior. This seems to us a first step towards an attempt to rationalize the decision making process.

Acknowledgements

The work presented in this paper is part of the ACC project that is supported by PREDIT and the French Minister of Transportation, mainly concerning the funding of the Ph.D. Thesis. We also want to thanks the other members of the ACC project, especially T. Artières, P. Gallinari and C. Tijus.

6. References

Brézillon, P. (1994) Context needs in cooperative building of explanations. First European Conference on Cognitive Science in Industry. Luxembourg, pp. 443-450.

Brézillon, P. (2003) Focusing on context in human-centered computing IEEE Intelligent Systems, 18(3): 62-66.

Brézillon, P. (2005) Task-realization models in Contextual Graphs. Modeling and Using Context (CONTEXT-05), A. Dey, B.Kokinov, D.Leake, R.Turner (Eds.), Springer Verlag, LNCS 3554, pp. 55-68.

Brézillon, P. and Gonzalez, A.J. (2006) Tale of two context-based formalisms for representing human knowledge. In: M. Ali & R. Dapoigny: IEA/AIE 2006, Springer Verlag, LNAI 4031, pp. 137-145.

Brézillon, P. and Pomerol, J.-Ch. (1997a) Joint cognitive systems, cooperative systems and decision support systems: a cooperation in context. Proceedings of the European Conference on Cognitive Science. Manchester, UK, pp. 129-139.

Brézillon, P. and Pomerol, J.-Ch. (1997b) User acceptance of interactive systems: Lessons from Knowledge-Based and Decision Support Systems. Failures & Lessons Learned in Information Technology Management, June, 1(1): 67-75.

Brézillon, P. and Pomerol, J.-Ch. (1999) Contextual knowledge sharing and cooperation in intelligent assistant systems. Le Travail Humain, 62(3), Paris: PUF, pp 223-246.

Brézillon, P, Cavalcanti, M., Naveiro, R. and Pomerol, J.-Ch. (2000) SART: An intelligent assistant for subway control. Pesquisa Operacional, Brazilian Operations Research Society, 20(2): 247-268.

Brézillon, P., Pomerol, J.-Ch. and Pasquier, L. (2003) Chapter 6: Learning and explanation in a context-based representation: Application to incident solving on subway lines. In: R. Jain, A. Abraham, C. Faucher and J. van der Zwaag (Eds.) Innovations in Knowledge Engineering. International Series on Advanced Intelligence, pp. 129-149.

Brézillon, P., Brézillon, J. and Pomerol, J.-Ch. (2006) Decision making at a crossroad: a negotiation of contexts. Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering. ISBN 2-921145-58-8, pp. 2574-2583

Campbell, B. and Goodman, J. (1988) HAM: A General Purpose Hypertext Abstract Machine. Commun. ACM, 31(7) (1988) 856–861.

Clancey, W.J. (1992) Model construction operators. Artificial Intelligence Journal, 53, 1-115.

Compton, P., and Jansen, B. (1988) Knowledge in context. A strategy for expert system maintenance. In: J. Siekmann ed. Lecture. Notes in Artificial Intelligence Heidelberg: Springer-Verlag, pp. 37-49.

Forslund, G. (1995). Toward cooperative advice-giving systems. The expert systems experience, Ph. D. Thesis 518, Linkoping University, Sweden.

GADGET (1999). An EU-project "Guarding Automobile Drivers through Guidance, Education and Technology". WP 3 report. Editor Siegrist, S. BFU, Switzerland.

Gonzalez A.J.and Ahlers R.H. (1998), "Context-based representation of intelligent behaviour in training simulations", *Transactions of the Society of Computer Simulation*, 15(4), 153-166.

Grimshaw, D.J., Mott, P.L., and Roberts, S.A.(1997) The role of context in decision making: some implications for database design. European Journal of Information Systems, 6, 2 (1997) 122-128.

Gruber, T. (1991) Justification-based knowledge acquisition. In H. Motoda, R. Mizoguchi, J. Boose & B. Gaines (Eds.), Knowledge. Acquisition for Knowledge-Based Systems Amsterdam, The Netherlands: IOS Press (1991) 81-97.

Hendrix, G. (1975) Expanding the utility of semantic networks through partitioning. Proceedings of the Fourth IJCAI, pp. 115-121 (1975).

Henninger, S. (1992) The knowledge acquisition trap. Proceedings of the IEEE Workshop on Applying Artificial Intelligence to Software Problems: Assessing Promises and Pitfalls (CAIA-92). Monterey, Canada, March (1992).

Laird, J.E., Newell, A. and Rosenbloom. P.S. (1987) "SOAR: an architecture for general intelligence." Artificial Intelligence, 33: 1-64.

Leplat, J. and Hoc, J.M. (1983) Tâche et activité dans l'analyse psychologique des situations. Cahiers de Psychologie Cognitive, 3, 49-63.

McCarthy, J. (1993) Notes on formalizing context Proceedings of the 13th IJCAI 1, 555–560.

Pomerol, J.-Ch. (1997) Artificial Intelligence and Human Decision Making. European Journal of Operational Research, 99, 3-25.

Pomerol, J.-Ch. (2001) Scenario development and practical decision making under uncertainty. Decision Support Systems, 31, 197-204.

Pomerol, J.-Ch., Brézillon, P. and Pasquier, L. (2002) Operational knowledge representation for practical decision making. Journal of Management Information Systems, 18(4): 101-116.

Schank, R.C. (1982) Dynamic memory, a theory of learning in computers and people Cambridge University Press.

Sowa, J.F. (2000) Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks Cole Publishing Co., Pacific Grove, CA.

Vergnaud, G. (1985) Concepts et schèmes dans la théorie opératoire de la représentation. Les Représentation, Psychologie Française, 30(3) : 245-252.

Watson, I. and Perera, S. A. (2000) Hierarchical case representation using context guided retrieval. Knowledge Based Systems Journal. 11(5-6) : 285-292.

Woods, D.D., Roth, E.M., & Benett, K. (1990). Explorations in joint human-machine cognitive systems. In Robertson S. Zachary W. & Black J.B. (Eds.) *Cognition, Computing and Cooperation*. Ablex (pp. 123-158).

Adapting e-learning interfaces to personality type

W.J.G. Breuer — T.L. van der Weide — V. Dignum — F. Dignum

Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands
willi@mbreuer.nl, {tweide, virginia, dignum}@cs.uu.nl

ABSTRACT. *People have different styles of learning. Extensive research has been done on how to individualise learning material to better suit a person's learning style. Electronic Learning (e-learning) is used increasingly and provides new ways to adapt learning material to an individual. Personality theories give guidelines how to adapt material to an individual. As these guidelines allow many interpretations, the purpose of this paper is to understand how to implement these guidelines. An e-learning course was implemented that presented content following these guidelines. By implementing an e-learning course that presented information that followed these guidelines, we tested whether there was correlation between personality type and how well the information was learned*

KEY WORDS: *E-learning, MBTI Personality Type, ILS, Witkin's field (in)dependence, Multi Agent Systems*

1. Introduction

Human actors and software agents interact and cooperate increasingly often in a wide range of applications such as teaching, tutoring, and decision support. Interaction can be improved by taking into account the different ways people reason and react. Nowadays, most software systems are not aware of the cognitive processes of their human users. We think that understanding cognitive processes leads to effective solutions to support people. Our approach is tested by adapting two e-learning courses to the student's personality type. The hypothesis is that the adapted course is easier to learn.

People have different personalities, which affect their daily activities, emotions, the ways they interact, and how they learn. In Sociology and Psychology, instruments abound that enable the analysis and classification of personality and learning characteristics of people. In particular, the personality theory Myers-Briggs Type Indicator (MBTI), and the Index of Learning Styles (ILS) report human specific preferences for the learning and intake of information, motivation relevant in decision making, communication style, and planning style. MBTI and ILS state that communication adapted to personality preferences requires less effort and results in better understanding. Hence, our hypothesis is that the interaction between an actor and a software agent will be more efficient when the agent adapts its communication to personality preferences. However, the guidelines given by MBTI and ILS are rather abstract and leave a lot of room for interpretation and implementation. This study is a first step to understand how to implement personality and learning style theories in software. We have constructed an intelligent computer teaching system that adapts the way of giving information to a person on the basis of personality type and learning style. Future steps may include personality-based reasoning, automatic adaptation to personalities and machine learning of personality preferences.

This paper is organised as follows. Section 2 discusses several existing approaches towards optimised learning like the MBTI and ILS and related theories. Section 3 describes the implementation of an e-learning course that follows found guidelines to adapt material to an individual. Section 4 explains the experiment that we conducted. Results are presented in Section 5 and Section 6 provides conclusions and discussion.

2. Related work

In this section, we briefly introduce an existing theory aiming to categorise persons into personality types, namely the Myers-Briggs Type Indicator (MBTI). Furthermore, we describe the Index of Learning Styles (ILS) and the Witkin's field (in)dependence. These two learning-/cognitive style theories are well known models in education sciences for understanding learning preferences.

2.1. *Personality type instrument: Myers-Briggs Type Indicator (MBTI)*

MBTI (Myers *et al.,* 1985) is one of the most frequently used personality-type theories. Isabel Myers and Katherine Briggs began developing the MBTI around 1940 and based their work on C.G. Jung's theory of human personality that was published in 1921. MBTI is a questionnaire that classifies a person on four dichotomous dimensions, which together make up the personality type (Martin, 1997). The four dimensions are explained next.

– Extraversion (E) - Introversion (I)

Extraversion-types have a wide range of friends and know many people. They get energy from outwards and are eager to express their views. Introversion-types prefer to know just a few people well and get energy from inwards; they are reluctant to exhibit themselves socially. One might call these types respectively a "people person" vs. "reserved person". This dimension is also referred to as the complementary attitude (or orientation) of energy.

– Sensing (S) - iNtuition (N)

Someone that naturally gives more attention to information that comes in through his five senses is classified as Sensing. If more attention is paid to meanings, patterns and possibilities in information, a person is classified as an Intuition-type. Sensing-types pay much attention to facts and are prone to missing new possibilities. Intuition-types focus on the big picture, rather than finding out all facts. Also referred to as the mental perception function.

– Thinking (T) - Feeling (F)

Thinking-types put more weight on objective principles and impersonal facts; they enjoy technical and scientific fields where logic is important. They make decisions based on principles and logical consequences. Feeling- types put more weight on personal concerns and the people involved. Consequently they make decisions based on values and the consequences for people. Also referred to as the Mental judgment function.

– Judging (J) - Perceiving (P)

Judging-types prefer to be more structured (planned), are likely to come to conclusions quickly and prefer a decided lifestyle. Perceiving-types prefer a more flexible (spontaneous) and adaptable lifestyle. They enjoy the process of coming to closure, and are more comfortable being open-ended. Also referred to as the attitude towards the outer world.

Although MBTI looks fairly simple, Fleenor *et al.,* (2001) explains the complexity of the relations among the functions (S-N, F-T) and the attitudes (E-I)/(J-P). MBTI, if measured correctly, is still only a preference, everybody has a next strongest preference (the auxiliary function) and a least strong preference (the inferior function). Someone that scores, for example, Perceiving, can work on the development of his/hers Judging side, though he/she will always have a preference for the Perceiving-type.

Consider a person classified as ESTJ, meaning that he prefers Extraversion, Sensing, Thinking and Judging. Myers *et al.,* (1985) tell us that such a person prefers being practical, realistic, decisive, and prefers to quickly move to implement decisions. ESTJ's like to organise projects and people to get things done, focusing on getting results in the most efficient way possible. An ESTJ takes care of routine details and has a clear set of logical standards, which are followed systematically. In addition, ESTJ's are forceful in implementing their plans.

Many questionnaires are available to determine the MBTI personality type. The official questionnaire consists of 93 forced-choice questions and results in a personality. Keirsey (1998) provides an alternative questionnaire to discover the MBTI personality type. The Keirsey Temperament Sorter II uses 70 questions to measure the full MBTI type. For availability issues and ease of use we have used the Keirsey Temperament Sorter II during our experiment.

To confirm that MBTI can be used in learning we refer to Lawrence (1993), one of the most prominent researchers in the field of learning preferences. Lawrence reviewed 130 papers that used MBTI in studies of learning. He found that about one-fourth tested the learning style hypothesis; another fourth gave some indirect evidence of learning style preferences relating to MBTI. In addition, almost all of these researches support the theory as presented by Myers. In his book, he gives an overview of the learning preferences associated with the dimensions of the MBTI-types.

2.2. *Learning-/Cognitive style instruments*

Some learning-style theories, like the Experiential Learning Theory of Kolb (1984) followed Jung in recognizing that learning styles result from individuals' preferred ways for adapting (Kolb *et al.,* 1999). Felder *et al.,* (1988) even used Jung's Sensing/iNtuition dimensions as one of the dimensions in their learning and teaching styles framework `Index of Learning Styles` (ILS) which is one of the most popular en influential learning style models (Stash, 2007).

2.2.1. *Index of learning styles*

Felder states that students take in and process information in different ways. As examples he mentions seeing – hearing; reflecting – acting; reasoning logically - intuitively and analysing - visualising. Felder states that mismatches can lead to lower grades, as teaching methods can also vary (e.g. lecture or demonstration vs. self-discovery; or focusing on principles vs. application). ILS is a self self-scoring instrument that assesses the preferences of the student on four dimensions of the Felder-Silverman model. After a 44-items questionnaire, it indicates the preferences of the subject on the following dimensions (Felder *et al.,* 2007):

– Sensing learners - Intuitive learners

Concrete, practical, oriented toward facts and procedures vs. conceptual, innovative, oriented toward theories and meanings. This dimension is taken directly from MBTI (Sensing/iNtuition) and yields the same result (Felder *et al.,* 2005).

– Visual learners - Verbal learners

Indicates the preference between visual representations of presented material (e.g. pictures, diagrams, flow charts) vs. written and spoken explanations.

– Active learners - Reflective learners

Does the subject learn by trying things out, working with others, or does he learn by thinking things through, working alone. This dimension is related to Kolb's Active/Reflective, and MBTI's extravert/introvert.

– Sequential learners - Global learners

Prefers learning linear, orderly, in small incremental steps or learn holistic, in large leaps.

2.2.2. *Witkin's field (in)dependence*

Witkin *et al.,* (1977) developed a cognitive style model that defined two distinct styles: field dependence and field independence. To test which the style of persons, they are asked to perceive an object, without being influenced by a surrounding field. Each complex figure includes an embedded simple figure, which should be identified as quickly as possible. Field independent subjects are relatively uninfluenced, whereas the surrounding field influences field dependent people. The Witkin's Field Dependency experiments, like the Group Embedded Figures Test (GEFT; Oltman *et al.,* 1971), show that field independent people are able to work in more complex environments than field dependent people. Moreover, Lyons-Lawrence (1994) notes that students perform better when given instruction adapted to their style.

Research by both Davis (2004) and Corman *et al.,* (1988) links Witkin's research to MBTI. Both state that N-types are field independent and that most S-types are field dependent.

3. Designing e-learning environments

This section discusses how to design an e-learning environment using the MBTI and ILS. Several papers (Durling *et al.,* (1996), McCaulley *et al.,* (1980), Claxton *et al.,* (1994) and Claxton *et al.,* (1996)) report that the MBTI's Sensing/iNtuition dimension is most influential on the learning style. Consequently, our design focuses on the Sensing/iNtuition dimension.

Lawrence (1993) gives an overview of the off-line learning preferences associated with the dimensions of the MBTI types. For example, Sensing personalities prefer to learn using "Demonstrations, Films, TV". Dewar and Whittington (2000) translated Lawrence's theory into an e-learning course and confirmed Lawrence's ideas on this preference. Using Lawrence's research as starting point, we selected the following three user-interface characteristics to be adapted to personality and learning type:

– Word and text abstraction: concrete vs. abstract text;

– Illustrating type: animation vs. still image;

– Screen structure: one page vs. multiple pages.

3.1. *Concrete vs. abstract text*

Research shows that Sensing-types learn better using concrete texts, whereas iNtuition-types learn better using abstract texts (e.g. Kummerow *et al.,* 1997 and Dewar *et al.,* 2000). In online environments, Sensing-types appear to be very concrete in their communication, and dislike 'fancy' terminology and design (Fuchs, 2001). According to Berens (2006) and many others, Sensing-types prefer learning in a practical context and tend to dislike learning in an abstract context. Sensing-types prefer tangible realities, whereas iNtuition-types prefer thinking and talking about concepts and patterns.

Research by Jensen *et al.,* (1998) confirms that iNtuition-type students responded positively to abstract content and negatively to hands-on content. During the course restructuring at the US-Air force academy they discovered that Sensing-types responded in exactly the opposite manner. Dewar *et al.,* (2000) give a similar view: iNtuition-types dislike precise work with many details. Felder *et al.,* (2002) found comparable results: iNtuition-types performed significantly better than

Sensing-types in courses with a high level of abstract content. The contrary was observed in courses of a more practical nature.

3.2. *Animation vs. still image*

Van Oostendorp *et al.,* (2007) concluded that "animations are not preferable over `good old static graphics`" from an experiment that tested the effect of static versus animated diagrams on understanding, convincedness, and cognitive workload. However, as described earlier, Lawrence (1993) states that Sensing types would benefit when the content contains demonstrations like films. In other words, how much is learned from a text could benefit from an interface that is adapted to the student's personality type. Sutherland (1998), McCaulley (1990) and Jin *et al.,* (2000) confirmed the hypothesis that Sensing types prefer content containing demonstrations like films and that INtuition-types prefer reading and doing independent study (Lawrence, 1993).

3.3. *One page vs. multiples pages*

Lawrence (1993) explains that Sensing types like a sequential step-by-step approach and that iNtuition-types prefer to find their own way through material. Fuchs's research (2001) on SJ-types is saying that a system should provide a clear structure with little possibility for error. In conclusion, sensing-types prefer texts that are organised on one screen, where the writer dictates how to move through the content. INtuition-types prefer the content structured such that the user can decide how to go through it.

As mentioned before, Davis (2004) and Corman *et al.,* (1988) were able to link Witkin's model to MBTI. Davis (2004) and Corman *et al.,* (1988) suggest that iNtuition-types are field independent and that most Sensing-types are field dependent. It should seem logical that multiple screens (which is in our view a more complex environment) for iNtuition-types is supported the explanation of Field Independence. Handal *et al.,* (2004), who examined the concepts of field dependent and field independent cognitive styles within the context of computer-based instruction, confirm that it "appears to suggest that hypermedia learning environments, such as multimedia CD-ROMs and websites, provide an environment where Field Independent learners have more opportunities to succeed."

4. Initial experiments

Based on results described in section 3, we developed the hypothesis model that Sensing-type learners are able to learn better in an e-course if they are provided with animations and concrete text on a single screen; and that iNtuition-type learners learn better provided with still illustrations in addition with abstract text on multiple hyper-linked screens.

4.1. *Preparation of e-learning material*

We adapted two existing e-learning courses to model this hypothesis. The courses were originally developed to test the effect on understanding, convincedness, and cognitive workload on students between static or animated diagrams (Van Oostendorp *et al.,* 2007). One course concerned the study of the function of the human heart and the other the mechanics of a flushing toilet system. Although the courses were very simple their availability in our department and the fact that the material was in Dutch, were determinant to our choice for this material. However, based on the results we will describe below, our decision now would be not to use these courses. In the following, we will describe how the content and presentation of these courses were adapted to the characteristics and expected preferences of iNtuition and Sensing-types.

4.1.1. *Sensing-type interface*

Concrete changes in the interface optimised for Sensing-types consisted of the use of the exact part names when referring to something. Other changes include text that was fully re-written to be very concrete (e.g. using more examples). For the same reason, we adapted some of the words used.

The Sensing course contained a video demonstrating how all parts are linked together and how e.g. the rinsing tube works. The video did not use any audio (as in audio-visual as suggested by Golanty-Koel (1978)), but included some titles to explain what was displayed.

The Sensing-type interface also included a `table of contents` at the beginning of the course for navigation. It consisted of a single screen, with headers to indicate topic changes.

4.1.2. *Intuition-types interface*

The course interface for the iNtuition type contained a single (still) diagram and explanation in more abstract text form. There was only a single diagram as we were able to explain everything in just this single diagram. We achieved abstract texts by using variables as pointers to concepts. We also explained things differently, e.g.

like using terms such as ‘atmosphere in room A’ (siphon in toilet-course) instead of showing a clear image how the siphon works. Where possible we broke the information into multiple hyper-linked screens, though we tried to maintain enough useful information on the separate screens. We decided to link the pages with each other by using both the table-of-contents and some links within the text where this was suitable.

4.2. Measuring the student’s performance

The e-learning courses were designed in such a way that no prerequisite knowledge about the system or the material was needed. To be absolutely sure no prerequisite knowledge helped the subject, the test itself was done twice with two different e- learning courses, and prior-knowledge was measured with in total four multiple-choice and Likert-5 scaled questions/statements. To measure how well the student masters the content, the student made an exam consisting of three groups of questions: Retention of part names; Comprehension; Questions to test whether participants could solve problems with their new knowledge. There was a single question on the retention of part names, where all corresponding part names of a specific item in the course had to be given. Four multiple-choice questions followed to measure comprehension. Finally three open questions to test the ability to solve problems followed. Subjects were required to rate their confidence on all questions on a scale that ranges from zero to 100% scale, with an increment of 10% per step. At the end of the questionnaire, six mental effort-questions were posed. Subjects were asked to state their effort needed to learn the course, based on a 7 points Likert scale.

4.3. *Setup of the experiment*

The experiment went as follows. (1) The experiment started with an introduction that told our subjects the test was about an experiment to test 'study types' in order to prevent them from manipulating any possible outcome. (2) Next the Keirsey Temperament Sorter II (as MBTI questionnaire) had to be filled in. (3) Subsequently it was tested how much the subject already knew about the content to come by making the subject answer several questions. After a short break we presented the subject either with content optimised for his personality or not and measured the time (unnoticeable) until they felt confident they knew every little detail. The next step was to test what the subject had learned from the content using the 3 groups of questions (4) and a little questionnaire about how the subject perceived the material (5). Again after a small break the user was presented with the second course in the same mode as before (6) and the questions (7/8) followed. Finally we asked the participant to answer (9) the Felder's Index of Learning styles to measure X's learning style. Participants were native Dutch speakers and received a small fee for their participation.

5. Results

A total 61 of students from various educational backgrounds, aged between 18 and 30 years participated in the experiment (age mean = 21.9, standard deviation (SD) = 2.82). Almost all participants were students of academic level (Bachelor or higher level; 93.4%) or were studying on a University of professional education (Bachelor level; 4.9%). Just over half (55.0%) of the participants was female. See figure 1 for an overall view of the personality types.

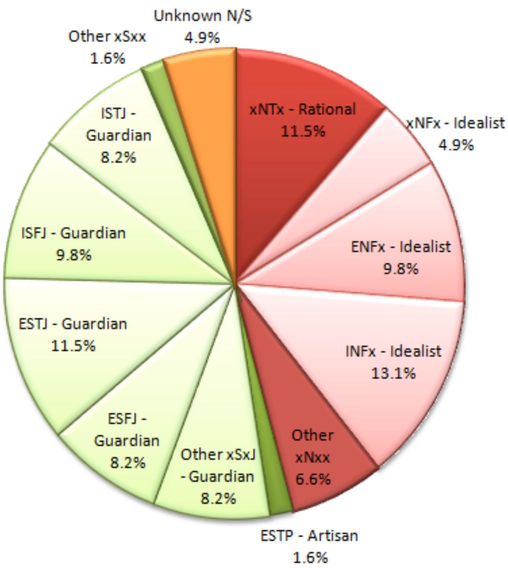


Figure 1. Participant distribution

The results show that MBTI and ILS measures were correlated along the Sensing/iNtuitive dimension. The encountered differences were not significant. ($X2=16.8$; $p<.01$).

Our hypothesis model states that video/animation improves MBTI Sensing-types' learning and that texts improve MBTI iNtuitive-types' learning. ILS measures this in the video/animation vs. verbal-dimension. During this initial experiment, more than 75% of our participants indicated they were Visual learners. The ILS Visual/Verbal-dimension seems unrelated to the MBTI S/N-dimension [$X2=.862$]. In this experiment, iNtuition-types outperformed Sensing-types independent of the interface used. The effect of the MBTI S/N-type on the score is significant $[F(1,53)=7.89; p<.01]$. The interface mode effect is not significant $[F(1,53)<1]$.

6. Discussion

The results show no relation between the presented interface and the performance of the test-subject (in general, nor of a specific personality type) when our hypothesis model and this specific e-learning course is used. This could mean that our S-interface is not implemented right, or that the literature is wrong or misunderstood. Furthermore, the results do not confirm that video/animation or text improves learning for S-types or N-types respectively. However, we can conclude that N- types have a higher average score, indicated a lower mental-effort and were (although not significantly) more confident. Godleski (1984) found that N-types - among engineering students - consistently outperformed sensors except in “real-world” courses like process design and cost estimation, in which the sensors did better. Our test could have been, unintentionally, specialised towards N-types due to the education material.

Another possible explanation for our results is that the topics of the e-learning courses were too simple to enable the subjects to make full use of the differences. We can clearly see the scores grades are focused above the mean and the reported mental effort on both courses is quite low. A more difficult course could show better use of the subjects’ differences in future research. A more complex e-learning course, which is not already optimised for a specific personality type, will probably give different results. The ILS-questionnaire suffices to determine the subject’s personality type, as the MBTI N/S dimension is clearly measured in the ILS N/S-dimension. The differences between word and text abstraction, illustration-type, and screen structure would be clearer if tested separately. The video/text-characteristic can better be determined using the ILS visual/verbal learner-dimension and not by the MBTI N/S-dimension. Including audio for sensing types, to fully create an audio-visual course, might increase performance and should be tested. In addition, creating a time-constraint on learning the courses could be considered in future research.

7. Conclusions and future work

This study is a first step to understand how to implement personality and learning style theories. This is required to answer our hypothesis that adapting communication to personality type improves interaction between a human actor and a software agent. The experiment, conducted on 61 subjects, shows that Sensing-types prefer concrete and visual materials, and iNtuition-types prefer abstract, but not verbal materials. Overall, the experiment shows that iNtuition-types score higher and require less mental effort than Sensing-types. There was no correlation between the confidence that the subjects had in their answers, and the interface mode and the time it took to complete the test. The results do not confirm the hypothesis, while other studies do confirm our hypothesis. Some possible explanations are that the openness for interpretation of the MBTI and ILS guidelines allowed us to make wrong implementation decisions, that the content of our tests was biased towards the iNtuition-type, or that the content was too simple to cause significant differences.

The initial experiment described in this paper, shows that following the intuitive guidelines provided by the personality and learning classification of MBTI and ILS do not necessarily result in more effective interaction between human actors and software agents. This is in contradiction with other studies that do report improved interaction. To further confirm the applicability of personality theories, deeper investigation is needed of how these guidelines can be formalised such that they can be implemented successfully. In specific, we will pursue the approach taken in this paper on more complex learning domains, which provide a richer differentiation between students. Another issue to consider is the automatic recognition of personality preferences, how to reason using personality-based concepts, and how to design reliable and reproducible experiments that test whether changes lead to actual improvement.

In this paper, we explored one aspect of the interaction between agent and humans, namely that of the differences in interface in learning environments. There are many applications, which could be improved by MBTI and ILS. Examples are e-learning, (serious) games, decision support, and cooperation between human actors and software agents. In general, agent designers are concerned on overall strategies for successful argumentation, including cooperativeness, competitiveness, explanation, and reaction speed. In the future, we intend to extend this research to include other aspects of argumentation, in special applied to training and decision support applications. In addition, it will be important to consider continuous interaction and automated adaptation to changing circumstances. We believe that interactions between people and agents will increase in number and quality, and therefore it is important to identify the means that will make such interactions more personal and satisfying for people, that is, approaching the way human-human interactions occur.

References

Berens L., *Understanding Yourself and Others: An Introduction to the 4 Temperaments 3.0.*, Telos Publications, 2006.

Claxton R., McIntyre J., “Empirical relationships between need for cognition and cognitive style: implications for consumer psychology”, *Psychological reports,* vol. 74 no. 3, 1994, p. 723-732.

Claxton R., McIntyre J., Clow K., Zemanek, J., “Cognitive style as a potential antecedent to values”, *Journal of Social Behaviour and Personality*, vol. 11, 1996, p. 355-373.

Corman L., Platt R., “Correlations among the Group Embedded Figures Test, the Myers-Briggs Type Indicator and demographic characteristics: A business school study”, *Perceptual and Motor Skills*, vol. 66, 1988, p. 507-511

Davis G., The relationship between learning style and personality type of extension community development program professionals at The Ohio State University (Dissertation), 2004.

Dewar T., Whittington D., “Online Learners and their Learning Strategies”, *Journal of Educational Computing Research,* vol. 23, no. 4, 2000, p. 415-433.

Durling D., Cross N., Johnson J., “Personality and learning preferences of students in design and design-related disciplines”, *International Conference on Design and Technology Educational Research*, 1996

Felder R., Felder G., Dietz E., “The Effects of Personality Type on Engineering Student Performance and Attitudes”, *Journal of Engineering Education,* vol. 91, no. 1, 2002, p. 3-17.

Felder R., Silverman L., “Learning and Teaching Styles In Engineering Education”, *Engineering Education*, vol. 78, no. 7, 1988. p. 647-681

Felder R., Soloman B., Index of Learning Styles, http://www.ncsu.edu/felder-public/ILSpage.html, 2007

Felder, R., Spurlin, J., “Applications, Reliability and validity of the Index of Learning Styles”, *International Journal of Engineering Education,* vol 21, no 1, 2005, p. 103-112

Fleenor, J., Mastrangelo P., Test review of the Myers-Briggs Type Indicator Form M., 2001

Fuchs R., “Personality Traits and their Impact on Graphical User Interface Design: Lessons Learned from the Design of a Real Estate Website*”, 2nd workshop on attitude, personality and emotions in user-Adapted Interaction*, 2001

Godleski E., “Learning Style Compatibility of Engineering Students and Faculty”, *Proceedings Annual Frontiers in Education Conference,* 1984

Golanty-Koel R., The relationship of Psychological Types and Mass Media Preferences to the values of non-academic high school students (dissertation), 1978. In Wilkins (1996).

Handal B., Herrington A, “On being Dependent or Independent in Computer Based Learning Environments”, *E-Journal of Instructional Science and Technology*, vol. 7, no. 2, 2004.

Jensen D., Murphy M., Wood K., “Evaluation and Refinement of a Restructured Introduction to Engineering Design Course Using Student Surveys and MBTI Data”, *Proceedings of the ASEE Annual Conference,* 1998.

Jin B., Austin D., “Personality types of therapeutic recreation students based on the MBTI”, *Therapeutic recreation journal First Quarter,* 2000.

Keirsey D., *Please understand me II. Temperament Character Intelligence,* Prometheus Nemesis Book Company, 1998

Kolb D., *Experiential Learning: Experience as the Source of Learning and Development,* Prentice-Hall, 1984

Kolb D., Boyatzis R.E., Mainemelis C., “Experiential Learning Theory: Previous Research and New Directions”, Perspectives on Thinking, Learning, and Cognitive Styles, Lawrence Erlbaum Associates, 2000

Kummerow J., Barger N., Kirby L. Worktypes, Warner Books, 1997

Lawrence G., *People types & tiger stripes (Third edition),* Center for Applications of Psychological Type, 1993

Lyons-Lawrence C., Effect of Learning style on performance in using computer based instruction in office systems, 1994. In Wilkins (1996).

Martin C., *Looking at Type: The fundamentals,* Center for Applications of Psychological Type 1997.

McCaulley M., Natter F., *Psychological (Myers-Briggs) type differences in education*, Center for Applications of Psychological Type Inc, 1980

McCaulley M., “The MBTI and Individual Pathways in Engineering Design”, *Engineering Education*, vol. 80, 1990.

Myers I., McCaulley R., *Manual: A Guide to the Development and Use of the Myers-Briggs type indicator*, Consulting Psychologists Press, 1985

Oltman P.K., Raskin, E., Witkin H.A., *A manual for the embedded figures tests,* Consulting Psychologists Press, 1971

Oostendorp, van H., Beijersbergen M., Animated diagrams: Their effect on understanding, confidence and mental effort, 2007

Stash N., Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System. SIKS Dissertation Series, no. 11, 2007.

Sutherland V., Are you the type for e-learning? An ABI white paper., IBM Advanced Business Institute, 1998.

Witkin H., Moore C., Goodenough D., Cox, P., “Field-Dependent and Field-Independent Cognitive Styles and Their Educational Implications”, *Review of Educational Research*, vol. 47, no. 1, 1977, pp. 1-64.

Wilkins D., “Are We Using the Wrong Teaching Method in Our Criminal Justice Classes”, *Journal of criminal justice education*, vol. 7, 1996.

Younis N., Salman R. Ashrafi R., Efficacy of Present e-learning Content to Student Personality Types, International journal of information technology, vol. 1, no 3, 2004

Basic concepts for human / robot shared authority

Stéphane Mercier* — **Catherine Tessier*** — **Frédéric Dehais**** — **Charles Lesire***

* *ONERA*
2 avenue Edouard Belin
BP 4025 - 31055 Toulouse Cedex 4
{firstname.surname}@onera.fr

** *ISAE*
10 avenue Edouard Belin, BP 54032 - 31055 Toulouse Cedex 4
frederic.dehais@isae.fr

RÉSUMÉ. *Dans le cadre de la supervision de mission d'agents artificiels (robots, drones...) par un opérateur humain, la question du partage des rôles et de l'autorité est une problématique avérée. Entre le contrôle manuel de l'engin par l'opérateur qui n'est pas infaillible et l'autonomie totale souvent irréalisable, un équilibre doit être trouvé. De plus, l'action de différentes entités sur des objets communs requiert une coordination et est souvent source de conflits. Nous présentons ici les concepts élémentaires d'une approche visant à ajuster dynamiquement l'autonomie d'un agent par rapport à un opérateur, sur la base d'une modélisation formelle des constituants de mission.*

ABSTRACT. *In the context of supervisory control of artificial agents by a human operator, the definition of the autonomy of an agent remains a major challenge. A trade-off must be found between manual control by the operator who is fallible and full autonomy of the agents which is often hard to achieve. Additionally when an agent and a human decide and act simultaneously using the same resources, conflicts are likely to occur and coordination is mandatory. We present the basic concepts of an approach that will aim at dynamically adjusting the autonomy of an agent relatively to its operator, based on a formal modelling of mission ingredients.*

MOTS-CLÉS : *Autonomie adaptative, Partage d'autorité, Systèmes multiagents, Interactions homme-machine.*

KEYWORDS: *Adaptive autonomy, Authority sharing, Multiagent systems, Human-Machine interactions.*

1. Introduction

While there is no universal definition of autonomy, this concept can be seen as a relational notion between entities about an object (Castelfranchi *et al.*, 2003; Brainov *et al.*, 2003) : for instance, a subject X is autonomous with respect to the entity Z about the goal G. In a social context entities like other agents or institutions may influence a given agent thus affecting its decision-making freedom and its behaviour (Carabelea *et al.*, 2006).

In the context of a physical agent evolving in the real world (i.e. an uninhabited vehicle) under the control of a human operator, autonomy can be seen as the ability of the agent to minimize the need of human supervision and to act alone (Schreckenghost *et al.*, 1998) : the primary focus is then rather the operational aspect of the autonomy than the social one. In this situation pure autonomy is just a particular case of the agent - operator relationship, precisely consisting in not using this relationship.

However in practice, as automation within complex missions is not perfectly reliable and is usually not designed to reach the defined objectives alone, human supervision is still needed. Moreover it seems that human intervention significantly improves performance over time compared to a neglected agent (Goodrich *et al.*, 2001; Goodrich *et al.*, 2007) (see Fig. 1).

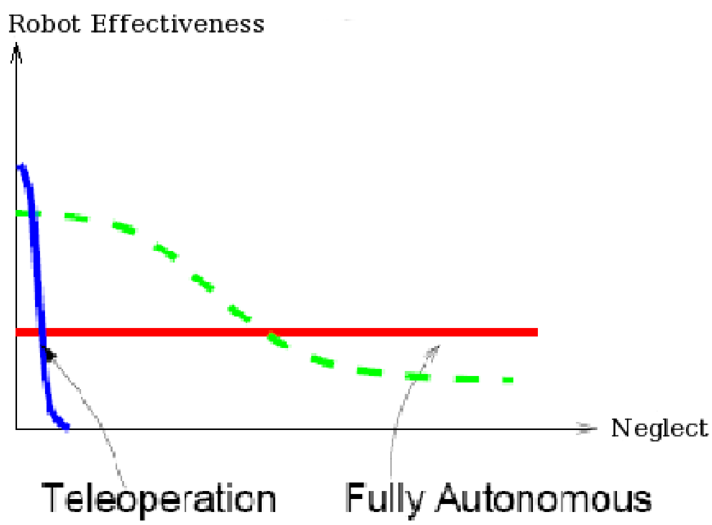


Figure 1. *performance vs neglect*

Adjustable Autonomy

(Sheridan *et al.*, 1978) first proposed a classification for operational autonomy based on a ten-level scale. This model remains quite abstract as it takes into account neither the environment complexity nor the mission context. However it provides an interesting insight into the interactions between an operator and an agent. This model has later been extended, using the same scale applied to a four-stage cognitive information processing model (perception, analysis, decision-making and action) (Parasuraman *et al.*, 2000). Based on the same principles other scales for autonomy classification have also been proposed, e.g. (Bradshaw *et al.*, 2003).

Other approaches aim at evaluating an agent's autonomy in a given mission context, like MAP (Hasslacher *et al.*, 1995), ACL (Clough, 2002) or ALFUS (Huang *et al.*, 2005). The latter proposes to evaluate autonomy according to three aspects : mission complexity, environmental difficulty and human interface. However this methodology aggregates many heterogeneous metrics and the meaning of the result is hard to evaluate. Moreover a qualitative step is necessary especially to set weights on the different tasks composing a mission and evaluate their importance.

The idea that operational autonomy can be graduated leads to the concept of adjustable autonomy or shared authority. The main principle is that machine and human abilities are complementary and they are likely to provide better performance when joined efficiently than when used separately (Kortenkamp *et al.*, 1997). A physical agent is thus capable of evolving at several predefined autonomy levels and switches levels according to the context. A level is defined by the complexity of the commands (Dorais *et al.*, 1999) or the ability to perform tasks without the need of operator's interventions (Goodrich *et al.*, 2001). The major limitation we can see in these approaches is the *a priori* definition of the levels, the static distribution of tasks among entities at each level and the fact that the number of levels is necessarily limited. Interactions between the agent and the operator are thus restricted to a given set and are determined by autonomy levels, there is no possibility of fine dynamic task sharing.

To add more flexibility, (Scerri *et al.*, 2003) endow agents with learning capabilities based on Markov Decision Processes (MDP) allowing them to better manage the need for human intervention. Agents can define their own autonomy levels from the user's provided intentions. However this method does not seem to be directly applicable to critical systems as the behaviour of learning agents facing unexpected situations is hard to validate. Moreover the operator's interactions are restricted to the agent's needs.

The approach of (Myers *et al.*, 2001) adds more human control on the agent. Levels are not defined in a static way but come from a norm : permissions and restrictions describing the agent's behaviours are set by the operator. In order to do so, the operator has to create a complete set of rules like ”In case of medical emergency, consult the operator to choose landing location”. The major issues associated with such an approach are the high number of rules to provide and the risk of conflict between rules. Anyway the autonomy of the agent is completely human-supervised and the agent has no possibility to adapt by itself.

Sliding autonomy(Brookshire *et al.*, 2004) consists in determining whether a task should be executed by the agent alone or by the operator using manual control ; there is no direct reference to autonomy levels. Roles are not shared at the mission level but are reconsidered for each action to realize. However it seems that the range of human-agent interactions is really restricted as each task is performed either ”completely autonomously” or ”completely through teleoperation”.

In contrast, collaborative control is an approach aiming at creating dialogs between the operator and the agent (Fong *et al.*, 2002) : the agent sends requests to the human operator when problems occur so that she/he could provide the needed support. This is again a restriction of all possible interactions : only dialog is used whatever the circumstances. In practice almost all interactions are initiated by the agent's requests and the operator acts almost exclusively as a support, she/he has not much initiative.

(Sellner *et al.*, 2006) have studied two authority sharing modes on a simulated space assembly task, SISA (System-Initiative Sliding Autonomy) where only the agent can request the operator's support and MISA (Mixed-Initiative Sliding Autonomy), where the operator can also intervene anytime. The allocation between the agent and the operator is realized separately for each task according to statistics to determine which entity will be the most efficient, which does not seem sufficient for a critical mission where errors are not allowed. However sharing at the task level is an interesting idea as it provides the most adaptive solution to the mission.

As shown by the literature review it is often interesting to join human and machine abilities to carry out a mission and adjustable autonomy seems a good principle. However the fact that the human operator also is fallible is often neglected. While it seems reasonable that the operator should keep the control over the agent, in most of the studies the operator's inputs are not evaluated and accepted “as they are” by the agent. Moreover the simultaneous decisions and actions from an artificial agent and a human agent might create misunderstandings and lead to conflicts and dramatic situations (Dehais *et al.*, 2005).

2. Context of the Study, Assumptions and Objectives

In this paper we focus on the autonomy of artificial agents (e.g. uninhabited vehicles, autopilots. . .) supervised by a human operator and achieving objectives for a given mission. Such agents evolve in a dynamic environment and face unexpected events. Consequently real-time reactions to these events in order to avoid dangerous situations and the loss of the agents themselves are compulsory. Additionally we consider systems where most of operational tasks can be associated with procedures, i.e. tasks must be executed in a precise order and respect strict constraints (as it is the case in aeronautics).

In an ideal context the agents would be able to achieve the mission completely independently from the operator, a case that is hardly likely to occur in reality. However this is a necessary ability for the agents as communication breakdowns between the agents and the operator may occur during the mission. Beyond this extreme case the agents may request the operator's help anytime for any task when an issue arises. On the other hand the operator her/himself must be free to intervene at any stage of the mission in order to adjust the agents'behaviours according to her/his preferences but also to correct their possible mistakes or improve their performance.

One of the main challenges is conflicts. The human operator's inputs may interfere with the agents'plans and break their consistency anytime, even if the inputs are intended to improve a task or to correct an agent's mistake. As an agent and the operator may both execute actions on their own, it is of great importance that they should remain coordinated so that they should not use the same resources at the same time for different purposes. For example if the autopilot of a UAV [1] and the operator simultaneously "decide" to move the vehicle in different directions, inconsistencies are very likely to appear in the flight and lead to an accident. Therefore conflicts must be detected and solved as soon as possible.

Finally our main objective can be summarized in the following question : why, when and how should an agent take the initiative ? When the environment has changed and the agents'plan needs to be updated ? When the operator's inputs are inconsistent with the procedures and with security constraints ? Or when they create conflicts with the current goals ?

3. Main functions

Planning and situation assessment are the two key "high-level" functions for autonomous agents. However, a third function, the authority sharing function, is proposed

1. Uninhabited Air Vehicle

so as to manage the interactions between an operator and an agent. This function will be based on conflict solving.

3.1. *Planning*

A mission consists in a set of *objectives* the agents should reach. To do so the agents will execute *tasks*, each task being supposed to provide an expected result while respecting some *constraints* (security, physical limits, authorizations, etc.) Each task that is executed uses and produces *resources*, and is a resource itself to reach an objective.

As the agent has to react to unplanned events occurring during the mission, the plan has to be continuously updated. This replanning process is a mandatory ability of the agent and is requested by the authority sharing function.

3.2. *Situation Assessment*

Situation assessment (Lesire *et al.*, 2006) constantly analyzes the state of the system : the current and possible future states are estimated according to the plan and procedures and to the evolution models of the environment, of the system itself and of all other relevant objects.

As far as the recognition of the operator's intentions are concerned, the only available information comes from her/his inputs in the system. If a pattern is recognized from these inputs and can be associated with one or several known procedures, this constitutes a valuable clue about the operator's non-explicit goals and may contribute to anticipate her/his future actions.

More precisely situation assessment compares the expected results of the tasks performed by the agents and the operator with the actual results and detects gaps that may appear. This allows potentials conflicts to be detected in the current state but also in an anticipated manner.

A conflict is a mismatch between the plan and its execution appearing as an inconsistency at the level resource. The situation assessment function identifies the conflicting resources and determines the characteristics of the conflict (involved entities, occuring time of the conflict, etc.)

Finally, this information is transmitted to the authority sharing function.

3.3. *Authority Sharing*

All issues in the execution of the mission correspond to a type of conflicts in the allocation of resources, between the entities or between the plan and its execution.

The authority sharing function gets information about conflicts from the situation assessment function : depending on the category of the conflict, appropriate solving methods can be executed. In any case, if a solution for a conflict exists, this is always realized through resource reallocation among the agents.

This can be done through the planning function. From the information coming from the situation assessment, the authority sharing function may add constraints to the planning process so that a new consistent allocation of resources can be found.

4. Basic concepts

The first step to get a formal and operational definition of adaptive autonomy is to formalize the basic concepts of a mission operated by physical agents and an operator.

Mission
A mission is a set of objectives to be reached by the agent(s) and the human operator.
Example :
M = { go to zone, detect fires, return to base }.

Resource
A resource is an item contributing to satisfying a mission objective. It can be a physical object, energy, a permission, a piece of information, a task, an algorithm, a logic condition...
A resource is written :
$r =< item\_id, type, time\_interval, value[time\_interval], \mathcal{R}_{cons}, \mathcal{R}_{prod}, src >$,
with $item\_id$: the identifier of the resource ;
$type$: the type of the resource (additive or absolute, exclusive, task, etc.) ;
$time\_interval = [t_{start}, t_{end}]$: the time interval defining the existence of the resource ;
$value[time\_interval]$: a set of dated discrete values taken by the resource on $time\_interval$;
$\mathcal{R}_{cons}$: the set of resources consumed or needed by this resource ;
$\mathcal{R}_{prod}$: the set of resources produced or affected by this resource ;
and src : the origin of the resource (see definition of $source$ below).

Example :
$r1 =< energy, additive, [3h05m17s - 3h06m20s], [1, ..., 1], \{battery\}, \{\}, src >$
with $src =< engine, 3h03m02s, Emaxx1 >$ (see definition of $source$ below).

Source
A source defines the origin of a resource.
A source is written :
$source =< r_{prod}, t_{prod}, a >$
with r_{prod} the producing resource ;
t_{prod} the production time ;
and a the producing entity (agent, operator...).

Example :
$source1 =< navigation1, 10h20m50s, Emaxx1 >$
$source2 =< piloting2, 11h42m20s, Emaxx1 >$

Tasks
Tasks are particular resources. They are created and scheduled by the planning algorithm in order to satisfy mission objectives.

Example :
Let $nav1$ be a task realized by the robot on operator's request. This resource is written :
$$nav1 =< navigating, task, [t_{start} - t_{end}], initiated, \{map, navAlgorithm\}, \{waypointsList\}, src >$$
with $\{map, navAlgorithm\}$ the resources needed to perform task $nav1$;
$waypointsList$ the resource produced by task $nav1$;
and $src =< GUIRequest, t_{prod1}, operator >$ specifying the origin of $nav1$, a request from the operator through the graphical user interface at t_{prod1}.
It is only when resources consumed by task $nav1$ have been allocated over time that this task takes the value $instantiated$ and its times t_{start} and t_{end} are set in the plan.
A *task* resource can take the following values :
$\{initiated, instanciated, executing, done, aborted, paused, failed\}$.

5. Experimental Environment and Scenario

Our concepts for adaptive autonomy will need further theoretical and software developments. Nevertheless, the framework for experimentations in real conditions with human operators interacting with "autonomous" vehicles is already being designed.

5.1. *The Scenario*

The scenario is the localization and assessment of a fire by a UGV [2] in a partially unknown area.

2. Uninhabited Ground Vehicle

The mission for the UGV and the operator consists in looking for starting fires around a factory or a facility and determining its properties (localization, size, dynamics) so that it could be quickly put out. The area is hardly accessible, dangerous and partially unknown (no precise and updated map available). Additionally, the scenario could be extended with the possibility for the UGV to carry an extinguisher. This would allow the UGV to directly put out a very starting fire or delay a fire evolution in a given area, e.g. close to sensitive items. As the extinguisher would be very small, its use would have to be carefully chosen. Figure 2 shows the scenario.

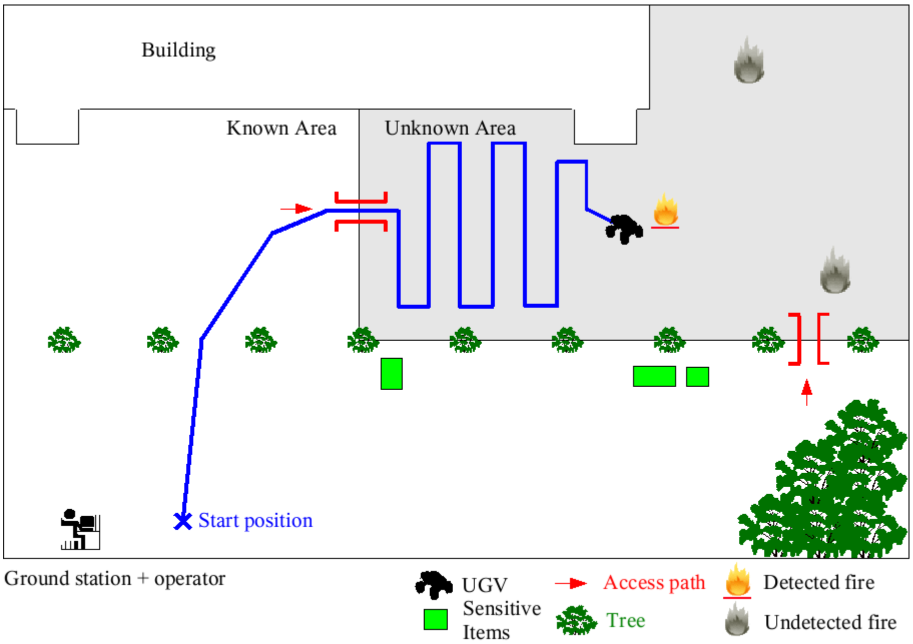


Figure 2. *the scenario*

Several operational assumptions are made :

– The area where the UGV evolves is divided into two parts : the start area which is known (a map is available), the search area which is partially unknown ;

– the known area includes obstacles to avoid, but there are localized on a map ;

– the human operator has no direct visual contact with either the UGV nor the outdoor environment ;

– there are sensitive items in the known area, which have to be protected against the fire threat coming from the partially unknown area ;

– the fires may evolve, possibly blocking known paths or endangering the UGV ;

– a fire evolution is determined by the objects that can burn ;

– the access paths to the search area are limited and narrow, making the access to the zone difficult.

The roles of the agents are the following :

The UGV :

– detects and locates a starting fire in a partially unknown area ;

– identifies the fire : gets its size and assesses the threat by getting closer ;

– puts out starting fires or delays a fire evolution by using the fire extinguisher ;

– updates the map of the partially unknown area to be able to get back to the start base ;

– avoids obstacles and fires, manages its resources, preserves its own existence.

The operator :

– supervises the UGV in real time and monitors its decisions / actions ;

– corrects or improves the UGV decisions / actions ;

– detects, identifies and possibly puts out starting fires by interacting with the UGV ;

– makes sure the UGV is able to come back to the start station.

Finally some hazards may impair the mission :

– communication breakdowns between the UGV and the operator ;

– dynamic and uncertain environment in the search area (obstacles, fires) ;

– possible loss of GPS positioning ;

– sensor failure.

5.2. *The Experimental Set-up*

ISAE is developing an experimental set-up composed of a ground station and several Emaxx UGVs - see Fig. 3 and 4.

The UGVs may be controlled using a remote control (in case of problems) or a graphical interface (normal use), but are also capable to rejoin a set of waypoints autonomously thanks to their sensors (GPS, inertial sensors, odometry). Algorithms are currently being developed to be implemented onboard (ARM 7 & 9 electronic cards, Linux OS) in order to give them decisional abilities (planning, situation assessment).

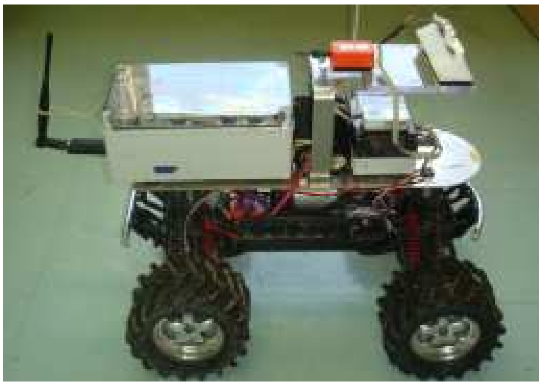


Figure 3. *an Emaxx UGV*

Figure 4. *the current control station for the Emaxx*

A wizard of Oz user interface is also being developed as it offers greater possibilities to control ”unexpected” events during the experiments (e.g. : communication breakdowns, sensor failures).

6. Conclusion and Future Work

We have presented the general principles and some basic concepts for an approach of operational adaptive autonomy. Resources, including the operator's tasks, are the key items to determine the contribution of each entity to the mission's objectives. Conflicts can be detected and classified depending on the entities that disrupt the plan. Consequently task reallocation within the system is performed so that conflicts could be solved safely with every entity being aware of what is being done.

Task reallocation will take into account the current capacities of the agents and operators, the operators'desires, the mission constraints, the priorities of objectives, all this being aggregated by the authority sharing function and transmitted to the planning function. Early conflict detection will allow agents to adapt their behaviours to the estimated operator's intentions as long as main constraints and objectives are respected, therefore improving the overall system performance. However, whether the operator intervenes or not, the agents are still expected to have the means to react "alone" to key issues.

Another aspect of adaptive autonomy is the fact that agents should be able to alleviate the operator's workload, e.g. relieving her/him of routine tasks and let her/him focus on key tasks of the mission. Again this is based on mutual situation monitoring and assessment and a better allocation of resources (including tasks) within the system when the context changes.

Current work focuses on a formal definition of mission execution including the dynamic aspects of the basic concepts we have defined, particularily conflicts, and on the fine identification of what precisely is involved in task reallocation. These concepts have to be operationalized then implemented on a real UGV platform (Emaxx) in order to conduct experiments with human operators [3]. Reliability, overall performance and the operator's satisfaction will allow us to assess our concepts for adaptive autonomy in real conditions.

7. References

Bradshaw J., Sierhuis M., Acquisti A., Feltovich R., Hoffman R., Jeffers R., Prescott D., Suri N., Uszok A., Van Hoof R., « Adjustable autonomy and human-agent teamwork in practice : An interim report on space application », in *Agent Autonomy* (Hexmoor *et al.*, 2003), chapter 11, 2003.

Brainov S., Hexmoor H., « Quantifying Relative Autonomy in Multiagent Interaction », in *Agent Autonomy* (Hexmoor *et al.*, 2003), chapter 4, 2003.

3. see section 5

Brookshire J., Singh S., Simmons R., « Preliminary Results in Sliding Autonomy for Coordinated Teams », *Proceedings of the AAAI 2004 Spring Symposium*, Stanford, CA, 2004.

Carabelea C., Boissier O., « Coordinating agents in organizations using social commitments », *Electronic Notes in Theoretical Computer Science, Elsevier*, 2006.

Castelfranchi C., Falcone R., « From Automaticity to Autonomy : the Frontier of Artificial Agents », in *Agent Autonomy* (Hexmoor *et al.*, 2003), chapter 6, 2003.

Clough B. T., « Metrics, Schmetrics ! How The Heck Do You Determine A UAV's Autonomy Anyway ? », *Proceedings of the Performance Metrics for Intelligent Systems Workshop*, Gaithersburg, Marylan, 2002.

Dehais F., Goudou A., Lesire C., Tessier C., « Towards an anticipatory agent to help pilots », *Proceedings of the AAAI 2005 Fall Symposium "From Reactive to Anticipatory Cognitive Embodied Systems"*, Arlington, Virginia, 2005.

Dorais G., Bonasso P., Kortenkamp D., Pell B., Schreckenghost D., « Adjustable autonomy for human-centered autonomous systems », *IJCAI'99 Workshop on Adjustable Autonomy Systems*, Stockholm, Sweden, 1999.

Fong T., Thorpe C., Baur C., « Collaboration, dialogue and human-robot interaction », *10th International Symposium on Robotics Research*, Lorne, Victoria, Australia, 2002.

Goodrich M., McLain T., Crandall J., Anderson J., Sun J., « Managing Autonomy in Robot Teams : Observations from four Experiments », *Proceeding of the ACM/IEEE international conference on Human-robot interaction*, 2007.

Goodrich M., Olsen D., Crandall J., Palmer T., « Experiments in adjustable autonomy », *Proceedings of Ijcai 2001 Workshop on Autonomy, Delegation and Control : interacting with autonomous agents*, 2001.

Hasslacher B., Tilden M. W., « Living Machines », , Los Alamos National Laboratory, 1995.

Hexmoor H., Castelfranchi C., Falcone R., *Agent Autonomy*, Kluwer Academic Publishers, 2003.

Huang H., Pavek K., Novak B., Albus J., Messin E., « A Framework For Autonomy Levels For Unmanned Systems (ALFUS) », *Proceedings of the AUVSI's Unmanned Systems North America 2005*, Baltimore, Maryland, June, 2005.

Kortenkamp D., Bonasso P., Ryan D., Schreckenghost D., « Traded control with autonomous robots as mixed initiative interaction », *Proceedings of the AAAI 1997 Spring Symposium on Mixed Initiative Interaction*, March, 1997.

Lesire C., Tessier C., « A hybrid model for situation monitoring and conflict prediction in human supervised "autonomous" systems », *Proceedings of the AAAI 2006 Spring Symposium "To Boldly Go Where No Human-Robot Team Has Gone Before"*, Stanford University, California, 2006.

Myers K., Morley D., « Human directability of agents », *K-CAP 2001, 1st International Conference on Knowledge Capture*, Victoria, Canada, 2001.

Parasuraman R., Sheridan T., Wickens C., « A model for types and levels of human interaction with automation », *Systems, Man and Cybernetics, Part A, IEEE Transactions*, vol. 30, n° 3, p. 286-297, 2000.

Scerri P., Pynadath D., Tambe M., « Adjustable Autonomy for the Real World », in *Agent Autonomy* (Hexmoor *et al.*, 2003), chapter 10, 2003.

Schreckenghost D., Ryan D., Thronesbery C., Bonasso P., Poirot D., « Intelligent control of life support systems for space habitat », *Proceedings of the AAAI-IAAI Conference*, July, 1998.

Sellner B., Heger F., Hiatt L., Simmons R., Singh S., « Coordinated Multi-Agent Teams and Sliding Autonomy for Large-Scale Assembly », *Proceeding of the IEEE, 94(7)*, 2006.

Sheridan T., Verplank W., Human and Computer Control of Undersea Teleoperators, Technical report, MIT Man-Machine Systems Laboratory, Cambridge, MA, 1978.

“The Human Side of Decision” session

11 – 12:30 am

Wednesday, June 11, 2008

Emotion, personality and decision-making

Relying on the observables

Gilles Coppin*

* *Telecom Bretagne*
CS83818, 29238 Brest Cedex 3
gilles.coppin@telecom-bretagne.eu

ABSTRACT. *This paper proposes a way to address the quantitative description of decision strategies for expert decision-makers in order to take into account the effects of personality, emotion or mood on decision-making. Most of common taxonomies used in popular models to decribe user profiles and behaviors seem to be difficult to apply in empirical cases. Instead we propose to rely on computational models inspired by cognitive psychology. These models and the related methodology allow to extract meaningful data structures from the behaviors of decision-makers. This data can be used to propose robust definitions of decision-styles. We then discuss the impact of this refined modeling on decision-support system functionalities.*

RÉSUMÉ. *Cet article présente une approche permettant de décrire les effets de facteurs personnels tels que émotion ou personnalité sur les processus de décision. La plus grande partie des taxonomies utilisées dans les approches traditionnelles du domaine semblent en effet difficiles à appliquer en pratique du fait qu'elles ne prennent pas en compte les comportements observables des décideurs. Nous proposons précisément d'appuyer la modélisation et la compréhension des décideurs sur la mise en évidence de structures régulières de comportements, qui permettent de définir de façon plus robuste la notion de style de décision.*

KEYWORDS: *Decision models, emotion, personality, cognitive modelling*

MOTS-CLÉS : *Décision, émotion, personnalité, modèles cognitifs*

1. Introduction

This paper is devoted to the modeling and analysis of *individual differences* in decision-making. These differences usually include some features related to the personality of the decision maker and others features related to his/her emotional states. It seems that these two kinds of features have usually been studied separately, so that it remains difficult to understand or predict what could be the real effect of combining a given personality profile with a transient emotional state. This paper strives to develop a unified representation for the categories of decisional behaviors using a cognitive model of decision making. Moreover, we show how such a unified representation could lead to a new definition of *decision styles*. Decision styles are of the outmost importance for the design of decision-aid systems.

The paper is structured as follows: first, we present a brief survey of personality and emotion theories as related to decision-making, from the psychological, cognitive and computational perspectives. Second, we present a set of models of decision-making that seem to have been often ignored by previous studies and that we find useful to describe individual decision-making behaviors and strategies. We especially try to identify which parts of these models can support the expression of individual differences. Lastly, we present the elements of a methodology that is adapted to our approach and some preliminary results. We briefly discuss potential applications of these modeling approaches to decision-aid systems.

2. Individual differences and decision : a brief state-of-the-art

2.1. *Taxonomic and process oriented approaches*

It is quite difficult to present a complete state-of-the-art on personality, emotions, moods and cognition when considering the huge amount of research on these topics that has been done for decades (such attempts were however done in (Revelle, 1995) or in (Schwarz, 2000)). The recent works of Damasio on somatic markers theory (Damasio, 1994) motivated new and broader dynamics in the domain by making obvious that neuro-psychology and neuro-imaging would add new crucial perspectives on the subject (see also (Montague, 2006)). Therefore we shall try to limit our introductory survey to studies that are directly connected to decision-making; even with this precaution, we do not claim to be exhaustive in our presentation.

Most of the theories and models that were proposed up to now can be separated into two kinds of approaches:

– A first type of approach consists of mapping classes of user profiles to classes of decision behaviors. The user profiles are either in terms of personality or emotion but do not combine both. The second set of classes corresponds to possible decision-making styles. The construction of the mapping of individual profiles to decision styles is commonly based upon statistical inferences.

– A second type of approach focuses on the description of the principles that underly emotion and moods, and on the modeling of their impact - possibly in terms of competition or collaboration - on usual cognitive functions such as reasoning, memory or decision-making.

We examin briefly these two types of approach in the following paragraphs.

2.2. *Taxonomies*

2.2.1. *Personalities taxonomies*

As mentioned by (Revelle, 1995), defining taxonomies has been a recurrent tradition in personality theories since Plato. The most famous contemporary taxonomy was defined by Jung and instrumentalized by Meyers-Briggs (Carlyn, 1977). According to the Meyers-Briggs indicators, one personality may be described by four bipolar dimensions:

– sensing vs. intuition (SN),

– thinking vs. feeling (TF),

– extraversion vs. introversion (EI) and

– judgment vs. perception (JP)

Together these four dimensions describe the general attitude of the individual towards his/her environment, self, and logics. Almost as famous, the *Big Five* (McCrae *et al.*, 1996) and the *Even Bigger Three* (Eysenck, 1991) models (respectively, B5 and EB3) propose to classify individuals according to five or three dimensions amongst Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness. Let us only indicate here that these dimensions seem to broaden the focus on social features (Openness, Agreeableness), but that the main cognitive dimensions of personality description (Extraversion, Neuroticism) remain almost unchanged. For an in-depth discussion, one can refer to (Revelle, 1995).

Still according to (Revelle, 1995), an important point is that the number and the definition of attributes as well as the dimensions involved in the definition of classes have not been fixed yet, and still can be considered to be under discussion. In fact, many studies still focus on re-expressing these initial taxonomies and on finding an ever more adequate way of combining their related dimensions (Gray, 1994).

2.2.2. *Emotion / mood taxonomies*

Emotions and moods have also been expressed through many taxonomies (see (Schwarz, 2000) and (Frijda, 1994) for in-depth surveys). One of the most famous systems was proposed by Clore (Clore, 1994), who distinguished between emotions focused on self and focused on others, and between positive and negative (i.e. valence-oriented) emotions. Distinctions between emotions and moods may be related to the time duration of the phenomena, but also have to take into account the notion of intentionality (Frijda, 1994).

More recently, it was proposed (Lerner *et al.*, 2000) to go beyond the concept of positive or negative valence for emotions and define them in terms of cognitive appraisals (ATF, Appraisal Technique Framework (Lerner *et al.*, 2006)) instead. Stating that emotions that are gathered in the same category[1] could lead to very different behaviors, Lerner proposed that emotions should be expressed in terms of classes that reflect the cognitive evaluation of situation (perceived level of control of situation, perceived level of uncertainty).

2.2.3. *Decision styles*

There are fewchoixes for the different classes of decisional behaviors and styles. Two models seem to be emerging: the General Decision Making Style (GDMS) and the Decision Style Theory.

The GDMS framework (Scott *et al.*, 1995) proposes to define a decision style as one of five categories: *rationale, intuitive, dependent, avoidant* or *spontaneous*. Some of these categories are based upon the way the decision maker processes information. For instance *rationale* decision-makers are characterized by a "comprehensive search for information, inventory of alternatives and logical evaluation of alternatives", while *intuitive* ones "by attention to details and [...] relying on premonition and feeling" (Scott *et al.*, 1995). GDMS categories also take into account the relationship to others during decision-making (*dependent* decision makers search for advice) or the global attitude of the decision-maker towards the decision process itself (*avoidant* decision makers try to avoid making decisions, *spontaneous* ones want the decision process to converge rapidly).

In a different way, Rowe and Boulgarides propose a Decision Style Theory (DST) (Rowe *et al.*, 1992) that distinguishes four categories that are based on the ways we perceive the stimuli and the ways we choose to respond. According to DST, decision-makers may be *analytical, directive, conceptual* or *behavioral*. *Directive* decision-makers have a strong need for structure and are oriented towards tasks to be done more than towards people. *Analytical* decision-makers are also oriented towards tasks and technical aspects but can tolerate ambiguity. They evaluate the situation with abstract thinking. *Behavioral* style corresponds to focusing on people and social aspects. *Conceptual* decision-makers can tolerate ambiguity and are also focused on people.

2.2.4. *Mapping personality and emotion onto decision-styles: few robust results*

A next step usually consists of statistically analyzing and inferring the correlations between personality or emotions profiles and decision styles. It is unfortunately extremely difficult if not impossible to draw robust conclusions from most experi-

1. Fear and anger in a negative valence category for instance

mental and statistical results provided by past experiments (see (Thunholm, 2004) for instance). We see at least two reasons for this common hinderance.

First, definitions and classes provided by the decision-styles theories seem to be very vague. Most of the analyses only refer to an ambiguous qualitative dichotomy between *heuristic* and *analytical* decision-styles[2]. Although many experiments attempt to prove a correlation between negative emotions and moods and analytical decision-makers (while positive traits are tied to heuristic decisions (Schwarz, 2000)), they do not take the context or a concrete and measurable behaviour of decision-makers into account.

Another important concern is also related to the definition of a decision style: *decision styles* seem often to be indistinguishable from *cognitive styles*, which are themselves very close to psychological profiles used for personality definition. Thus, it seems to us, for instance, that the GDMS categories rationale and intuitive are very close to the thinking and intuition features of the initial taxonomy proposed by Jung[3]. According to (Leonard *et al.*, 2005), the four DST styles are re-expressions of the respective associations of dimensions proposed by the Myers-Briggs indicator. When aspects of input categories overlap output categories, it is therefore understandable that an inferential statistical approachto data nalysis leads to confusing and contradictory results. We reach some kind of tautological association where the features characterizing a decision style could already be part of the personality profile of the decision maker, and not result from it.

A last but major concern is that these approaches do not allow to think about cumulating or combining effects of personality and emotions on the decision maker's behaviours. Taxonomy-based studies happen to focus on personality, emotion or mood separately while it would be necessary to take into account the whole set of influencing factors to answer a question such as: "would the decision style of an extravert, but angry and stressed, decision-maker remain heuristic or shift to analytical?".

2.3. ***Cognitive processes, emotions and decision***

Another way of dealing with emotions and moods effects on decision consists in relying on cognitive processes *models*. This kind of approaches was especially used when taking into account the emotion in individual differences. Often referring to Damasio's work (Damasio, 1994), but as well to LeDoux (LeDoux, 1996) or Sloman (Sloman, 2002), a lot of studies have proposed to describe emotion as a

2. Lerner uses almost the same bipartition defined by *deep thought* against *shallow thought* (Lerner *et al.*, 2006).
3. Scott and Bruce admit themselves that these dimensions should be generally mixed in order to describe the real decision process.

specific process that could be complementary of the usual rational one. Forgas's Affect Infusion Model (Forgas, 2000) proposes for instance to distinguish different strategies for information processing (*direct access, motivated*, *heuristic*, *substantive*) and to identify when and where affect could impact them. According to this model, only the most *cognitive* processing strategies amongst the four may be influenced by emotion: *direct access* and *motivated* strategies barely involve conscious and complex retrieval in memory or processing of information, and therefore should not be influenced by affective markers related to the situations, while *heuristic* and *substantive* strategies could be. But the main concern with these approaches remains the same: the quantification of effects still relies on imprecise terms such as heuristic or analytical[4] and does not provide an easy way to analyze real decision cases and behaviors.

At last, different approaches directly address computational models, and propose to embed some parameters that could represent emotions and moods within the decision process into simulations or even in autonomous agents, like in EMA (Emotion and Adaptation) (Gratch *et al.*, 2006) for instance. In (Gmytrasiewicz *et al.*, 2000), one can also find a model of decision processes for agents based upon utility theory and its parameterization from emotions and moods. An interesting approach is also proposed by (Hudlicka *et al.*, 2004) with the MAMID methodology where individual differences, including personality, emotions and moods, directly impacts quantitative parameters of cognitive processes such as memory, attention, or reasoning.

2.4. *Finding new inspiration in psychology of decision*

Let us summarize the problems we relate to existing methods. First, the definition of decision styles are too imprecise to support statistical inference and do not allow any cumulative effect analysis. The definition of decision styles is also mostly qualitative and too close from psychological profiles to allow the matching with empirical observations. Second, when using an approach based upon processes, most models use a similar vague description of classes of decision behaviours. When quantitative, models of decision rely on utility models that are known to be difficult to validate experimentally and are far from cognitive psychology theories.

We however aim at a decision models that integrate all individual differences and all specific features of a decision maker. Further, only this condition can enable us to design personalized and customized decision support systems that takes into account at the same time personality and emotions of the decision maker. Some models issued from the psychology of decision allow to better fulfil these requirements. We present them in the following paragraphs.

4. Forgas mentions the "most cognitive" processes to be more influenced by emotion, for instance, but how could we define a *most* and a *least* cognitive process?

3. Cognitive models for individual decision-making

3.1. *Work hypotheses*

Cognitive approaches of decision making mostly rely on the assumption that decision-making may be defined precisely with information processing techniques. Decision-making involves two main kinds of tasks: choice and judgment. In a choice task, the alternatives are compared, the ones to the others, in a judgmental task a label has to be attached on each alternative. Another important feature is related to the level of expertise of the decision maker, as it may have some influence on the cognitive processes invoked during the decision processes. We make the hypothesis in this paper that our decision-makers are experts. A main characteristic of expertise in decision is the low amount of the information processed to perform a decision (while a novice uses an overcrowded amount of information, an expert one uses what is just enough but relevant (Shanteau, 1988)). Obviously such a phenomenon is balanced by the high quality and the appropriateness of the used information. We will focus in the following on expert decision makers, especially because the notion of decision-style should be more easily related to some stable, experienced and measurable decisional behaviors. We also assume a multi-attribute framework to describe alternatives

3.2. *Cognitive models and dominance structures*

Three main kinds of cognitive models for decision-making have been proposed. A first one asserts that decision-making may be described as the chaining and use of elementary strategies that are applied on alternatives and/or attributes (Montgomery *et al.*, 1976). A second one describes decision-making as problem solving (Huber, 1986). A third one proposes to model decision-making as a search for dominance structures (SDS) (Montgomery, 1983). We shall rely in the following on a computational version of SDS, the Moving basis Heuristics (MBH), which was proposed by J.P.Barthélemy and E.Mullet in 1986 (Barthélemy *et al.*, 1986). We chose to rely on the MBH as it allows exhibiting some references combinations of attributes and values (called aspects) quite easily and from simple observations of the decision-maker.

3.3. *Individual differences and dominance structures*

Let us suppose that a judgmental decision task attached to an expert consists in selecting or rejecting alternatives that can be described along 10 attributes a_1 ... a_{10}. For sake of simplicity, let us imagine as well that each attribute has 5 possible ordered values 1 to 5. Writing a_i^j the attribution of value j to attribute i, an alternative - dedicated to be either selected or rejected - can then be described by a vector of aspects such as $A = [a_1^3, a_2^4, ..., a_{10}^2]$ for instance, meaning simply that A has a value 3 for its attribute a_1, 4 for its attribute a_2, etc.

Following the dominance search and MBH approaches, we assume that the expert's decisions will always rely on a limited set of information. More precisely, this means that when analyzing one alternative to be evaluated, the expert will only consider some specific and limited subsets of aspects that could win the decision. Thus, one could imagine that a combination such as [a_3^3, a_5^2] (attribute a_3 with value 3 and attribute a_5 with value 2) is used as a reference and allows by itself to decide for a positive judgment for all alternatives that will be *at least* at these levels for attributes a_3 and a_5, whatever could be the values of the other attributes. $SD1 = a_3^3, a_5^2$ is then one reference for the expert decision-maker, namely one of his/her dominance structures. $[a_2^4, a_4^3]$ or $[a_3^1, a_7^2, a_9^4]$ could be possible other dominance structures, but $[a_3^2, a_5^1]$ could not, as being dominated by SD1 ($a_3^2 \leq a_3^3$ and $a_5^1 \leq a_5^2$). In simple words, expert decision will then consist in browsing the set of dominance structures and in checking that at least one of them is satisfied by the alternative.

It was demonstrated that it was possible to extract the set of dominance structures related to an expert decision making through the the analysis of his/her decisions under interactive questioning (Barthélemy *et al.*, 1995). Of course, the complete exploration of the combinatorial space is generally not possible: the size of the problem may be too large, or even when limited, too large for an expert decision maker to accept to spend so much time to browse explicitly the whole space. Efficient solutions may be however proposed with light supplementary hypotheses (Lenca, 1997).

In other words, it is possible to extract from expert decision makers' behaviours some sets of data structures that represent their respective decision anchors and references, and that describe their respective and personalized decision process. We consider these sets are the key for expressing individual differences on decision making, as we explain in the following paragraphs.

3.3.1. *Defining decision-style from (observed) decision-maker's dominance structures*

Let us now suppose, as an example, that

$$SD1 = a_3^2, a_2^5$$

,

$$SD2 = a_4^1$$

,

$$SD3 = a_4^5, a_{10}^2$$

could be attached to the description of the behaviors of one expert decision maker DM1. Hence, (s)he would decide an alternative to be accepted if this alternative presents either a combination of attributes a_2 and a_5 respectively greater than 3 and 2, or an attribute a_1 greater than 4, or a combination of attributes a_4 and a_6 respectively greater than 6 and 10. Suppose now that another expert DM2 on the same decision task could be described, after the same kind of analysis, by the structures

$$SD1 = a_1^4, a_2^4, a_3^3, a_7^2 \quad , \quad SD2 = a_2^3, a_5^2$$

,

$$SD3 = a_1^4, a_4^3, a_9^3$$

,

$$SD4 = a_2^3, a_7^4, a_8^3, a_9^2, a_10^2 \quad , \quad SD5 = a_5^4, a_3^3, a_7^5$$

,

$$SD6 = a_1^3, a_4^2, a_9^2$$

Obviously, we would have now some support to say that the second decision maker bases his/her decisions on more in-depth analysis of alternatives attributes, and has more references to take to justify his/her decisions: *in-depth analysis* can actually be attached to the average number of attributes per dominance structure and to the amount of dominance structures attached to the decision task.

We can further define two levels of empirically-based decision-styles for the respective expert decision makers:

– at a *categorical* level, DM2 can be said to be more analytical than DM1, on the basis of the measured number of dominance structures used in the decision task and on the average number of aspects composing the dominance structures. DM1 could be said, at the opposite, to be heuristic.

– at the *individual* level, each DM can be uniquely defined by the contents of his/her dominance structures. Decision-styles could be then compared on a set-comparison (two expert decision makers can happen to share few or a major part of their rules).

We have represented on figure 1 the dominance structures that were extracted from the observation of expert decision makers in professional orientation (Barthélemy *et al.*, 1986). Each of the sixteen experts had to analyse a same set of about one hundred student applications, and to decide whether the student was to be accepted or not. Students applications described their grades (1 to 8) in different disciplines (A to I). According to the previous notation, dominance structures are encoded by the name of the attributes - A to I - and the related value - 1 to 8, the + operator being equivalent to a logical *or*. Thus the expert 1 exhibited choices that could be explained by a grade in discipline A greater than 5, or a grade in C greater than 6, the combination of two grades of 4 for disciplines A and C or a same kind of combination for disciplines C and I.

From this analysis, the expert number 7 and the expert 11 exhibit different decision style: the first one apply a one-attribute unique rule as the second one uses 9 rules, most of them being double-attribute . Both of them remain heuristic in their strategies, but we can now distinguish between different ways of being heuristic. Experts 4 and 7 could this way share a same categorical style (one decision rule only, with one attribute involved only), and their respective individual style would be defined by their respective threshold (in this particular case, both expert use the same attribute in their unique decision rule).

Expert	*rules*
1	A5+C6+A4C4+C4I4
2	A5F5+A4I5+B3F6+C2G6+D3E6
3	A6+A4B6+A4G6+B6C4+B3D6+B5I6+C4I5+D3E6+D3J6+E3H6
4	A5
5	I6+A5B3+A2C6+A3D5+B3D6+D3J6
6	C6+A4E5+A5F5+A5G4+B4F5+B5I6
7	A4
8	F5+A3C6+A3D4+B5D4+B3H6
9	A6+A5B5+B6C3+B2D6+C4D5+C5D4+C3G6+D3E5
10	A4B6+A5B3+A3D6+A4F6+A4G6+C6D5+D3J6+E6F5
11	A5+A3C6+A3D6+A4E4+A4F5+A4H3+B5I6+C3E6+C3G6+C4I5
12	A3C6+A3D6+A2F6+B4F5+B3H6+C3G6+C4I6+E4J6
13	A5+A3D6+A4G6+B6C4+B3E6+B3F6+B4G6+C5D4+D6F3
14	B6+E6+A5B3+A3C6+A4F6+B3D6+C4I5+D3J6
15	A5+G6+H6+A4I5+C3I6+F6G4
16	A5+I6+A3C6+A3E6+A4F5+A3I6+B2C6

Figure 1. *Representation of dominance structures underlying decision strategies of 16 experts*

3.4. *Extending the decision-style to emotions and moods*

It is not too difficult now to think about taking into account of emotions in a similar way: still developing our purposefully simplified example, one could analyze the effect of emotion on decision-making through a differential effect measured on the sets of dominance structures. Thus stating that our expert decision-maker DM1 shifts for instance from

$$SD1 = a_2^3 a_5^2, SD2 = a_1^4, SD3 = a_6^4 a_{10}^2$$

to

$$SDe1 = a_1^3, a_5^2, a_6^3, SDe2 = a_1^3, a_4^2, a_9^2$$

$$SDe3 = a_1^4, a_6^4, a_{10}^2, SDe4 = a_2^3, a_7^4, a_8^3, a_9^2, a_{10}^2$$

under the effect of sadness could allow an objective quantification of "going towards analytical". The same means can be used to analyze the effect of aggregated effects, such as shyness added to anger added to stress for instance, in cumulating the differential evolution of dominance structures to achieve at least a qualitative understanding of the process.

Effects of emotions and moods on decision-making thus can be defined and possibly measured from the variation of the apparent dominance structures that can be extracted from the expert decision behavior. If we call N1 the number of structures and N2 their average number of aspects, we may add to the definition of one decision-maker's individual style his/her typical variations of strategies when facing given emotional states. In the previous toy-example, variations of $N1 = +1$ and $N2 = +1.8$ can be attached to the decision-maker DM1. At a categorical level, one could expect some of the expert decision-makers to globally show limited variations of decision structures under effect of emotion or mood.

4. Defining decision-styles from empirical observations and observables

Describing the stable decision strategies of a decision-maker thanks to his dominance structures allows in a way to think of a new point of view on individual differences in decision-making. What matters really in order to understand and further to assist a decision maker is no more related in priority to his or her personality category or emotional state. What matters is the form and content of the decision maker's set of dominance struictures, their content and their possible variation depending on mood and emotion. This approach especially allows us to evaluate properly the additive effects to be expected from the personality, the emotion and the mood of the decision-maker.

In other terms, the method should better consist in first identifying the main forms of sets of decision structures and, only after this first step, try to correlate them with combinations of emotions, moods or personnality. We do not claim that this would mean to redefine the categories of moods or emotions *from* the decision styles, but at least, this should give a more practical and concrete way of managing individual differences in decision making.

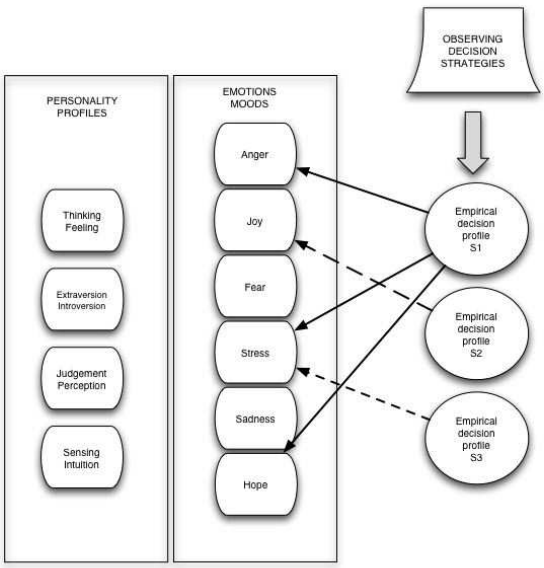


Figure 2. *Going from classes of observed decision behaviors backward to emotional states and personalities*

5. Individual and personalized decision support system

As we propose to rely on existing and observable human decisional expertise - or minimally recurrent and stable behaviors, we have to forget about classical decision support systems and change our mind about the role of an assisting system. The assisting system is not considered as providing the decision maker with solutions that refer to an absolute reference or optimum any longer, but on the contrary as proposing a kind of structured mirror image of decision maker's actions and strategies, aiming at the extraction and synthetic formalization of one unique decision-style.

What can be the operational interest of this approach, if the decision maker is an expert and if the data that are initially used for extracting the most pro-eminent strategies come from him / her? The answer is at least double:

– first, the expert decision maker may not be able to express his / her own expert strategies. Proposing to mirror these decisional strategies allow reaching a better level of meta-cognition for the expert DM that may be usually drowned into daily action.

– second, this explicit expression of decisional strategies may be used as reminders and controls for the expert himself / herself, or even as guides for novices

This approach was developed and validated on different cases of industrial process control that involved expert decision makers. Strategies of control were extracted from the daily observation of behaviors, then synthesized and displayed towards the experts themselves as a mirror of their know-how. This on-line non-intrusive assistant allowed to elicit expert strategies, and to capitalize and work further on refining the process control (Coppin *et al.*, 2007).

But it should be possible now to go further in the management of decision support. Affective computing as defined by (Picard, 1997), that is to say taking into account emotions and moods in the management of man machine interaction, has become a major challenge and could possibly be the key for a more efficient man-machine cooperation. This is even more crucial when dealing with decision-making. Overwhelming flows of information addressed to a decision maker that is known to decide from few robust and simple dominance structures could rapidly lead to bad performances and failure. Hiding information or shifting part of them to the machine when the decision-maker is more analytical and needs more exhaustive browsing his/her numerous dominance structures could not be efficient either. So defining personal and individual decision-styles from personality and emotion, and identifying them from current behaviors, could allow to adapt the communication, the information display and even the task sharing between the user and the machine.

6. Conclusion

In order to personalize the decision assistance, it is necessary to better focus on the effects of emotions and moods on the behaviors of decision-makers, and especially the expert ones. We have presented in this paper an alternative way to classify these effects. Starting from observable data and interpreting them from a cognitive psychology point of view, we can extract formal and comparable descriptions of decision strategies, i.e. the dominance structures, and then map more safely the effective decisional behaviors with measured or inferred personality profiles of states of emotion. When doing this, we expect to make decision styles explicitly based upon observable behaviours instead of defining them *a priori*, and to validate these styles definition with the feedback of the decision makers themselves. These perspectives should be studied in coming laboratory experiments.

Acknowledgements

The author especially thanks MIT/HAL Pr. Missy Cummings, Mark Ashdown, Yves Boussemart and Vanessa Esch for their fruitful and relevant comments and suggestions on this paper.

7. References

Barthélemy J.-P., Coppin G., Guillet F., “ Smelting process control: from experimental design to acquisition of expertise”, *International Conference on Industrial Engineering and Production Management*, vol. 2, p. 2-11, 1995.

Barthélemy J.-P., Mullet E., “ Choice basis: A model for multiattribute preference”, *British Journal of Mathematical and Statistical Psychology*, vol. 39, p. 106-124, 1986.

Carlyn M., “ An assessment of the Myers-Briggs Type indicator”, *Journal of Personal Assessment*, vol. 41, n˚ 5, p. 461-473, 1977.

Clore G., *The nature of emotion*, Oxford University Press, chapter Emotions require cognition, p. 181-191, 1994.

Coppin G., Cadier F., Lenca P., “ Some considerations on cognitive modeling for collective decision support”, *HICSS 2007*, Hawaii, 2007.

Damasio A., *Descartes’ error: emotion, reason and the human brain*, Putnam, New York, 1994.

Eysenck H., “ Dimensions of personality: 5 or 3? criteria for a taxonomic paradigm”, *Personality and individual differences*, vol. 12, p. 773-790, 1991.

Forgas J., *Feeling and thinking: the role of affect in social cognition*, Cambridge University Press, chapter Affect and information processing strategies: an interactive relationship, p. 253-280, 2000.

Frijda N., *The nature of emotion*, Oxford University Press, chapter Varieties of affect, emotions and episodes, moods, and sentiments, p. 59-67, 1994.

Gmytrasiewicz P., Lisetti C., “ Using decision theory to formalize emotions in multi-agent systems”, *ICMAS*, 2000.

Gratch J., Marsella S., “ Evaluating a computational model of emotion”, *Journal of Autonomous Agents and Multiagent Systems*, vol. 11, n˚ 1, p. 23-43, 2006.

Gray J., *The nature of emotion*, Oxford University Press, chapter Personality dimensions and emotion systems, p. 329-331, 1994.

Huber O., *New directions in research on decision making*, North Holland, chapter Decision making as a problem solving process, 1986.

Hudlicka E., Zacharias G., *Proceedings of the 2004 Winter Simulation Conference*, chapter Approaches for modeling individuals within organizational simulations, 2004.

LeDoux J., *The emotional brain*, Simon and Schuster, New York, 1996.

Lenca P., Acquisition automatique et analyse de processus de décision. Application au domaine bancaire (*Automatic acquisition and analysis of decision processes. Application to bank domain*), PhD thesis, École Nationale Supérieure des Télécommunications de Bretagne, Université de Rennes I, 1997.

Leonard N., Scholl R., Kowalski K., " Information processing style and decision making", *Science*, vol. 310, p. 1680-1683, 2005.

Lerner J., Keltner D., " Beyond valence: toward a model of emotion-specific influences on judgment and choice", *Cognition and emotion*, vol. 14, n° 4, p. 473-493, 2000.

Lerner J., Tiedens L., " Portrait of the angry decision maker: how appraisal tendencies shape anger's influence on cognition", *Journal of behavioral decision making*, vol. 15, p. 115-137, 2006.

McCrae R., Costa P., *The five-factor model of personality: Theoretical perspectives*, Guilford, New York, NY, chapter Toward a new generation of personality theories: Theoretical contexts for the five-factor model, p. 51-87, 1996.

Montague R., *Why choose this book?*, Dutton, 2006.

Montgomery H., *Analysing and Aiding Decision Processes*, North-Holland, chapter Decision rules and the search for a dominance structure: toward a process model of decision-making, p. 343-369, 1983.

Montgomery H., Svenson O., " On decision rules and information processing strategies for choices among multi-attribute alternatives", *Scandinavian Journal of Psychology*, vol. 17, p. 283-291, 1976.

Picard R., *Affective computing*, Cambridge Press, 1997.

Revelle W., " Personnality processes", *Annual Review of Psychology*, vol. 46, p. 295-328, 1995.

Rowe A., Boulgarides J., *Managerial decision making*, MacMillan publishing company, New York, 1992.

Schwarz N., " Emotion, cognition and decision making", *Cognition and emotion*, vol. 14, n° 4, p. 433-440, 2000.

Scott S., Bruce R., " Decision making style: The development and assessment of a new measure", *Educational and Psychological Measurement*, vol. 55, n° 5, p. 818-831, 1995.

Shanteau J., " Psychological characteristics of expert decision makers", *Acta Psychologica*, vol. 68, n° 1-3, p. 203-215, 1988.

Sloman S., *Heuristics and biases: The psychology of intuitive judgment*, Cambridge University Press, New York, chapter Two systems of reasoning, p. 379-396, 2002.

Thunholm P., " Decision-making style: habit, style or both", *Personality and individual differences*, vol. 36, n° 4, p. 931-944, 2004.

The influence of stress and mood on time-constrained decision making in crisis situations

To what extent can we predict human performance?

E.W.M. Rondeel — M.H. Kempen — N.J.E. Wijngaards — C.H.M. Nieuwenhuis

D-CIS Lab / Thales Research & Technology Netherlands
P.O. box 90, NL-2600 AB, Delft, The Netherlands

Eefje.rondeel@icis.decis.nl

ABSTRACT. *In this position paper a number of hypotheses are posited concerning the effect of measurable human factors, such as subjective stress, arousal and mood, on the performance of human decision making; taking into account the amount of risk involved in the decision. The proposed domain of application is crisis management: a situation in which time-limits, uncertainty and possibly dire consequences provide an ideal context to assess the validity of our hypotheses. Experimentation involves both people in management functions and non-management functions. The final objective is to provide the basis for a demonstrator which can measure mood, arousal and subjective stress on the job, provide runtime feedback and ergo positively influence human decision making processes.*

KEY WORDS: *decision making, stress, mood, risk, crisis management.*

1. Introduction

The focus of this position paper is on human decision making in critical circumstances, in which lives can be at stake and in which a wrong choice can have detrimental consequences. Decisions to be made in such risky situations are critical for the prevention of death, injury, loss and damage. During crisis management large amounts of information have to be processed; information that is frequently incomplete, faulty, uncertain, conflicting and untimely, and which often has to be processed under severe time constraints. In the context of crisis situations, decision making is crucial, especially when the impact of the decision is tremendous. It is well known that human factors play an important role in decision making in general and this is expected to be even more so during crisis situations. The suggested research is to lead to the improvement of the quality of the human decision making process in a crisis situation.

We claim that the human factors subjective stress, arousal and mood have an influence on human decision making, and that the amount of risk concerned with the decision moderates this influence. Past research has not yet addressed the specific influences of subjective stress, arousal and mood on decision making processes in particular decision tasks or contexts. Additionally, two important characteristics of stress, namely subjective stress and arousal, have not yet been disentangled explicitly in previous experiments. Subjective stress is defined here as a person’s inability to cope with a certain situation. Arousal is defined as a physiological state that is characterized by an increased heart rate and blood pressure. Mood is defined here as a diffuse and affective state, ranging from positive to negative. Finally, risk is defined as a concept that signals a potential negative impact on something valuable. It should be noted that there is no general consensus on an exact definition of the above mentioned concepts. However, in order to clearly define the scope of our claim and our hypotheses, it is important to outline what the authors aim at by the use of the concepts that are to be investigated.

The results from the proposed studies are to be incorporated in a software agent which is able to monitor a human decision maker and predict the quality of the human decision making process. For this purpose, specific heuristics are needed on the basis of the valuation of mood, subjective stress and arousal, combined with a determination of the level of risk. Valuation of these human factors is assumed to be a set of discrete values. Devising a complete set of experimentally grounded heuristics is too expensive in terms of time and effort. The joint valuation of mood, subjective stress and arousal in the context of risk is therefore disentangled such that specific hypotheses can be tested, leading to the development of heuristics. These heuristics provide the basis for the software agent.

The posited hypotheses are presented in Table 1. Most hypotheses are derived from existing literature and previous research, as indicated in the table. The large columns represent the three human factors being studied both for negative/low and positive/high states. The rows represent the riskiness of the decision, i.e., is the decision to be made in the context of a potentially high risk outcome (high impact) or in the context of a low risk outcome (low impact). Each cell indicates the expected quality of the decision making process. The expected performance (quality) will be measured by investigating the decision making *process* and not the outcome of the decision, since the outcome of a decision is beyond experimental control. The quality of the decision making process is measured by investigating how much information is used in order to make a decision, whether people fully explore all possible outcomes and if decisions are not taken too hastily.

	Positive mood	Negative mood	Low Subjective stress	High subjective stress	Low arousal	High arousal
Low risk	+ + [1]	- [1]	+ / - [1]	- [4]	+ / - [1]	+ + *
High risk	- [2]	- - [3]	- [1]	- - [5]	- [1]	- *

Table 1. *Expected performance on decision making tasks for the different conditions. Performance is in comparison to a control group; performance can be: - -, -, +/-, +, or ++. Expected performance in the low subjective stress and low arousal group are based on research on the influence of risk on decision making in general.*
*[1] Based on, for example, Isen, 2000. [2] Based on, for example, Williams, 2004. [3] Based on, for example, Raghunathan & Pham, 1999. [4] Based on, for example, Kontagiannis & Kossiavelou, 1999. [5] Based on, for example, Kowalski-Trakofler, Vaught, & Sharf, 2003. * Indicates a well-educated guess by the authors.*

Although some hypotheses have already been proven in previous research, all hypotheses need to be investigated as human subjects may exhibit all valuations of all factors during experimentation. The hypotheses will eventually lead to a theory about the influence of subjective stress, arousal and mood with risk as a moderating factor, on decision making processes in critical circumstances. In the remainder of this paper, the task domain is outlined, relevant research topics are discussed in more detail and our own research approach to substantiating our claim is described. Finally, the relevance and implications of our claim are discussed.

2. Task domain

We claim that investigating the influence of mood, subjective stress and arousal on decision making is necessary in a *naturalistic setting*, implying that the human factors to be investigated have to be embedded in naturalistic decision making research. Naturalistic decision making entails the art of decision making with limited time, knowledge and resources (see for example Todd & Gigerenzer, 2001): it is a framework that directs attention towards experienced decision makers working in the field, thus strongly favoring the use of models based at least partly on heuristics. Naturalistic decision making can be described as an attempt to understand how people make decisions in real world contexts that are meaningful to them (see for example Whyte, 2001). Naturalistic decision making deals with real world tasks rather than classical laboratory experiments, such as gambling tasks, making it a very fruitful paradigm to substantiate our posited hypotheses.

Another characteristic is the *context* in which the decision making performance is to be investigated. Subjective stress, arousal and mood are investigated in a crisis situation in which the public safety is disturbed; people, the environment or important material are at risk; the incident that causes the crisis situation is sufficiently wide spread and the incident requires a fully operating management team and operational team[1]. Since crisis situations are characterized by a highly changing environment and new events, the decision making process concerns decisions to be taken within several minutes.

The focus of this position paper is primarily on the decision making performance of people in management teams, such as mayors, and not necessarily on the decision making quality of people in operational teams, such as fire-fighters. The reason for this is that people in these management functions do not routinely operate in a crisis situation, that is, they do not encounter decision making in a crisis management context on a day-to-day basis. Nevertheless, people in management functions are required to make important decisions in crisis situations. Investigating the influence of subjective stress, arousal and mood on decision making with a focus on people in management functions has a high ecological validity. For example, during a flood, a mayor of a city has to decide whether or not to evacuate citizens. When lives are at stake, and not to forget the mayor's respectability, stress levels can be very high, which might lead to a decreased decision making performance.

Equally important is that decision making is investigated in a naturalistic setting as a *dynamic and situated process*, in which one decision may influence the next one and in which decisions in the past cannot be undone. The decision maker's setting in time is depicted in Figure 1. At each time step t the decision maker is observing the world and acting upon it. Not only the events in the world in which the person is situated influence the decision making process, but also human factors, such as mood, subjective stress and arousal. The person's interpretation of the impact of possible future outcomes is represented by the factor risk. The granularity of the time steps t is chosen such that at each time step t at most one decision is made. Actions taken at previous time steps influence the situation at time step $t+1$. For example, the outcome of a decision might be interpreted as a 'good' outcome, which may elevate the decision maker's mood and reduce stress. This 'mind state' will in turn influence the process to make the next decision.

[1] The terminology of operating management team and operational team is common in descriptions of disaster and crisis management procedures and organizations

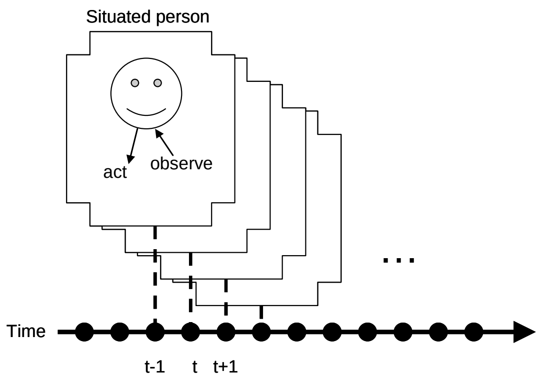


Figure 1. *Naturalistic decision making as a dynamic and situated process. A decision at a certain point in time may influence decisions at future points in time.*

3. Research background

Our hypotheses, as presented in Table 1, are based on a solid foundation of research on decision making and human factors. In this section, relevant research is discussed regarding decision making in general, subjective stress, arousal and mood, as well as the role of risk for each of these factors in decision making.

The expected utility theory (Von Neumann & Morgenstern, 1947) is a well known theory for formalizing decision making, stating that people make decisions by weighing the severity and the likelihood of the possible outcomes of different alternatives. However, decision making has been found to be influenced by many factors, ranging from time constraints to a person's mood state. Also the complexity of the decision plays a role. Research shows, for example, that the more complex the decision, the more people will use heuristics (e.g., Tversky & Kahneman, 2004).

3.1. Psychological and physiological characteristics of stress

Hammond (2000) stated: 'we know almost nothing about the extent to which decisions are affected by stressful circumstances, much less the manner in which the decisions are influenced by high-stress environments' (Hammond, 2000, cited from Kowalski-Trakofler et al., 2003, p. 280). However, we already do know that stress is an important factor in decision making and research on stress and decision making is of high relevance in a large number of situations. Especially the effect of stressful conditions on human judgment in emergency situations is of importance, since judgments made in these situations can have detrimental consequences. Stress occurs when environmental demands tax or exceed the adaptive capacity of an organism, resulting in psychological and biological changes (Cohen, Kessler, &

Gordon, 1995). Although these psychological and biological changes are often linked, they can also occur invariantly (Huey & Wickens, 1993). In the proposed research, these psychological and biological changes are investigated separately. *Subjective stress* will be used to refer to psychological changes, whereas biological changes will be referred to by *arousal*. The inability to cope with stressful situations depends on an individual's perception or interpretation of that situation; whether an individual experiences *subjective stress* thus depends on the perceived experience of the stressful situation (Gillis, 1993). As mentioned above, stress can be accompanied by biological changes, e.g. arousal. However, an increased arousal level does not necessarily mean that a person cannot cope with a certain situation. It is therefore not surprising, that in research on stress, results have been found that support both the notion that stress leads to improved performance, but also to performance degradation (see for example Poulton, 1976). Disentangling these two characteristics is therefore of high importance, because it is likely that they have different influences on the decision making process.

3.2. Subjective stress and decision making

Several stressors can lead to psychological and biological changes, for example time pressure, potential threat, workload, noise or even the weather. The main findings from previous research about the influence of stress (both subjective stress and arousal) on decision making are that stress leads to the narrowing of attention, a lack of concentration, over-reliance on heuristics and biased decisions. It has also been found that stress restricts cue sampling, decreases vigilance, reduces the capacity of working memory, causes premature closure in evaluating alternative options, and results in task shedding (Kontogiannis & Kossiavelou, 1999). Stress, (as caused by limited time) also leads people to not rely on learned procedures to perform a decision. Although this can be partly compensated for by the use of checklists, stress might have far more complex influences on decision making that cannot simply be compensated for by the use of checklists. Furthermore, people under stress seem to focus on more important information but also on more negative aspects in information (see for example Kowalski-Trakofler et al., 2003).

The role of risk. Risk is a major factor as it encodes the subjective relevance of the decision that has to be made in naturalistic decision making. In everyday usage, risk is often used synonymously with the probability of a known loss. The overall results that have been found in research on risk and decision making is that when risk is high, people become risk avoidant. Thus, if a situation concerns risk (for example lives are at stake), and time is limited, people tend to become more cautious and adopt risk-avoiding behavior, primarily focused on avoiding short term losses. Subjects in such situations also tend to adopt a simpler mode of information processing in which alternatives are not explored fully to determine the decision's consequences (cited from Kowalski-Trakofler et al., 2003). People under stress have also been found to offer solutions before all alternatives have been considered and they scan these alternatives in a non systematic approach.

3.3. Arousal and decision making

A large amount of research on stress and decision making can also be applied to arousal and decision making. Indeed, it seems that if the level of arousal increases, attention is restricted and cues that might be relevant to a certain task will be neglected (e.g., Huey et al., 1993). A high level of arousal also has an effect on attention selection, i.e., high arousal levels lead a person to focus on a restricted set of environmental or internal sources of information (Huey et al., 1993). However, it is also sometimes found that a higher level of arousal leads to better performance (e.g., Hockey, 1970), especially on vigilance tasks (see for example Kennedy & Coulter, 1975). The task at hand seems to be correlated to whether decision making performance increases or decreases. Important to note is that when subjective stress is high, the level of arousal is also likely to be high, but not necessarily the reverse. Whether an increased level of arousal has a positive influence on decision making performance is therefore only an educated guess at this moment.

The role of risk. As mentioned in section 3.2., when risk is high, people are more likely to avoid risky decisions. However, a high level of arousal is associated with risk seeking behavior (e.g., Mano, 1994), suggesting that people with a high level of arousal engage in a worse decision making strategy when the risk concerned with that decision is high. The particular influence of arousal on the decision strategy also seems to depend on the kind of decision and whether the arousal level is accompanied by positive feelings (e.g., joy) or negative feelings (e.g., anxiety). For example, it seems that when arousal is high and the feeling is negative, risk taking can be destructive. A high level of arousal combined with negative feelings seems to lead to a preference for a high risk, high pay-off choice over a low risk, low pay-off choice (e.g., Leith & Baumeister, 1996). However, Raghunathan and Pham (1999) found that conditions in which arousal is high (e.g., anxiety), led to risk aversion. Thus, findings on the influence of arousal on risk are not yet straightforward.

3.4. Mood and decision making

The last human factor to be investigated is mood. Mood is a diffuse and affective state, ranging from positive to negative. In normal functioning, the human adaptive system is believed to accommodate changes in moods by influences of external events. Schwarz (2002) states that mood, being a diffuse affect, serves as a source of information about the state of the environment. A negative mood is reflected on to the environment and lets the person in question believe that the situation or environment is problematic. A positive mood lets someone believe that the situation is benign. Therefore, mood determines the cognitive processing style that is being used in a certain situation. The influence of mood on processing style is widely tested and indeed it is found that mood activates different processing styles. People in a negative mood are more precise, judge more carefully and deliberately and perform better at analytical reasoning tasks. People in a positive mood, on the other hand, engage in an intuitive, heuristic, top-down processing style (Schwarz, 2002).

Rondeel (2007) for example found that a positive mood let to better performance on a change blindness task when peripheral changes had to be detected, whereas people in a positive and negative mood did not differ in terms of physical aspects of visual perception.

Even when people lack the presence of internal knowledge structures, such as heuristics, a person's mood guides what type of information is noticed: people in a negative mood focus more on local information and people in a positive mood focus more on global information. This difference in focus in different mood states, based on the Easterbrook hypothesis, is called the levels-of-focus hypothesis (Clore, Wyer, Dienes, Gasper, Gohm, & Isbell, 2001). Gasper and Clore (2002) investigated this hypothesis and indeed found that a negative mood seems to inhibit a global focus. Furthermore, a positive mood leads to increased flexibility in thinking, whereas a negative mood leads to more analytical and critical thinking. In addition, there is less distortion in decision making when people are in a positive mood compared to a negative mood. In fact, a positive mood has even been found to promote efficiency and thoroughness in decision making (Isen, 2000; Isen & Means, 1983). A possible explanation for this could be that people in a positive mood integrate information at an earlier stage in the decision making process (Estrada, Isen, & Young, 1997).

The role of risk. Another theory about mood and decision making states that people in a positive mood are looking for the maintenance of their mood state, and will therefore make a decision that maintains their current mood. For example, Thompson (2005) found that individuals in a happy mood were more willing to maintain their mood status by being risk-aversive. Other research, conducted by Isen and her colleagues (e.g., Isen & Geva, 1987; Isen & Patrick, 1983) supports this view but also shows that a positive mood increases riskiness in a low risk task, in which success is more likely. It thus seems as if the outcome of the decision plays a role. For example, people in a positive mood underestimate the chance of losing something, but they do rate the potential loss as bigger. Other research again suggests that people in a negative mood are more willing to take risk, when they believe this choice will lead to a positive outcome (Raghunathan et al., 1999). Williams (2004) investigated the impact of mood on managerial perceptions and found that a positive mood led to an increased perception of willingness to undertake risky business propositions. Thus, research of mood on decision making and risk is still unequivocal, but it does imply an effect of mood in real life situations.

4. Research approach

To test the posited hypotheses and to meet our practical objective, a pragmatic scientific approach is adopted, including empirical data gathering, the development of a theory about the influence of mood, subjective stress and arousal on the decision making process, the development of a measurement tool of mood, subjective stress and arousal, experimental data gathering and finally formulating heuristics for a software agent capable of monitoring and influencing the quality of the human decision making proces.

Subjects. The target group for subjects is two-fold. First, the results are to be generalized to people in management functions in the context of crisis situations. Therefore, research focuses on people in managerial functions, such as mayors. Second, pilot studies are foreseen to test several aspects of the to-be-conducted experiments. In addition, further experiments are necessary to develop reliable and effective measurements for subjective stress, arousal and mood. For these latter pilot studies and experiments, large groups of subjects are required, which can also exist of people with non-managerial functions. Control groups are used to serve as a comparison group to evaluate the decision making process (see also section 1) and will consist of subjects with a 'neutral' mind state, i.e., a neutral mood, no induced subjective stress and no induced arousal.

Empirical data gathering. The first step will be to gather empirical data by interviewing people in a managerial function, such as mayors, about the role of subjective stress, arousal and mood in decision making. This is partly done by interviewing people during or after a crisis exercise and partly by interviewing mayors who have encountered a crisis situation in their career (for examples of how mayors experience crisis management, see Jong & Johannink, 2007). Gathering empirical data in crisis management exercises is a perfect starting point for developing a theory about the influence of subjective stress, arousal and mood on decision making. A questionnaire will be developed which provides a good account of the person's own experiences, expectations and evaluations of the crisis situation, with the primary focus on the role of subjective stress, arousal and mood in the decision making process in the encountered crisis situation.

Subjective stress, arousal and mood measurement. In parallel with the gathering of empirical data, experiments will be conducted to develop a reliable measure for subjective stress, arousal and mood. The measure should not include questionnaires about personality traits but should use observable parameters. This measure should have the ability to disentangle the characteristics of stress, namely subjective stress and arousal. In addition, a reliable measure for mood states will be explored and developed. Measurement devices for psycho physiological responses will be used and correlated with subjective measurement scales, such as the Profile of Mood States (POMS; McNair, Lorr, & Droppleman, 1992). Another reliable measure for subjective stress is the stress hormone cortisol level (e.g., Roelofs, Elzinga, & Rotteveel, 2005). Additionally, EMG (Electromyography) is a measurement method that is investigated as measure for stress or mood. Measurement methods for arousal include skin conductance levels and heart rate. We focus on measurement devices which are very low in intrusiveness and burden the participant as little as possible. Pilot studies should indicate how the abovementioned measures correlate and how mood, subjective stress and arousal are related in a typical crisis situation.

Experimental data gathering. To investigate the influence of subjective stress, arousal and mood on decision making, several experiments will be conducted with a large group of subjects. The professional skill of the subjects is an important factor in these experiments, warranting specific selection of subject groups (non-managers, managers, mayors, etc.). These experiments serve to test the hypotheses as well as the measures for subjective stress, arousal and mood. The influence of the factor risk will also be determined in this series of experiments. A task or scenario will be chosen in which mood, subjective stress or arousal is manipulated. Performance will be measured as well as subjective stress, mood and arousal.

Experimentation setting. One likely setting for the experiments is that subjects are involved in a simulation game in which they are the mayor of a small city. In the simulated scenario, a disaster occurs and the subjects need to make some important decisions. In addition, their subjective stress level, arousal level or mood is being manipulated. The scenario should be sufficiently appealing to the subjects so that they can empathize with the developing crisis and its effect on 'their' city. In order to obtain reliable results, the situation should have a high subjective relevance for the decision makers. In addition, the simulation game should take sufficient time for subjects to allow them to undergo different levels of subjective stress, arousal and mood. The scenario employed in the experiments needs to be well-composed and specifically tailored for use in experimentation. The amount of choices (plot-lines, decision points, etc.) should be limited and predefined such that the scenario is repeatable across subjects. The number of choices should be sufficiently rich to allow for a 'freedom of choice' perception of the subject. As good methodology dictates, the reliability and usability of the scenario is to be tested in a pilot study.

Stress and mood induction method. Crisis situations almost automatically involve stress, adjustment to novel situations is necessary and decisions are inevitably based on incomplete and sometimes even faulty information. An example to induce stress experimentally is the Trier Social Stress Test (TSST, Roelofs et al., 2005). However, also the increase of workload can be chosen as stress induction method. By letting participants perform two parallel tasks (for example handling a crisis situation while also responding to auditory signals), workload increases, which is known to induce stress levels (e.g., Klein, 1996). Mood could be manipulated by the use of subliminal priming of negative or positively valenced faces (e.g., Ruys & Stapel).

Tool to demonstrate the theory. The final step of the research is to contribute to the development of a software agent that predicts the decision making competence of an individual user and that gives feedback about the quality of the decision making process by combining the results from the aforementioned phases. The software agent[2] is to autonomously monitor and assist a human decision maker, in a more coaching role than usually associated with agents as personal assistants (e.g., Nwana, 1996). The feasibility of such an agent hinges on:
-measuring a person's mood, subjective stress and arousal level on the job in a non-intrusive, reliable manner.
-determining the amount of risk involved with the decision.

---

[2] For background literature on software agents see for example Weiss (1999); for background literature on intelligent systems see for example Russel and Norvig (2003).

-determining in an experimentally grounded manner about how a specific mood state, subjective stress and arousal level influences a person's capability in the context of risk to perform a certain decision making task, by also taking the person's past performance into account.
-warning to inform a person about his or her possible biased performance.
-interfering in more extreme circumstances through a mixed initiative approach in which the software agent deliberately manipulates the human's decision making process and is able to (i) provide instructions to improve the quality of the current decision making task, or (ii) explicitly abort the current decision making task to prevent a decision making process of too low quality.
-In addition a situation needs to be created which fosters a working, trust-based relationship between the person and the software agent.

The tool can be used for example in the context of a flood crisis, in which the mayor of a certain city is confronted with a large amount of decisions. The tool can provide the mayor with information about his or her decision making process quality, based on which the mayor can decide to let someone else perform certain actions or tasks or to rethink the decision.

5. Discussion

Decision making is a very important topic in safety situations, especially when the impact of the decision is tremendous. Human factors play an important role in human decision making and this is expected to be even more so during crisis situations. In the socio-technical approach of the D-CIS Lab, and especially in the context of the Interactive Collaborative Information Systems (ICIS) project, one aim is to improve human effectiveness and in particular to investigate how human factors, such as subjective stress, arousal and mood, influence phenomena such as decision making. Eventually, the aim is to improve the effectiveness of actor agent teams by building on the advantages of both actors and agents while addressing (situated) shortcomings.

The expected impact of our substantiated claim is the improvement of the quality of a decision in a crisis situation. In a crisis situation, one has to use the resources that are available and the people that are present. Subjective stress and arousal play an important role in these situations. Furthermore, these people might not be specialized in crisis situations, and therefore they cannot rely on a large amount of experience. Instead, they need to be assisted as much as possible. This might be accomplished for example by the use of feedback, provided during the performance of the task. Also important is to investigate the influence of more subtle human factors such as mood. Since mood has proven to impact cognition in several ways, including judgment and decision making, it is of significant relevance to point out the specific influences and to construct a method to measure mood and to give feedback, so that human decision making can be optimized.

Our perspective detailed in this paper on naturalistic decision making in crisis situations provides a basis for fundamental and pragmatic research results. The focus of our work is on crisis situations and management functions; previous research has not focused on this group of decision makers or on this particular situation as such. Included in this investigation is whether risk plays a role in the decision making process. The combination of these factors is expected to reliably replicate a real life situation in which decisions have to be made. The emphasis is leveraging fundamental research into applied research: the final goal is to develop a software agent that can measure a person's subjective stress level, arousal level or mood state and that can predict decision making performance in a certain situation in which time constrained decision making is required. This tool should also be able to give feedback or intervene with the decision making process. This tool is to be used for further research in the field of naturalistic decision making and for operational purposes both. Additional research should aim at improving the tool for use in the field, for example by focusing on feedback methods for improving performance, since the way of providing feedback is essential (Bickmore & Schulman, 2007).

Acknowledgements

The research ideas reported in this position paper are partly addressed in the Interactive Collaborative Information Systems (ICIS) project (http://www.icis.decis.nl), supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024. The ICIS project is hosted by the D-CIS Lab (http://www.decis.nl); the open research partnership of Thales Nederland, the Delft University of Technology, the University of Amsterdam and the Netherlands Organization for Applied Scientific Research TNO.

6. References

Bickmore, T., & Schulman, D., "Practical Approaches to Comforting Users with Relational Agents", *Conference on Human Factors in Computing Systems*, San Jose, CA, 28 April – 3 May 2007, New York, ACM, p. 2291-2296.

Clore, G.L., Wyer, R.S., Dienes, B., Gasper, K., Gohm, C., & Isbell, L., "Affective feelings as feedback: Some cognitive consequences", in L.L. Martin & G.L. Clore (Eds.), *Theories of mood and cognition: a user's guide*, 2002, Mahwah, NJ, Erlbaum, p. 27-62.

Cohen, S., Kessler, R.C., & Gordon, L.U., "Strategies for measuring stress in studies of psychiatric and physical disorders", in S. Cohen, R.C. Kessler, & L.U. Gordon (Eds.), *Measuring Stress: A Guide for Health and Social Scientists*, 1995, New York, Oxford University Press, p. 3-28.

Estrada, C.A. Isen, A.M., & Young, M.J., "Positive Affect Facilitates Integration of Information and Decreases Anchoring in Reasoning among Physicians", *Organizational behavior and human decision processes*, vol. 72, no. 1, 1997, p. 117-135.

Gasper, K., & Clore, G.L., "Attending to the big picture: mood and global versus local processing of visual information", *Psychological Science*, vol. 13, no. 1, 2002, p. 34-40.

Gillis, J.S., “Effects of life stress and dysphoria on complex judgments”, *Psychological Reports*, vol. 72, no. 3, 1993, p. 1355-1363.

Hammond, K.R., *Judgments Under Stress*, New York, Oxford University Press Inc, 2000.

Hockey, G.R.J., “The effect of loud noise on attentional selectivity”, *Quarterly Journal of Experimental Psychology*, vol. 22, no 1, 1970, p. 28–36.

Huey, B.M., & Wickens, C.D., *Workload transition: Implications for Individual and Team Performance,* Washington, DC, National Academy Press, 1993, p. 94-121.

Isen, A.M., & Geva, N., “The Influence of Positive affect on Acceptable Levels of risk: The Person With a Large Canoe has a Large Worry”, *Organisational Behaviour and Human Decision Making Process*, vol. 39, no. 2, 1987, p. 45-154.

Isen, A.M., & Means, B., “The influence of positive affect on decision-making strategy”, *Social Cognition*, vol. 2, 1983, p. 18-31.

Isen, A.M., & Patrick, R., “The Effect of Positive Feelings on Risk-Taking: When the Chips are Down”, *Organisational Behaviour and Human Performance*, vol. 31, no. 2, 1983, p. 194-202.

Isen, A.M., “Positive affect and decision making”, in Lewis, M., & Haviland-Jones, J.M. (Eds.), *Handbook of Emotions*, 2000, New York, Guilford, p. 417-457.

Jong, W. & Johannink, R., *Als het dan toch gebeurt. Bestuurlijke ervaringen met crisisbeheersing*, In Dutch, Enschede, Bestuurlijk Netwerk Crisisbeheersing, Universiteit Twente, 2007.

Kennedy, R.S., & Coulter, X.B., “The interactions among stress, vigilance, and task complexity”. *Human Factors*, vol. 17, no. 1, 1975, p. 106-109.

Klein, G., “The effect of acute stressors on decision making”, in Driskell, J.E., & Salas, E. (Eds.), *Stress and Human Performance*, 1996, Mahwah, NJ, Lawrence Erlbaum Associates Inc., p. 49-88.

Kontogiannis, T., & Kossiavelou, Z., “Stress and team performance: principles and challenges for intelligent decision aids”, *Safety Science*, vol. 33, no. 3, 1999, p. 103-128.

Kowalski-Trakofler, K.M., & Vaught, C., & Sharf, T., “Judgment and decision making under stress: an overview for emergency managers”, *International Journal of Emergency Management*, vol. 1, no. 3, 2003, p. 278-289.

Leith, K.P., & Baumeister, R.F., “Why bad moods increase self-defeating behavior? Emotion, risk taking, and self-regulation”. *Journal of Personality and Social Psychology*, vol. 71, no. 6, 1996, p. 1250-1267.

Mano, H., “Risk taking, framing effects, and affect”, *Organizational Behavior and Human Decision Processes*, vol. 57, no. 1, 1994, p. 38–58.

McNair, D.M., Lorr, M., & Droppleman, L.F., *The profile of Mood States (POMS) Manual*, San Diego, CA, EdITS, 1992.

Nwana, H.S, “Software Agents: an Overview”, *Knowledge Engineering Review*, vol. 11, no. 3, 1996, p. 205-244.

Poulton, E.C., “Arousing environmental stresses can improve performance, whatever people say”, *Aviation, Space, and Environmental Medicine*, vol. 47, 1976, p. 1193-1204.

Raghunathan, R., & Pham, M.T., “All negative moods are not equal: motivational influences of anxiety and sadness on decision making”, *Organizational Behavior and Human Decision Processes*, vol. 79, no. 1, 1999, p. 56-77.

Roelofs, K., Elzinga, B.M., & Rotteveel, M., “The effects of stress-induced cortisol responses on approach-avoidance behaviour”, *Psychoneuroendocrinology*, vol. 30, no. 7, 2005, p. 665-677.

Rondeel, E.W.M, In the mood for change: the influence of mood on change blindness, unpublished master’s thesis, Radboud University Nijmegen, Netherlands, 2007.

Russel, S. & Norvig, P., *Artificial Intelligence, a modern approach*, New Jersey, Prentice Hall, 2003.

Ruys, K. I., & Stapel, D. A., “Emotion elicitor or emotion messenger?: Subliminal exposure to two faces of facial expressions”, *Psychological Science*, in press.

Schwarz, N., “Situated cognition and the wisdom of feelings: cognitive tuning”, in Feldman, L., & Salovey, P. (Eds.), *The wisdom in feelings*, 2002, New York, Guilford, p. 144-166.

Thompson, C.A., Fundamental social motives and risk-taking tendencies, Florida State University, D-Scholarship Repository, Article #115, 2005, http://dscholarship.lib.fsu.edu/undergrad/115.

Todd, P.M., & Gigerenzer, G., “Putting Naturalistic Decision Making into the Adaptive Toolbox”, *Journal of Behavioral Decision Making*, vol. 14, no. 5, 2001, p. 353-384.

Tversky, A., & Kahneman, D., “Judgment under uncertainty: heuristics and biases”, in Shafir, E. (Eds.), *Preference, Belief, and similarity: Selected Writings*, 2004, London, MIT Press, p. 203-220.

Von Neumann, J., & Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton, NJ, Princeton University Press, 1974.

Weiss, G., *Multiagent systems: a modern approach to distributed artificial intelligence,* London, MIT Press, 1999.

Whyte, G., “Perspectives on Naturalistic Decision Making from Organizational Behavior”, *Journal of Behavioral Decision Making*, vol. 14, no. 5, 2001, p. 353-384.

Williams, S., “The Impact of Mood on Managerial Perceptions”. *Research and Practice in Human Resource Management*, vol. 12, no. 2, 2004, p. 128-139.

Behaviour and norms: some experimental results for complex systems modeling

Roger Waldeck

Telecom-Bretagne, Departement LUSSI, Technopole Brest Iroise CS 83818, F-29238 Brest Cedex 3, France
and ICI (EA - 2652)-Université de Bretagne Occidentale
Roger.waldeck@enst-bretagne.fr

ABSTRACT. *This paper is a review of some experimental results suggesting a way for modeling social interactions. We focus on two classical games, the ultimatum and the dictator game. Using observed facts in laboratory, we ask whether individual rationality could be explained by the presence of a norm of behavior. Norms suggest to individuals what is the appropriate behaviour in social interactions. They serve as a reference point for triggering sanctions by other players or emotional feelings for deviants. Actions are judged according to a norm and norms are themselves context dependent (actions available, framing effects, culture, decisional context). Norms emerge. Belief about the right « norm » depends on everyday observations. Behaviour is shaped by norms and norms by behaviour leading to complexity especially when interaction are localized.*

RESUME**.** *Nous passons en revue des résultats sur les jeux de l'ultimatum et du dictateur pour appréhender la modélisation des interactions sociales. On montre que le choix des actions des individus ne se fait pas seulement par rapport à la distribution des conséquences (présence ou non de préférences sociales) mais que cette distribution est souvent jugée par rapport à des critères normatifs dépendants du contexte (ce qu'il est normal de faire dans une situation pour une population donnée). La présence ou non d'alternatives de choix compatibles avec la norme affectera le jugement porté par les individus sur une distribution particulière de conséquences. De plus, les croyances quant aux comportements appropriés conditionnent les comportements eux-mêmes et entraînent l'apparition de normes qui peuvent évoluer en fonction du contexte par un phénomène d'apprentissage de la nouvelle norme.*

KEY WORDS: *norms , rationality, complexity, Ultimatum and Dictator game.*

MOTS-CLÉS: *normes , rationalité, complexité, jeux de l'Ultimatum et du Dictateur.*

1 Introduction

There is a non trivial link between norms and individual behaviour. From the individual point of view, norms prescribe how individuals should behave in a given circumstance. They shape beliefs by suggesting the appropriate behaviour in a social context and by setting the frame in which actions are judged.. Olstrom [2000] states that “a norm,…, is the propensity to feel shame and to anticipate sanctions by others at the thought of behaving in a certain , forbidden way…”. From an individualism methodology point of view, one can say that norms are emerging from the actions of interacting individuals (Boudon [1998], Opp [2001]). Since behaviour is shaped by norms and norms by behaviour there is a non trivial retroaction leading to complexity. We will review some experimental results which will suggest how the usual utilitarian framework is affected by norms. We focus on two classical games, the ultimatum game (UG in the following) and dictator game (DG) which have received much attention in the past two decades. By surveying observed facts in the laboratory, we discuss the alternative hypotheses concerning individual preference modeling and the formation of expectations with respect to norms.

There is compelling evidence that individuals adhere to cooperative norms (empirically or factually observed) both in repeated and non repeated interactions. In public good games or prisoner dilemma, where individual and collective interests are in conflict, people cooperate although “rational” behaviour would suggest them to deviate (Leydyard [1995]). Since Adam Smith, economic theory has considered that self-interest in the driving factor of homo-eoconomicus. The selfishness approach may explain cooperation by the repetition of the game (Axelrod and Hamilton [1981], Fudenberg and Maskin [1986] , Kreps *et al.* [1982]. In that case, cooperation is grounded on the concept of reciprocal altruism (Trivers [1971]): cooperation in bilateral interactions occurs only for the gain of future reciprocation. However self-regarding behaviour does not explain cooperation involving anonymous interactions in one shot games and experimental studies have shown that considering self-interest alone is not enough for explaining behaviour observed in the laboratory[1]. Two types of theories have been put forward to explain experimental data. Social preference models (Fehr et Schmitt [1999], Bolton et Ockenfels [2000]) posit that individuals are not only interested in their own gains but also by the distribution of gains among individuals. Reciprocal models (Rabin [1993], Dufwenberg and Kirchsteiger [1999], Falk and Fischbacher [2000], Charness and Rabin [2002] Bowles and Gintis [2003]) posit that the gains of others enter positively or negatively in the individual preference function depending on whether “others” are considered as being nice or mean to the player. Models based on reciprocity involve a concept of equilibrium based on psychological games (Geanakoplos *et al.* [1989]). Utility of a player is now a combination of the monetary outcome of the play plus a reward or punishment function depending on

[1] The list would be to long but some examples are Camerer and Thaler [1995], Dawes and Thaler [1988], Camerer [2003], Fischbacher U. *et al.* [2001], Fehr and Gächter [2000], [2004], Fehr and Fischbacher [2004], Ledyard [1995].

intentions in the strategic play. Since intentions require beliefs of higher order (beliefs of the player on what others will play, belief about what others think the player will play, etc…), the cognitive capacity required concerning coherence of beliefs of different levels gives to norms a crucial role. They reduce the cognitive overload needed to "compute" the best action and enable individuals to correlate their actions on a common belief. In some sense, a norm becomes a shared belief on how to play. However, the concept requires that the adequate normative behaviour must be evident for at least some players, actions must be observed, and at least some people must be willing to punish deviants from the desired norm so that it becomes popular or self-sustained. Otherwise, actions may converge to an undesired norm (full defection in public good games for example).

In section 2, we present the experimental results of the Ultimatum game. In section 3, we consider whether the proposed split relates to a concern for fairness or merely to a strategic behavior. The dictator game allows us to discriminate (imperfectly) between both assumptions and show us that social distance and some reciprocal preferences driven by expectation exist. In section 4, we consider the responders' rejection strategy and show that a lot can be explained by negative reciprocity although negative reciprocity is framed by norms. Notably causal attribution and contextual factors affect the rejection rate. In section 5, we conclude.

2 Experimental results and preference modeling

Individual rationality is explored in a game theoretical context by the use of an UG. Suppose that two individuals must share a pie of S =100 ECU in a one shot interaction. In all the following games, players play for real money (ECU are transformed into real money according to an exchange rate known to all individuals). Individual 1 (called the proposer in the following) offers a split (S-x, x) to the second player (called the responder) who may accept or refuse it. If the responder refuses they both end up with (0,0). Any split of (S-x, x) can be supported as an equilibrium given a proposer's belief about the responder's rejection rate. However, the subgame perfect equilibrium would predict (99, 1) as a rational outcome. The argument goes as follow: the responder would accept any positive offer if he has a preference structure characterized by an increasing utility function for money. Under the assumption that the proposer knows that the responder is rational and selfish, a selfish proposer offers the lowest positive unit of money in play to the responder.[2] Many experimental results show the following (Camerer [2003]): the mode and median of the distribution are generally around 40-50 % of the initial pie and the mean offer around 30 %. Only a few offers are in low categories 0%, 1-10 % or the hyper fair category 51-100 % . Offers between 40-50 % are seldom rejected but offers under approximately 20 % or so of the initial stake

[2] (100, 0) is another subgame perfect equilibrium where the responder accepts 0. However such an equilibrium prediction seems even less robust.

are rejected half of the time. These results are robust to the stake and notably the subgame perfect equilibrium is not the outcome (Cameron [1999], Slonim and Roth [1998]). How can we explain this divergence from theory? Since for one shot games, no incentive for intertemporal utility maximization exists, one assumption is that these offers reflect proposers' fairness consideration. Another assumption is that offers are strategic and maximize some utility of money given some beliefs about the responders' rejections rate. A DG (dictator game) discriminates between both assumptions since in this case the responder has no other choice but to accept the offer. In a DG selfish proposers should give zero. By contrast, any positive offer reflects a true concern for comparative payoffs and should be a measure expressing preferences for fairness. To present more formally the approach, consider the following simplification of UG where the proposer has only three actions. Either a split of a = (80,20) or b = (50,50) or c = (20,80). Multi-criteria decision analysis postulates that an individual has a set of actions A= {a, b, c} and a set of dimension N={1,2,..n} along which the agent evaluates each action. Moreover, for each dimension $j \in \{1,...,n\}$, the individual has a preference relation $\bullet_j$ over his set of actions in A defining a preference profile p = ($\bullet_1, \ldots, \bullet_n$). The problem for the agent consists in finding a global preference function $\bullet$ defined on the profile p. The dimensions reflect different objectives of the agent. If an agent maximizes only its own wealth then game theory traditionally assumes the existence of a Von Neumann and Morgenstern (VNM) utility function and the following normal form game is being played :

Proposer	**Responder** A	**Responder** R
a	$u_1(80),u_2(20)$	$u_1(0),u_2(0)$
b	$u_1(50),u_2(50)$	$u_1(0),u_2(0)$
c	$u_1(20),u_2(80)$	$u_1(0),u_2(0)$

Definition: *an egoist or self-interested individual is an agent that has a preference relation defined over its own outcome or function of its own outcomes .*

The agent may also have some distributional preference over outcomes. In that case some authors have assumed that agents have a preference profile defined by two criteria p = ($\bullet_1$, $\bullet_2$) the first criteria being his absolute gain and the second criteria a concern for social preferences including the other player's gain. For example, if player 1 is inequity averse then the absolute value of the difference between the profit of player 1 and 2 enters negatively in the utility function of player 1 (Fehr and Schmidt [1999]). More generally, the matrix becomes:

Proposer	Responder A	Responder R
a	$u_1(80,20)$, $u_2(80,20)$	$u_1(0,0)$, $u_2(0,0)$
b	$u_1(50,50)$, $u_2(50,50)$	$u_1(0,0)$, $u_2(0,0)$
c	$u_1(20,80)$, $u_2(80,20)$	$u_1(0,0)$, $u_2(0,0)$

Responders who reject positive amounts of money are irrational if they are only money maximizers. However, rejections may be explained by social preferences such as a preference for fairness. Strong reciprocity is another explanation that has been put forward in the literature (Gintis et al [2003], Gintis and Bowles [2001]):

Definition: *Strong reciprocity is the propensity of an individual to sacrifice resources to be kind toward those who are being kind or to sacrifice resources to punish those who are being unkind.*

Negative (strong) reciprocity is a motivation to adopt an action that harms someone else, at one's own material cost, because that person's intentional behaviour was perceived to be harmful to oneself. Perception, norms and context may thus determine what is considered as being kind or unkind. In that case, the general utilitarian model $u_i(x,1-x)$ is not sufficient and the function $u_i(x,1-x)$ should be context dependant. Notably the matrix of utility may change in different normative contexts. We show some experimental evidence on this. Ideally, we would like to shape the importance of the notion of norm in the definition of individual beliefs and preferences modeling.

3 Fairness or strategic offer?

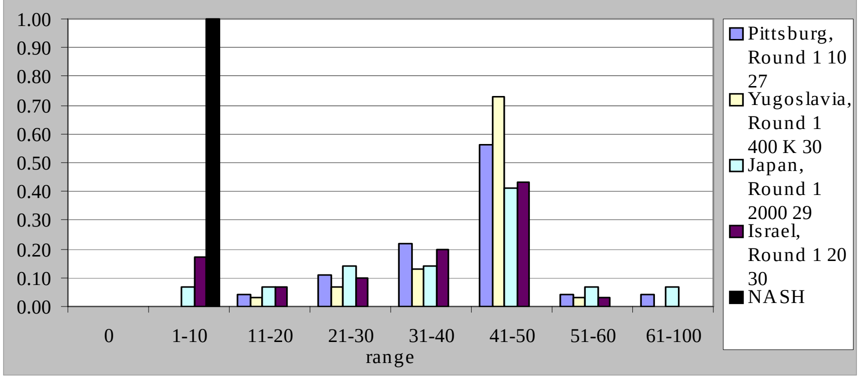
3.1 Some results for the UG game

Are individuals egoistic maximizers or altruistic beings? Roth *et al.* [1991] compared UG played in 4 developed countries (US, Japan, Yugoslavia and Israel) over 10 rounds in a stranger design. Modal offers were all at 50% of the pie in all 4 countries in the first round of play. Given the rejection rate for all data over 10 periods, the modal offer in the last period (round 10) corresponds to the offer maximising expected outcome. There is some rationality in the proposer's offer[3]. Countries' discrepancy in offers increased over periods except for Yugoslavia and United states which had no significant differences in offers in the first and last

[3] A stranger design insures that two players will meet only once in the repeated game. A stranger design should be equivalent to a one shot game since there is no incentive for an individual to build up a reputation. However, Abink *et al.* [2004] show that a group rationality exists. Some players reject more often when being able to build up a group reputation of being "tough" which in turn provides higher gains to responders in later periods. Indirect reciprocity (Alexander, [1987]) exists. Roth *et al.* [1991] did not control for this effect. This should reinforce the number of offers near 50%.

periods (mode at 50%) and remained the countries with highest offers (the mode of Israel was at 40% and Japan at 40 and 45%). A striking fact is that the distribution of offers in the first period are quite similar across countries (Israel and Japan being a bit more self-interested) revealing thereby some nearby distribution of beliefs (at the population level) about what should be an adequate offer. In addition, offers are far away from the subgame perfect equilibrium (figure 1) showing that proposers did not expect that responders were only selfish.

Figure 1: offers in 4 developped countries (1st round)



Moreover, rejections rates adapted over periods. Winter et Zamir [1997] showed a similar effect by considering an UG where real players played against a mix of virtual players (computer programs) and real players. They showed that real players adapt their beliefs and strategies to their environment: proposers learn to lower their offers in tough environments whereas responders learn to accept low offers in presence of tough proposers. As stated by Kahneman, Knetsch et Thaler [1986]: « Terms of exchange that are initially seen unfair may in time acquire the status of a reference transaction ». These first results seem to suggest that if some norm of fairness exists it is for strategic reasons (also Kagel *et al.* [1996]). Nevertheless, no country seems to be as egoistic as suggested by game theory. But high offers are merely a response to some belief about rejection rates and at best, only some individuals have fair preferences[4]. In other terms, most individuals are selfish but they know that an acceptable offer can only be around the fairness point. Henrich found probably was the most self-interested culture. The Machiguenga trib average offer was at 0.26 with a mode at 0.15 /0.25 (72%) and there was only 1 rejection (over 10). In a subsequent study, Henrich *et al*. [2003] showed that the ultimatum (UG) or public good (PG) games are played quite differently in 15 different tribs located in different part of the world. Mean offers vary between 26% to 58% (compared to 43-48% for games with students) and modal offers vary between 15% and 50% (students 50%). The tribs rank from society where the rejection in the UG

[4] Children are generally more self-interested than adult suggesting that what is considered as an acceptable offer evolves with age (Murnighan et Saxon [1998], Harbaugh *et al*. [2000]).

is very rare although offers from the proposer are very low. In other groups, rejection rates are substantial although offers are fair (status seeking competition). Except for two groups (Hadza and Sangu farmers), offers are generally different from a maximizing behaviour of proposers (even by integrating risk aversion) given responders' rejection rate. In order to explain data, other factors such as social preferences or "norms of playing the game" have to be considered. The authors show that a large part of the observed variability at the group level is explained by two factors: market integration and gains from cooperation. Market integration is highly correlated to the complexity structure of the group organization (family level of society through chiefdoms) which might constitute an explanation. The fair norm is more probable in groups with a high degree of market integration and large gains from cooperation. At the individual level, there are two possible interpretations of the result: either individuals maximize some generalized preference function over the distribution of outcomes. Or individuals have relative stable beliefs (perhaps more or less dispersed within the group) on the way they should act. These beliefs are inherited by every day life situations (for example hunter groups of large catches are more used to share the pie in a fair way). Thereby deviations from prescribed actions in these rule abiding groups, are more likely to trigger sanctions by others, reinforcing the respect for the rule. On the other hand, groups with relatively disconnected members are probably less constrained by norm abiding rules. In such groups, behaviour could be more dispersed (no true norm) or more self interested (an egoistic norm). For example, Machiguenga and Tsimane tribs live in societies with little cooperation, exchange or sharing beyond the family unit. There is little fear of social sanctions and little care about the public opinion. Machiguenga had the lowest offers with only one rejection and Tsimane had no rejection (mean offer at 0.37).

3.2 Comparing Ultimatum games and Dictator Games

If responders are constrained to accept proposers' offers, strategic offers disappear. Forsythe *et al.* [1994] observed that contributions are lower in the DG than in the UG. In DG, 35% of offers were below 10% of the endowment whereas none was observed in the UG. Approximately 20% of individuals offered 41% or more compared to more than 60% in UG. Mean offers were at approximately 23% compared to 44% for UG. Thus a portion of offers in UG are strategic but not all individuals are purely self-interested, expressing apparently a concern for fairness. However, several results show than when social distance between players increases the mean contribution decreases (Bohnet and Frey [1999], Hoffman *et al.* [1994,1996], Haley K.J and Fessler [2005]. The fear of being observed and judged by pears seems to be an important concern for individuals and triggers expectations about normative behaviour. In turn, it suggests that pro-social behaviour in the DG may be an expression of reciprocity driven by expectations about what is an appropriate gift by the dictator. Pro-social behaviour are supported by emotions (Bowles and Gintis [2003]). Emotions are triggered whenever an individual feels that he is acting meanly according to a certain intrinsic norm of behaviour. However behaviour must be observed. Thus in the case of the dictator such feelings is more or less activated by social distance or observable behaviour although the proposer has no reason to fear any monetary sanction. Some authors (Gneezy and Rustichini [2000], Barr [2001]) have suggested that social disapproval may work more effectively than monetary sanctions which are often seen by individuals as a price to be paid for not respecting the norm (Gneezy and Rustichini [2000]). Dana *et al.* [2006] showed that when an exit option is offered to proposers (they can leave the game without sharing the pie and get 9$ out of a pie 10$ leaving nothing to the receiver) then a significant proportion of agent prefer to take the exit option rather than sharing even if this implies a suboptimal allocation ((9,1) is better than (9,0)). These exit choices are clearly related to the fact that the dictator wants to avoid the judgement of the receiver (even for an anonymous game). When a new game is played where the dictator is insured that the origin of the gift will be hidden (i.e. the receiver will just be told that he gets an additional sum without knowing that it was left by the proposer) then the exit choice is no more used (except for 1 decision over 24). If actions are judged with respect to norms, beliefs about the appropriate norm will although depend on the context in which actions are taken. Chery *et al*. [2002] show that depending on the effort made by the proposers to win their income, the split will be different. In a double blind treatment with effort, contributions fall to zero in the DG. The context affects thus the social preference of individuals and they may become completely selfish under certain conditions. A way of modelling the agents' preferences is to introduce a third attribute into the preference function of a player which is an expression of emotions like guilt, shame , empathy (Bowles and Gintis [2003], Gordon *et al*. [2005] , Phan and Waldeck [2008]). However such a component may be cultural, context and even individual specific. Moreover the activation of such a component will depend on the feeling (beliefs) about the appropriate norm.

4 Rationality on the rejection?

Why do individuals reject a substantial amount of money in UG? Two possible explanations are inequity aversion (Fehr et Schmitt (1999], Bolton et Ockenfels [2000]) or negative reciprocity (Falk and Fischbacher [2000]). Falk et Fischbacher propose an UG that discriminates between the two hypotheses. They looked at the responders' rejection rates when the proposer plays strategy U depending on an outcome if the proposer would have played D (figure 2). Outcomes are represented by a couple of numbers where the first number represents the outcome to the proposer and the second to the responder. They proposed 4 treatments by varying the possible outcome for the path D (for proposer) followed by A (for responder). The four treatments were: (.,.) = {(5,5), (2,8), (8,2), (10,0)}. If the responder is only concerned by the distribution of consequences, the rejection rate if U is chosen, should not depend on what happens if he chooses D. Indeed, Falk and Fischbacher showed that rejection rates (RR) when U depends on the outcome for the path (D,A) that is: RR(5,5) > RR(2,8) > RR(8,2) > RR(10,0). Since RR (10,0) > 0 some rejections may also be explained by fairness.

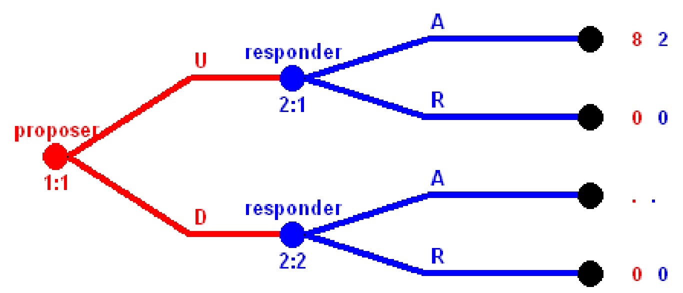


Figure 2

Responders may have two concerns when rejecting an offer. Either their beliefs about a norm of fairness are not met engendering disappointment. Or they may feel that the proposer's intention was mean. In the first case, rejection is independent on the origin of the offer. Blount (1995) shows that individuals have a need to infer causes and to assign responsibilities for why outcomes occur[5]. In a first study, she looked at the importance of causal attribution in the rejection of the responder. Three attribution processes were considered: Nature, a non interested Third Party Condition (3rd PC) and a self Interested proposer Condition (IC) [6]. Responders were asked to fix their Minimum Acceptable Offer (MAO) which was the lowest amount under which the offer was rejected. If Nature is the origin of the attribution then the observed MAO are near the subgame perfect equilibrium. There was no significant difference in MAOs between the 3rd PC and the IC. One possible explanation is first in the responders' expectations: in the 3rd PC, responders expect a symmetric distribution of offers with a peak at 50% whereas the expected distribution is right skewed for the IC. MAO may thus be strategically lower in the IC. A second explanation is in the importance of relative payoff comparison for the responder: in the 3rd PC, responders are less concerned by comparative payoffs (because there is no intention involved in the split). This effect should induce lower MAOs in the 3rd PC than in the IC. Indeed, this latter involves intentions and a greater responder's willingness to compare payoffs and to reject low offers. Observations show that the two effects compensate. An additional explanation of high MAOs in the 3rd PC is the breaking of the expected fairness norm. Blount advanced then the hypothesis that if the stress is put on the impact of responders' rejection on the passive player 1 (not

[5] See also Kagel *et al.* [1996] for similar results on the impact of intentions and norm breaking on rejections.

[6] In the 3rd PC, the offer was made by an non-interested third party whereas rejection affected a passive player 1 (in the role of proposer of the UG) different from the 3rd party.

responsible for the split) in the 3rd PC, it should then reduce MAOs. In a 3rd study, she varied the procedure of the experiment. Instead of asking for a statement of MAO, responders were made clear by filling out a questionnaire (like the question below) of the impact of rejections of the offer in the 3rd PC on an innocent player 1. In this 3rd study, the 3rd PC showed significantly lower MAO than the IC (with same procedure). It shows that stressing on the impact of receiver's decision on an innocent player, may change the reference point of evaluation (whereas all other treatments i.e. MAO statements and interested party condition with questionnaire showed no difference). Blount and Bazerman [1996] found a similar effect of the elicitation procedure. They showed to subjects questions like : *suppose that the other player proposes to take 6 dollars and gives you 4 ; you accept the 4 dollars or reject and get zero?* The MAO reported in the condition asking for MAO was 4 (out of 10) while in the second condition it was 2.33 showing a framing effect of changing the reference point. By stating MAO, the reference point is the fair norm (for those "groups"), whereas answering the question, the reference point becomes zero.

5 Conclusion

The experiments presented suggest that individuals in a game are not only concerned by the distribution over outcomes (selfish or social utility) but that the concern of doing the right thing (i.e. meeting the expectation of others) is important. Actions are judged according to a norm and norms are themselves context dependent (actions available, framing effect, culture, decisional context). Norms emerge. Belief about the right « norm » of behaviour depends on everyday observations of individuals. An implication is that individuals may have different feelings about the right norm especially in local interaction structure where information is localized. Expectations are evolving according to observations. In the case where beliefs about the right norm are shared, each individual propensity to adopt the norm may depend on sanctions (real or moral) when not behaving as suggested by the norm and diminish the individual utility. Moral costs may by proportional to the number of adopters of the norm. The following example may highlight the process: everybody in a city may be convinced that the socially collective optimal behaviour is not to steal. If the norm is respected then an individual caught when steeling may feel ashamed. But if every body else is steeling in the city the shame may disappear. Gordon *et al.* [2005] and Phan and Waldeck [2008] showed that polymorphic equilibria with cooperation may exist, depending on the distribution of moral cost. Convergence to this equilibrium depends on the number of initial cooperators .

REFERENCES

Abbink K., Sadrieh A., Zamir S. (2004). Fairness, Public Good, and Emotional Aspects of Punishment Behavior. Theory and Decision Volume 57, 1.

Alexander R. D. (1987), The biology of Moral Systems; New York: Aldine.

Axelrod, R. (1984). The Evolution of Cooperation, Basic Books.

Axelrod R., Hamilton W. D. (1981). The evolution of cooperation. Science, vol. 211, 27. , 130-1396

Barr A. (2001). Social dilemma, Shame Based Sanctions and Shamelessness: Experimental Results from rural Zimbabwe. Oxford University.

Blount S. (1995) When social outcomes aren't fair: the effect of causal attributions on preferences.

Blount S. Bazerman M.H. (1996). The inconsistent evaluation of absolute versus comparative payoffs in labor supply and bargaining. Journal of Economic Behavior & Organization, 30 (2), 227-240.

Bolton and Axel Ockenfels (2000). ERC: A theory of equity, reciprocity and competition, American Economic Review , 90, 166 - 193.

Bohnet I., Frey B (1999). Social distance and other-regarding behavior. American Economic Review, 89, 1, 355-339 .

Boudon R. (1998). Au delà du modèle du choix rationnel' ? Dans les modèles de l'action , sous la direction de B . Saint-Sernin, E. Picavet, R. Fillieule et P. Demeulenaere, *Paris Puf.*

Bowles, S. and H. Gintis (2001). Prosocial Emotions. Economy as a Complex Evolving System, III, in honor of Kenneth Arrow, Santa Fe Institute.

Bowles, S. and H. Gintis (2003). The Origins of Human Cooperation. The Genetic and Cultural Origins of Cooperation. P. Hammerstein. Cambridge, MIT Press.

Camerer C.F. (2003). Behavioral game theory , experiments in strategic interaction . Princeton University Press.

Camerer C.F. and Thaler R.H. (1995). Anomalies: Ultimatum, dictators and Manners. The Journal of Economic Perspectives, 9,2, 209-219.

Cameron L. (1999). Raising the stakes in the ultimatum games: experimental evidence from Indonesia. Economic Inquiry, 27, 47-59

Charness, Gary, and Rabin, Matthew (2002); "Understanding Social Preferences with Simple Tests", The Quarterly Journal of Economics 117 (3), 817 - 869

Chery T.L., Frykblom P., Shogren J.F. "Hardone the dictator , The American Economic Review, 92 (4), 1218-1221.

Dawes R.M., Thaler R.H. (1988). Anomalies: cooperation . The Journal of Economic Perspectives, 2, 3, 187-197.

Dana J., Cain D.M. and Dawes R.M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. Organizational Behavior and Human Decision Processes, 100 (2), 2006, 93-201.

Dufwenberg M., Kirchsteiger G. (1999) A theory of sequentail reciprocity. Discussion Paper, CentER, Tilburg University.

Elster, J. (1989). "Social Norms and Economic Theory." Journal of Economic Perspectives 3: 99-117.

Falk A., et Fischbacher U. (2000) Distributional Consequences and Intentions in a model of reciprocity.

Fehr E. et Fischbacher U. (2004). Third party punishment and social norms. Evolution and Human Behavior , 25, 63-87.

Fehr E. et Gächter S. (1999). Collective Action as a Social Exchange. Journal of Economic Behavior and Organization 39: 341-369.

Fehr E. et Gächter S. (2000). Cooperation and Punishment in Public Goods Experiments. American Economic Review 90(4): 980-994.

Fehr E. et Gächter S. (2004). "Social norms and human cooperation." Trends in Cognitive Sciences 8(4): 185-190

Fehr E. et Schmidt K.M. (1999). A Theory Of Fairness, Competition, And Cooperation. The Quarterly Journal of Economics, MIT Press, vol. 114(3), 817-868.

Fischbacher U., Gächter S. et Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public good experiment Economics Letter Elsevier, vol. 71(3), pages 397-404

Forsythe R., Horowitz J. Savin N.E., Sefton M. (1994). Replicability Fairness and Pay in experiments with simple Bargaining Games, 6(3) , 347-69.

Fudenberg D. , Maskin E.S. 1986. The folk Theorem in Repeated Games with Discounting or with Incomplete Information. Econometrica 54, 533-554.

Genakoplos J., Pearce D. Staccheti E. (1989). Psychological games and sequential rationality. Games and Economic Behaviour 1, 60-79.

Gintis H., Bowles S., Boyd R., Fehr E. (2003). Explaining altruistic behaviour in humans. Evolution and Human Behavior 24, 153-172.

Uri Gneezy; Aldo Rustichini. (2000) A fine is a price The Journal of Legal Studies, Vol. 29, No. 1, 1-17.

Gordon M., Phan D., Waldeck R. , Nadal J.P. (2005). Cooperation and free riding with moral costs. , Advances in Cognitive Economics, NBU Series in Cognitive Sciences.

Haley K.J., Fessler D.M.T.(2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. Evolution and Human behavior 26, 2005, 245-256

Harbaugh W.T., Krause K.et Linday S. (2000). Children's bargaining behaviour: differences by age, gender and height.

Henrich J., Boyd R., Bowles S., Camerer C., Fehr E. and Gintis H. (2003), Foundations of Human Sociality: Ethnography and Experiments in Fifteen Small-scale Societies (Oxford University Press, Oxford).

Hoffman E., McCabe K., Shachat K. and Smith V. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games. Games and Economic Behavior, 7(3), 346-380

Homans G. C. (1974). Social Behavior. Its elementary forms. New York, Harcourt.

Kahneman D., Knetsch J., Thaler R. (1986): Fairness as a constraint on profit seeking: entitlements in the market. American Economic Review 76: 728-741.

Kagel, Kim et Moser (1996). Fairness, in Ultimatum Games with asymmetric information and asymmetric payoffs, Games and Economic Behavior 13 (1), P100-110.

Kreps D., Milgrom P., Roberts J., Wilson R. (1982). Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. Journal of Economic Theory 27, 245-252.

Ledyard J.O. (1995). Public goods : a survey of experimental research . Handbook of Experimental Economics. Editeurs, Hagel J.H. et Roth A.E. Princeton University Press, 111-194.

Levitt S.D. , List J.A. (2007). What do laboratory experiments reveal about the real world. Journal of Economic Perspectives 21 (2), 153-174.

Murnighan J.K. et Saxon M.S. (1998). Ultimatum bargaining by children and adults. Journal of Economic Psychology, 19, 415-45.

Novak M. et Sigmund K. (1998). The dynamics of indirect reciprocity, J of theor Biol 194, 561-574.

Offerman T. (2002) Hurting Hurts More Than Helping Helps. European Economic Review 46, 1423-37.

Opp K. D. (2001). How do norms emerge ? an outline of the theory. L'explication des normes sociales sous la direction de Raymond Boudon , Pierre Demulenaere et Ricardo Viale, Presse Universitaires de France.

Olstrom E. (2000). Collective action and the evolution of social norms . Journal of Economic perspectives 14 (3): 137-158.

Phan D. , Waldeck R. (2008), Social norms, Emotions and cooperation in groups. 14 International conference in Computing in economics and Finance

Rabin M. (1993). Incorporating Fairness into Game Theory and Economics, in American Economic Review 83,5: 1281-1302.

Roth A. E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Smuel Zamir (1991). Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." American Economic Review, 81, 1068-95.

Robert Slonim and Alvin E. Roth (1998). Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. Econometrica, Vol. 66, No. 3, pp. 569-596.

Trivers R.L. (1971), “The evolution of reciprocal altruism”, Quarterly Review of Biology 46:35-57

Von Neumann J. and Morgenstern O (1944). Theory of Games and Economic Behavior. Princeton University Press, 2007.

Winter E. et Zamir S. (1997), Japanese Economic Review , Vol. 56, No. 3, (2005), 363 – 385.

Zamir S. (2001) Rationality and Emotions in Ultimatum Bargaining Annales d’Economie et de Statistique n° 61.

"Knowledge in HCP" session

11 – 12:30 am

Thursday, June 12, 2008

Collections: a Context for Singularities?

Francis Rousseaux*, Alain Bonardi and Benjamin Roadley***

** Reims University, 51100, Reims, France*
francis.rousseaux@univ-reims.fr ;benjamin.roadley@univ-reims.fr

***IRCAM, Place Stravinsky, 75004, Paris, France*
alain.bonardi@ircam.fr

ABSTRACT. *Many of our modern computerized activities, may they be personal, professional or even artistic, involve searching, classifying and browsing large numbers of digital objects. The tools we have at hand, however, are poorly adapted as they are often too formal: we illustrate this matter in the first section of this article, with the example of multimedia collections. We then propose a software tool for dealing with digital collections in a less formal manner. Finally, we see that our software design is strongly backed up by both artistic and psychological knowledge concerning the ancient human activity of collecting, which we will see can be described as a metaphor for categorization in which two irreducible cognitive modes are at play: aspectual similarity and spatiotemporal proximity.*

RÉSUMÉ. *L'informatique « à objets » a été inventée pour simuler nos activités de rangement d'objets dans des structures de classes identifiées et étiquetées, et son succès fut, comme l'on sait, immédiat. Une tendance innovante se fait jour depuis peu, caractérisée par la mobilisation de l'informatique « à objets » pour ranger nos collections, considérées comme des amas d'objets en attente de rangement dans des classes ad hoc, qu'il s'agirait cette fois de construire dans le même mouvement. En effet, faire collection est une activité plus originaire que classer, dans la mesure où elle autorise l'expérimentation concurrentielle des concepts 1° en extension (dans le cadre d'une disposition spatiotemporelle, même provisoire et éphémère) et 2° en intension (dans la perspective d'un ordre abstrait de similarités, même métastable). Reste que l'on fait toujours collection de quelque chose, ce qui interdit un typage de l'activité indépendant de ses objets et trouble ainsi les habitudes du modélisateur.*

KEY WORDS: *collections, singularity, categorisation, cognitive modelling, contiguity.*

MOTS-CLÉS: *collection, singularité, catégorisation, modélisation cognitive, contiguïté.*

1. Multimedia collections

1.1 Technological context

Our modern WIMP-based interfaces were created in the early 70s, they were used on computers with low storage capacities, slow processing speed, relatively low connectivity and low resolution monitors. These computers were first used in offices and administrations, where the desktop metaphor fitted very well. Then, personal computers brought this kind of hardware to people's homes, and the desktop metaphor still fitted as computers were mainly used for editing and filing documents.

Since those times, the technology has leaped forward, and today a large portion of the population uses a computer and connects to the internet on a daily basis. Here in France, 9 out of 10 people in the 18-24 age group use a computer and the internet daily (*Les Français et l'ordinateur*, phone survey by TNS SOFRES for the group Casino / L'Hémicycle, 15-16/04 2005). Computers are equipped with high storage capacity hard drives, powerful processors, high bandwidth internet connections, to name but a few technological trends. These are still evolving but the fact is that today more and more people are using their computers not only for editing and filing documents, but also for collecting music, films, images, books... Large amounts of these can be stored on hard drives and DVD-ROMS. The contents can be downloaded from the internet, or imported from digital devices such as cameras, which have also become mainstream.

Not surprisingly, a huge market has emerged from these multimedia collections. We can now choose from a myriad of computerized tools which assist us in finding, retrieving, recording, creating, editing, browsing and classifying multimedia contents. The variety of tools at hand seems to fit with the variety of uses involved in multimedia computing, from the most creative ones - such as graphic design, audio synthesis, etc - to the most formal ones - classification in particular. However, there doesn't seem to be many tools bridging the gap between these two seemingly opposing polarities.

1.2 Collecting: between formalism and creativity

Let us illustrate this situation. First, let us suggest that looking for new material and classifying are two important processes involved in collecting. Indeed, when someone decides to start building a collection he usually already possesses a few items. Then, to extend this collection, new items must be added. In order to do so, the collector goes into the world and looks for these new items. Then as the collection builds up, the need to arrange the items into categories will become clearer, as the collection cannot simply remain a messy stack of unordered items.

So, in order to illustrate our point, let us describe a particular example: the music collector. As we have said, our collector will surely possess some initial items; these may be some CDs or vinyl records. His first action involved in extending his collection could be a visit to the record shop for example. Here, the music is classified conformingly to the record companies' desires, which can sometimes be confusing for our collector, who is a fan of Jimi Hendrix, and just does not know where to look for his albums: in the blues section? Rock section? Is there a 'sixties' section? Anyway, despite finding them rather practical at first sight, our collector didn't create these labels, and finds it difficult adapting to them. However, as he browses through the shop, he also notices some nicely illustrated records, and discovers new artists he is interested in because their records are sitting next to Jimi's. Finally, when he has bought enough music records, and come back home, he will be able to start arranging his collection in a very personal and satisfying manner, which will be pleasing to the eyes, and also allow him to retrieve items quickly.

If he had decided to collect digital music, and go online to find new items for his collection, the process would have been rather similar. Commercial music download sites allow the user to browse through predefined music categories, thus implementing a kind of virtual record shop with the same problems mentioned earlier. The search tool however can come in handy, and allow the user to search for the name of an artist, a song, an album or even musical genre. All these are still editorial information, which aren't necessarily the most useful to the collector. Then, when the music is downloaded, the album consists of a group of compressed audio files, containing preset meta-tags, again storing editorial information. When browsing these files in his audio player, the songs are defined and classified automatically, not always according to the collector's desires. His final attempt is then to create a set of folders on his disk, and arrange his items in these folders. But how does he name these folders? What if he wants to arrange and browse the items in multiple ways? What if a particular item doesn't fit in any folder, or could be placed in two or three different categories? Pachet has also described many problems in the area of Electronic Music Distribution (Pachet 2003).

As we see from this example, the tools that the everyday user has at hand are too formal, and are poorly adapted to the growing activity of collecting multimedia contents. Indeed, what we have said for music can also be said for the other kinds of media, and can also be said for information research, file sharing, etc.

Attempts have been made at putting the human user back in control of the collecting process, rather than relying purely on predefined categories and automated research algorithms. However, it has become obvious that the other extreme of handing complete control over to the user isn't optimal either. Let us take a look at online content sharing sites, such as the famous FlickR™. There is no categorization here, but there are three main strategies when looking for photos: date, location, tags. The first two are self-explanatory, but the tags are more interesting here. When someone uploads a photo to the website, they can link a certain number of keywords, called tags, to this photo. Then, we can either browse through the most popular tags, or type a tag into a textbox for a more precise search. The users then have complete freedom on the way they choose to define their photos. But the problem is that many photos aren't tagged, and the photos that are, often have poorly named tags, making them difficult to retrieve. Therefore, we believe that an optimal solution to the problem of digital collections could lie somewhere between these two polarities: predefined categories and total user creativity.

1.3 Examples of tools attempting to bridge the gap

MusicBrowser is a software which aims at indexing large and unknown music collections, and also helping the user find "interesting" music in these collections (Pachet 2006).

When digital sound files are imported into the system, they are analyzed, and a database of their acoustic properties is created / updated. Then the user can browse through the collection in a traditional manner, relying on editorial information. He can also create his own categories intuitively. He starts by creating a category, and giving it a name. This can be totally subjective if he wishes, he may call it "evening music", "happy music" or "favorite", etc. He then adds a few songs to this category, before asking the program to finish classifying, based on acoustic similarities. Of course, the more categories there are, and the more examples there are, the easier it is for the system to classify the entire collection. However, if there are mistakes, the user may simply move a song from one category to another, and ask the system to start again. This creative feedback loop, between user input and automated algorithms, will eventually lead to a satisfying classification for the user, who will have saved a lot of time in the process. He will then be able to create other classifications of the same collection if he wishes, and switch instantly between any of them. He may also share these classifications or download others.

IMEDIA is a research project focused on indexing large collections of photos, and interactive searching and browsing (Boujemaa 1999). When photos are added to the system, they are analyzed and a database of visual descriptors is created / updated. One of the main features of the program is allowing the user to search for similar photos. At first, a list of random images from the collection is displayed, the user may browse them, or view another set of random images. When he sees a photo he likes, he can select it and ask the system to find similar ones. For example, if he chooses a photo of a beach, then the system will display a list of photos of beaches. Once again, if the user isn't completely satisfied with the results, a "relevance feedback" system allows him to select the errors, and the system will take this into account in order to display a more relevant list of results.

In these two systems, we have noticed a creative feedback loop between the human user's input (starting point, examples, relevance feedback…) and the computer (automated algorithms for classifying and searching). This helps the user build and browse his collection in a constructive process, leading to a result which neither he nor the computer could have achieved alone. Also, both editorial information and semantic information (invisible to the user) are taken into account. *IMEDIA* and *MusicBrowser* address the problems of music collections, and photo collections, but the same ideas may be applied to other media collections, such as texts or videos, for example. It is only a case of finding the appropriate descriptors. Also, both these ideas, interactive searching and browsing, can be transposed to different media.

We can even think further, and imagine a common environment for collecting multimedia files. This could be a system with a generic layout and set of functionalities that would give birth to different programs specialized in collecting certain types of media. In the next section, we shall present a software prototype that we have implemented in order to experiment with this idea. As we shall see in the next section, we have tried to create a program more suitable to the particular process of collecting, which has an element of subjectivity, evolves over time and doesn't rely purely on similarities, as in the *IMEDIA* system for example. Indeed, we sometimes wish to expand our collection with something completely different, now how would we do that? We also believe that this process lies somewhere between formal classification/automated algorithms and total creativity. There are more and more examples of this, such as the two projects described previously, and we will try to take this process even further.

2. An expemimental software for the creation of multimedia collections

ReCollection is a computer program for searching, arranging and browsing digital content.

As our collecting activities vary from one context to another, it is too ambitious to seek a general solution to the problem. Rather, particular application areas must be defined and isolated, in order for a specific answer to be given, however always relying on a set of basic principles. Here, we shall discuss the software prototype we have created for the digital opera / open form opera *Alma Sola* (designed by Alain Bonardi, IRCAM, Paris and performed at *Le Cube,* Issy les Moulineaux, October 2005).

2.1 A useful metaphor: the art collection

Artists and philosophers have described some very particular characteristics of collections. One of those, as noted by Wajcman, is that of *excess* in a collection (Wajcman 1999). This means that the number of collected items exceeds the collector's capacity of memorization, but also of physical storage and exposition in the gallery. Thus, there is a need for at least one *reserve,* where the excess can be stored. For example, the George Pompidou National Museum of Modern Art, Paris, owns about 59000 artworks, making it one of the largest modern and contemporary art collections in Europe. Obviously, all the items cannot be *exposed* in the galleries at once, so a very large portion is stored in the reserves. Often, the items in reserve are stored in heaps, in random locations, and they aren't always labeled, which makes it difficult to find and retrieve objects.

The reserve allows us to handle the excess in collections, which is a problem in many of today's computer applications. Our multimedia collections, for example, are becoming very large and we are often losing control over them.

On the other hand, objects which are currently exposed are found in the *gallery*. Here, the objects follow a spatio-temporal arrangement defining a finite number of visitation paths. The closeness in space of certain artworks and the chronological order in which they are approached are set carefully by the curator, as they strongly influence the visitors' experience. This aspect is also very important, and we shall discuss it later in detail.

2.2 The reserve

The *ReCollection* software has two main modes: reserve and gallery. The reserve allows us to store our objects which aren't exposed in the gallery. There are many objects in the reserve, and these are not always labeled; also they are rarely arranged in an orderly and tidy manner. So when we visit the reserve, we have no choice but to wander around, picking up objects, inspecting and identifying them one at a time. The reserve can also be compared to the attic, in which our family possessions are stored similarly. As we explore our attic, we can happen to pick up an old photo album, which we had completely forgotten about. This item will surely bring back memories and emotions. We can then choose to keep this album under our arm, as we continue to explore the attic, or we can leave straight away, and put it on our fireplace, for example, making it visible to visitors. It is all these pleasant and familiar experiences which we believe can be recreated thanks to the modeling of the reserve in our computer program.

Figure 1 – The reserve

The user can create any number of reserves. However, he must create at least one, and store at least one object in this reserve. When he is in reserve mode, he can only view one object at a time. When he decides to view another object, it is chosen randomly from the remaining items in reserve. During a visit, each object is viewed only once. If the user wants to view an item he has already visited, he may go through the history of items on the left side of the screen, as shown in figure 2. When he finds an object of interest, he can move it to the gallery. It will then be removed from the reserve, and saved in memory, with a group of objects waiting to be imported in the gallery. Then, in gallery mode, the user will see this heap of objects, and will be able to import it in the desired gallery, at the desired location.

2.3 The objects

The items in the *Alma Sola* collection are made up of three components:

- a photo of the performance,
- a sound recording of a few seconds of the singing,
- a text, the line which is sang in the corresponding sound file.

These are all regular files stored on disk (bitmap, wave and .txt formats). Each item also has a name. In a more general context, the objects can be made up of any one of these types of media, a video (though not implemented in this version), or any combination of these.

Also, each object has a set of descriptors attached. There is a specific set of descriptors for each type of media, which describe the contents of the object, for example the average volume of the sound, the brightness of the photo, the number of words, etc. Depending on the application, we could also include editorial information, such as date, author, etc.

These descriptors may be assimilated to the private properties of traditional computer objects. But in the context of collecting objects, we also need to account for other properties that come from the activities in which these objects collectively engage.

2.4 The gallery

A collective activity involving a number of objects at a time is their relative arrangement in the gallery space. To the location of objects in this space, we have added their color; these two properties make up an extra conceptual layer which is the framework for the creation and management of our collections.

In *ReCollection*, there is always at least one gallery, and the user can create as many as he wishes. There is always at least one item in a gallery, some basic content that the user can interact with, a starting point for his collection.

The objects can be placed and arranged manually in the gallery space, using *click and move*, just as in common user interfaces. The user can also rely on two algorithms to automatically dispose the objects. The first one, inspired by *cataRT software* (Schwarz 2006), calculates the objects' positions and colors according to descriptors chosen by the user. The second calculates the positions depending on a sample of objects selected by the user. A Principal Components Analysis (PCA) finds out which descriptors vary most amongst the objects of the sample, the system can then rearrange the whole gallery according to these descriptors, as in the first method.

The arrangements resulting from the algorithmic calculations can always be modified manually in order to correct them (in the eventuality of rather subjective descriptors), to build up a global figure, or to bring items together. This way, through creative human-computer feedback loops, meaningful global figures can emerge through the arrangement in space of collected items, as well as local figures, soft pseudo-categories which are heaps of objects brought together by the system and/or the human user. These pseudo-categories are the building blocks for the classes the collection is implicitly aiming for. They are easily and constantly updated; items are added and removed instantly by being moved in space. They are loosely defined and never completely closed off from others, allowing some objects to be lost somewhere in between several heaps, when they cannot be placed in any one category. In a nutshell, this system allows for the creation of collections in which classes are in constant evolution, and are built by exploiting not only the objects' degree of similarity, but also their relative location in space and time.

Furthermore, the user may wish to search for objects in the gallery or in the reserve, in order to build on these categories, look for new kinds, or even fill in gaps in the gallery space. For this, the *ReCollection* system has two search tools he can use. The first is a simple 'keyword query', which searches for a keyword within the text or names of the objects. The second is a 'search by similarity'. The user selects an object, or group of objects, and the system searches for items which are similar (according to the descriptors). In both cases, the search is carried out in both the gallery and reserve, and a list of results is displayed in the gallery, ordered by similarity.

Once all the items of interest have been imported from the reserve, through browsing or searching, and once they have been arranged in the gallery space, the user has a first *disposition* he can play with. When he will browse the gallery space, his experience will be influenced by the fact that certain objects are close in space, and in time of visitation. Although this is interesting in itself, the system can help the user go further, by defining a set of guided visits, which are simply an order of visitation of selected objects in the gallery.

The type of interface we have chosen to implement these functionalities is a 2D zoomable user interface (ZUI), inspired by Ken Perlin's *Pad* (Perlin 1993). All objects are in the same 2D space, which has no borders. The point of view can be moved vertically and horizontally, and the user can zoom in and out. If he zooms in on an item, until it fills the screen, the sound is played back. This kind of interface has been experimented; it has obtained good results, and has been proven reliable (Guiard 2001). Its intuitive approach is seducing to us, particularly in our goal of intuitively collecting digital media. Finally, the spatial metaphor takes advantage of the users' spatial memory and cognitive abilities (Seegmiller 1977, Hasher 1979).

Figure 2 - Gallery mode

In the next part of this article, we shall discuss in detail some key characteristics of collections, as identified by artists, philosophers and psychologists. This theory is at the foundation of our work, and it demonstrates the novelty and usefulness of the collections metaphor in computer science.

3. The strange epistemological status of collections

Object-oriented computer science was invented to assist the task of classifying objects in a structure where different classes are distinguished (Perrot 1998, Granger 1994, Baudrillard 2005). As we all know, this innovation quickly became a success.

3.1 Collections, between order and disorder

Recently, an innovative trend is mobilizing computer objects for the organization of our collections, considered like a group of objects waiting to be organized in ad hoc classes that must be created simultaneously (Pachet 2004, Serra 2004, Rousseaux 2005).

Because our collections seem to be nearer to order than disorder, attempting to assimilate them in classes is not so surprising. At least, collections look like they are waiting for their completion within a classification order, with the aim of turning into canonic achieved structures made of objects and classes. But something is also resisting this assimilation, as artists and philosophers have always noticed.

3.2 Artists' fascination for collection regimes

As a matter of fact, artists and philosophers have always been fascinated by the rebellious nature of collections and have demonstrated this in their own way

(Benjamin 1989, Wajcman 1999, Pomian 1987, Tourangeau). Here, for example, is the analysis of Gérard Wajcman (Catalog *for the inaugural exhibit of the Maison Rouge*) on the status of excess in a collection:

"Excess in a collection does not mean disordered accumulation; it is a fundamental principle: for a collection to exist as such-in the collector's eyes the number of objects must exceed the physical possibilities of exposing and storing the entire collection at home. Therefore, someone who lives in a studio can have a collection: it is only necessary for him to have at least one work he cannot hang in his studio. That is why the reserve is an integral part of collections. Excess also applies to the capacity of memorization: for the collection to exist, it is necessary for the collector not to be able to remember all the works he owns. In fact, the number of objects he owns must be so important that it becomes too important, so that the collector can forget one of them or leave a part of his collection outside of his home. To say it differently, for a collection to exist, the collector must not have full control over his collection anymore." (Wajcman 1999).

4. Computer scientists and collections

Undoubtedly impressed by artists and philosophers who considered the strange status of collections, "object-oriented" computer program designers understood that computer modeling of object collections would necessarily involve the creation of hybrid structures including private characteristics – by which the collected objects are usually referred to – but also including characteristics that come from the activities in which these objects collectively engage.

4.1 A parsimonious, conservative, and indeed seductive, approach

Often, the approach implicitly chosen to characterize a collection is parsimonious and consists of over-determining the private referencing of the collected objects through a minimal description of the collective activity's context, even if it means predicting that the collection shall become a class or set of classes.

This practice presents the unquestionable advantage of not fundamentally opposing the traditional modeling of objects. However, it does not always live up to the collectors' high standards.

Here it is important to distinguish between figural and non-figural collections. This subtle distinction, introduced in the 1970s by Piaget and his research teams of child psychologists (Piaget & Inhelder 1980), brings more light to the situation. If it is certain that (non-figural) collections that adapt well to the aforementioned parsimonious approach exist, it is because they are completely independent of their spatial configuration. In that, they are already close to classification, of which they can only envy the formal completeness. On the other hand, there are collections we can label as *figural* because both their arrangement in space and the private properties of the collected objects determine their meaning.

4.2 Collections versus classes

In their book *La genèse des structures logiques élémentaires* (lit: *The Genesis of Elementary Logical Structures,* page 25), Jean Piaget and Bärbel Inhelder provide a precise distinction between figural and non-figural collections, which are still called classes or categorical collections. For the authors, a class requires only two categories or relationships, both necessary and sufficient, for its actual definition as a class:

1. The qualities common to its members and to those of the classes it belongs to, as well as the specific differences that distinguish its own members from the members of other classes (comprehension);
2. The relationship of a part to the whole (membership and inclusion) determined by the quantifiers ”all”, ”some” (including ”one”) and ”none” applied to the members of the class in question and to other members of the classes it belongs to, defined as extensions of that class.

For example, cats share in common several qualities owned by all cats, some being specific and some others belonging also to other animals. But no spatial considerations ever enter into such a definition: cats may be grouped or not in the space without any change concerning their class definition and properties (1) and (2). Piaget then introduces *figural collections*, in which meaning defined by properties (1) and (2) is linked to the spatial arrangement of its elements. A figural collection composes a figure, through the spatial relationships between its elements, whereas non-figural collections and classes are free of any figure.

4.3 Figural **versus** *non-figural collections*

It is precisely these figural collections that object-oriented computing is promising more and more an effective modeling of, pushed by an ever-growing social demand for on-line digital media browsing and information research amongst multiple sources (Pédauque, Rousseaux 2006).

But as we now understand, figural collections adapt poorly to their assimilation into non-figural collections or classes. Although according to Piaget, collections are destined to become classes, in the same way as subjects will grow psychologically so as to improve their cognitive capacity to classify. Still referring to Piaget, it is a *radical lack of differentiation* that nudges figural collections out of the classical modeling field.

5. Conclusion

Husserl used to say that consciousness is always consciousness of *something*, that consciousness always *pre-dates* the subject and the object, and *puts them together* in the process. There are no subjects or objects already existing independently that meet in the world to fill out a journal of experiences (the subject) and perhaps adapt to each other by induction. In the same fashion, we could say that a collection is always a *collection of something*, in that the original process of categorization is the activity of collecting, implacably mixing abstraction and spatio-temporal arrangements, and producing as many metastable categories.

The current models for information search are too formal, and they assume that the function and variables defining the categorization are known in advance. In practice, however, when searching for information, experimentation plays a good part in the activity, not due to technological limits, but because the searcher does not know all the parameters of the class he wants to create. He has hints, but these evolve as he sees the results of his search. The procedure is dynamic, but not totally random, and this is where the collection metaphor is interesting.

The collector's experimentation is always carried out by placing objects in temporary and metastable space/time. Here, the intension of the future category has an extensive figure in space/time. And this system of extension (the figure) gives as many ideas as it does constraints. What is remarkable is that when we collect something, we always have the choice between two systems of constraints, irreducible one to the other. This artificial indifferentiation for similarity/contiguity is the only possible kind of freedom allowing us to categorize by experimentation.

Our prototype implements these ideas by allowing the user to dispose his objects in 2D space. This arrangement may be manual, automated or both; it may be based on similarity, spatial proximity or both. A global figure may emerge from this arrangement, influencing the browsing and also the extension of the collection. Local figures emerge, which are the temporary pseudo-classes illustrating the pre-categorization building process of collecting. The art gallery metaphor fits very well, as it adds further meaning to the arrangement of the collected items in space, and models the excess in collections thanks to the reserve.

Through exploiting space in this way, the software interface takes advantage of our cognitive abilities in dealing with spatial information, and also our ability to collect information and acquire knowledge. Our next step is experimentation in order to validate our work. This could simply take the form of a series of sessions in which both novice and experimented users are asked to build up collections using the software. Through user-feedback, we will have a first idea of how well the interface is understood, how useful the users find it and how easy it is to use. If this experiment is a success, as we believe it will be, we will continue our research and bring it to the next level. Through integrating new functionality focused on indifferentiation for similarity/proximity, we will be able to build specific tools for a variety of applications in which the user's activity may be – at least metaphorically – described as building a *figural collection*.

6. References

BAUDRILLARD, J. (2005) *The System of Objects*. Verso.

BENJAMIN, W. (1989) *Paris, capitale du XIX^e siècle — le livre des passages*. Le Cerf.

BOUJEMAA, N. AND NASTAR, C. (1999) "Content-based image retrieval at the imedia group of the inria". *10th DELOS Workshop Audio-Visual Digital Libraries*. Santorini.

BRÉZILLON, P. (1999) "Context in Human-Machine Problem Solving: a Survey". *Knowledge Engineering Review*. 14, 1-34.

GRANGER, G-G. (1994) *Formes, opérations, objets.* VRIN.

GUIARD, Y. BOURGEOIS, F. MOTTET, D. BEAUDOIN-LAFON, M. (2001) "Beyond the 10-bit barrier : Fitts' law in multiscale electronic worlds". *Proc. IHM-HCI 2001*. Springer-Verlag.

PACHET, F. (2004) "Les nouveaux enjeux de la réification". *L'Objet*. 10(4).

PÉDAUQUE, R. http://rtp-doc.enssib.fr/rubrique.php3?id_rubrique=13.

FOX, D., PERLIN, K. (1993) "Pad: An alternative approach to the computer interface". *Proc. ACM SIGRAPH'93.*

PERROT, J-F. (1998) "Objets, classes et héritage : définitions". dans *Langages et modèles à objets Etat des recherches et perspectives*. Collection Didactique. INRIA. pages 3-31.

PEYRET, J-F. (2006) "Trouver le temps". *colloque Ecritures du Temps et de l'Interaction*. http://resonances2006.ircam.fr/?bio=57. IRCAM.

PIAGET, J., INHELDER, B. (1980) *La genèse des structures logiques élémentaires.* Delachaux et Niestlé.

POMIAN, K. (1987) *Collectionneurs, amateurs et curieux.* Gallimard.

ROUSSEAUX, F. (2006) *Singularités à l'œuvre.* Collection Eidétique. Delatour.

ROUSSEAUX, F. (2005) "La collection, un lieu privilégié pour penser ensemble singularité et synthèse". *Revue Espaces Temps*. http://www.espacestemps.net/document1836.html.

ROUSSEAUX, F. (2006) "Par delà les Connaissances inventées par les informaticiens : les Collections?". *Intellectica*. 2005/2-3. n° 41-42

SCHWARZ, D., BELLER, G., VERBRUGGHE, B., BRITTON, S. (2006) "Real-time corpus-based concatenative synthesis with catart". *DAFx.*

SEEGMILLER, D., MANDLER, J.M., DAY, J. (1977) "On the coding of spatial information". *Memory and Cognition.*

SERRA, X. (2005) "Towards a Roadmap for the Research in Music Technology". *ICMC* 2005. Barcelona.

TOURANGEAU, S. *Collection création, parcours désordonné, propos d'artistes sur la collection.* http://collections.ic.gc.ca/parcours/laboratoire/livre/creation.html.

WAJCMAN, G. (1999) *Collection.* Nous.

A contextual approach for a fake-centered rationality

Charles Tijus[*] -- Patrick Brézillon[**]

[*] *Laboratoire CHArt, University Paris 8,*
2 rue de la Liberté,
93526 St Denis, France
tijus@univ-paris8.fr
[**] *LIP6, University Paris 6,*
104 avenue du Président Kennedy,
75016 Paris, France
Patrick.Brezillon@lip6.fr

ABSTRACT. This paper presents the continuation of a work on an approach of rationality. We argued that a formalism for task modelling called Contextual Graphs allows to increase clarity, and rationality, of the decision-making process. However, it often happens that there is no path towards a solution when facing a new and unknown environment, or a new field of investigation. Generally, theories for enlarging the way of thinking in such situations are based on analogy or case-based reasoning. Such processes supposed modes of thinking (having possible sources for hints, then use them to solve the target problem) that differ from observations of natural human reasoning (having a target problem, then find a possible source for hint). We advocate for the production of target's counterfactuals as provider of a substitute that is then used for solution. Moreover, we show how contextual categorization and the context-based formalism for procedure representation can be used to model the process of finding rational routes for solution.

RÉSUMÉ. Ce papier présente la poursuite d'un travail sur la rationalité. Nous pensons qu'un formalisme pour modéliser les tâches appelé "Graphes Contextuels" permet d'augmenter la clarté, et la rationalité des processus de prise de décision. Cependant, il est facile de tomber dans des impasses dans des un environnements nouveaux et inconnus ou à un nouveau champ d'investigation. Généralement, l'ouverture de notre manière de penser sont alors basées sur l'analogie ou le raisonnement à base de cas, ce qui supposent des modes de pensées (qui sont des sources possibles pour des suppositions et alors sont utilisés pour résoudre le problème-cible) qui diffèrent des observations sur le raisonnement naturel humain (avoir un problème-cible, et alors trouver une source possible de suppositions). Nous proposons que la production de contrefactuels de la cible comme substitut utilisé comme solution. Nous montrons aussi comment la catégorisation contextuelle et les Graphes Contextuels pour la représentation des procédures peuvent être utilisés pour modéliser le processus de découverte de routes rationnelles vers une solution.

KEY WORDS: a maximum of six significant words should be chosen and presented as key words.

MOTS-CLÉS: un maximum de six mots significatifs doivent être isolés sous forme de mots-clés.

1. Introduction

Brézillon and Tijus (2006) show how people solve everyday tasks. The authors use Contextual Graphs (CxGs) as an epistemic tool that provides more clarity and rationality for decision-making. CxGs are already used in several domains such as medicine, ergonomics, psychology, army, information retrieval, computer security, road safety, etc. [Brézillon, 2003; Brézillon and Pomerol, 1999]. The reason is because CxGs describe the "how-to-do-it" in a uniform formalism that computes commonalities and differences between alternative procedures through the uniform representation of elements of reasoning and of contexts.

People also know how to solve tasks with an evolving context by adapting known procedures to the context at hand. Such adaptive processes usually are modelled by case-based or analogy-based systems, such as finding sources in memory to be adapted for solving the target task. Since the task at hand is close to already solved tasks, the modelling is close to a recognition task: to recognize the target task as a variant of a known task [Tijus, 2003]. The scope of CxGs is to finding the closest situation as a source and the corresponding graph to be adapted to the target task [Brezillon and Tijus, 2006].

However, people may make erroneous decisions and try to applied incorrect procedures, because they are in a dead-end and failed for a while to get a route to a solution. These problem-solving situations are also opportunities for learning while finding a route for solution. Learning occurs when there is perception of results that appear to be errors. Such learning increases clarity and rationality for decision-making as well as contextual dependency of solution. This paper presents how CxGs and contextual categorization might be helpful for the discovery of adaptive solutions that improve decision-making rationality.

Classical problem solving theories [Newell and Simon, 1972], such as search and planning in a space of problem states, are one way of learning with little use of known procedures. Conversely, finding a solution through analogy making is one of the human performances that computer scientists have tried to model and simulate on computer. By contrast, we explore a cognitive mechanism that is based on the realization of counterfactuals to get substitutes as sources. Differently of analogy, the process does not involve matching sources to be modified and adapted to target, but supposes target adaptation to finding sources.

Hereafter, the paper is organized in the following way. We first present the limits of analogy and its case-based variant and the contextual categorization approach. Second, we define what is the problem of finding substitutes for innovative solution. Then, we present the tools that could be useful for innovative solution.

2. The analogy in solution building

Sir Ernst Gombrich, a well-known art specialist, has been disserting about a very ordinary hobby horse. "Should we describe it as an 'image of a horse'?" Is it rather "a substitute for a horse", a "horse-like" that can be ridden by a child. Thus, for Gombrich, there is no analogy between the horse and the stick. Gombrich concluded that “substitution may precede portrayal and creation communication." What kind of process could lead to use a stick as a horse? Suppose that the child wants to ride a horse and must find a horse, but without past experience of the hobby horse. Thus, it is difficult to imagine how analogical transfer from knowledge of sticks as source could be matched with the horse as a target. An explanation would be that the child sees the stick and decides to use it as a horse. Here again, it is difficult to decide how knowledge about horses could match the perceived stick.

Case-based or analogy-based systems try to find on large amounts of data analogous cases, or analogous decision-making situations for decision-making or explanation [Roth-Berghofer, 2004]. With Gombrich, we think that analogy is not the right way for explaining conceptual changes that accompanies solution building. For example, it is difficult to understand how analogical thinking could explain the Archimedes prototype of stories about how to find a solution, as given below [Tijus & Brézillon, 2006].

Archimedes had a problem. King Hiero of Athens purchased a new crown and asked Archimedes to find out if the crown was really made out of pure gold, or if it was contaminated with cheap silver. Archimedes could not come up with a solution. After a long day of worrying, he decided to relax with a warm bath. When he entered the tub, he noticed the water level rising. This was something he knew, but now he suddenly realized the water displacement was proportional to the volume of the immersed part of his body. He found a way to determine the volume of the crown, and thereby discovered the solution to his problem!

According to Gentner and Toupin (1986), the finding of possible sources for analogy is as follows. First, the goal being defined (the crown and the problem of evaluating if the crown is only made of pure gold), a matching process starts by carrying out a large number of comparisons between the components (objects, objects attributes and relations) of the sources and of the goal. Second, source and structure are mapped for "global" identification. Ripoll and Eynard (2002) discussed the order of these successive phases (encode target, find sources with local matches, match structure) in human analogical processes. With the additional problem of how components are selected for mapping [Kwon *et al*, 2002], since analogy is based on similarity with already encoded data, it would be difficult to find a never used solution: Archimedes had never used water displacement for computing the volume of a possible source. Archimedes did not solve the problem with his body in the past as required by case-base reasoning or by analogy-based systems. Archimedes got the target problem, solved the source problem, and at the same time, was solving the target problem.

Archimedes’ cognitive processes are based on recognizing and identifying the crown like the same kind of things than his body from the viewpoint of the water-level rising. Thus, this is categorization problem. We could conjecture the goal of Archimedes could have been that “the crown and the same weight of pure gold should have the same volume. Thus how to find the volume of the chased crown without destroying the crown?” Archimedes could think that the problem would be simpler “if kings had a gold cube on their head instead of chased crown!” Why “cube of Gold”? “Cube of gold” is a counterfactual that is inferred while searching how the problem could be solved. This is realized by searching the objects that would be a solution of the problem, and then by computing differences: How the solution object differs from the actual object? How to reduce their difference? For example, Archimedes could have find that he cannot transform the crown in a cube, but it could mould it. Having the idea of something that envelops the object. Taking a bath, Archimedes could then notice that the water envelops his body and could have perceived the water displacement. If possible, such a thinking process cannot be derived from analogy.

3. Using a substitute instead of the solution

First, it is better to learn from errors than from success [Young *et al*., 2006] because errors help changing misconception. In an impasse, people, which are asked to solve a difficult problem, either make a lot of trials with errors, or stop and start thinking about why the goal cannot be reached and what could be a good position for reaching it, and the conditions to be satisfied to be in that position [Zanga *et al*., 2004]. Consider the chess game. For example, "If the Queen was here, I would win! Where is my Queen? Is it possible to have a series of moves leading the Queen there?" Why having these kinds of assumptions, hypotheses and counterfactuals? Our proposal is that solution building is based on the selection of candidate objects by counter-factuals which is a reasoning to enlarge thinking.

The substitute theory is the following. Consider the hobby horse of the previous example. The child wants to ride [a horse]. As there is no horse, s/he first considers riding the horse (fact) and as with no horse, s/he considers the counterfactual riding something that is not a horse, but that could be something to ride (in the environment). Then, among things in the environment to ride, the stick appears to be a substitute. It could be the reason that the child sees the stick and want to play with. If the stick is the stick of a brush, that is used to do the housework, s/he considers the counterfactual doing something else than the housework, but to play with. Then among plays, find the one that can use the stick.

Thus, the solution building would be both procedural and semantic. The solution is not analogy (i.e. from past: from source), but arises from creation of substitutes (from future: supposing to having solved impasses: riding). Impasses are solved by deleting prerequisites (specific conditions: having a horse to ride), by creating unconditioned substitutes (anything that can be used for riding), by contextual counterfactual categories (look for no horse), that can be generated by Generalized Galois Lattice (creating the no-horse category of things that can be used for riding), by selecting the substitute if any (a stick) that is implemented in Contextual Graphs, CxGs (ride the stick as a horse).

There is some experimental evidences of creating fakes (counterfactuals) from fakes (misconceptions). Along the logical viewpoint, if you don't get what you want in a category A (what you think to be facts but appears to be fakes), then looks in the non-A category (what is supposed to be fakes). Zamani and Richard (2000) show that impasses provide useful information about situation properties that are relevant to the problem. Participants, which succeed in impasses, do not ask themselves how to leave the impasse (fact), but how would be the situation if it was not an impasse (fake). Counterfactuals could also be triggered when, for example, one asks someone to solving the problem in Figure 1. What would s/he think about calculating the black surface (fact) as the difference of two rectangles, namely the whole surface and the white surface (fake). Such a hint helps to solve the initial problem (fake to fact).

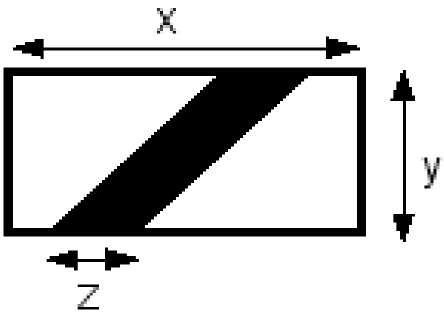


Figure 1. *The problem of finding the surface of the black part of the rectangle.*

In Figure 2, participants solve easily the passalong test of the top (top part of Figure 2), which consists of moving the squares from one position to another position by pushing them. However, many participants find more difficult to solve the second problem (bottom of Figure 2), because they consider the two left rectangles as two separate rectangles (fact) and move them separately, while if pushing them both as a whole the problem is the same than the top problem previously solved. Coloring the two rectangles with a same color, helps them conceptualizing them a large square (fake) and solving it (fake to fact) [Zamani & Richard, 2000].

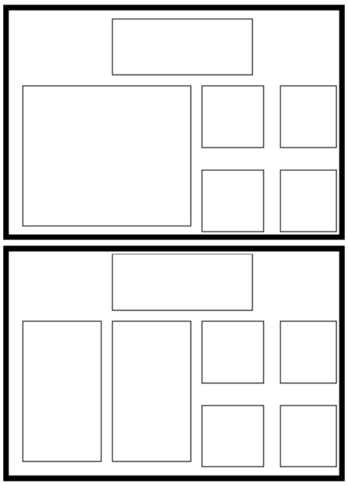


***Figure 2**. The Passalong test [Zamani & Richard, 2000]. Top problem and bottom problem are similar if the two left rectangles are perceived as being the large square of the top problem.*

4. Modelling rational route to solution

Contextual Categorization (CC) concerns the building of a category network to catch relationships between objects that are currently processed. For instance, the processing represented in Figure 3 is the building of a network of categories that comprise basic elements like, for instance, people in a list with attributes. In this example, the basic elements are individuals (O: I, J, K, L, …) with attributes about their location (P: Town: London, Paris; the countryside : Trifouilli) and their job (P: farmer, employee, …). The ProcOpe formalism (Poitrenaud et al., 2005) allows generalized Contextual Categorization model to operate on Galois Lattices to create a hierarchy of categories with transitivity, asymmetry and non reflexivity.

The O_n X P_m Boolean matrix indicates for each of the n objects O, if the object has the m properties P. In a lattice, whose complexity depends on the way properties are distributed over objects, the maximum number of categories is either 2^n-1, or m if m < 2^n-1. Let's suppose that (1) the task is to find people for loan and (2) from experience, one knows that employees in large town are good clients, as well as farmers in the countryside, although it is better to avoid employees in the countryside.

if (job: employee) & in big town (Paris, London),
if (job: farmer) in country (Trifoulli),
Then contact for loan
if not do not consider

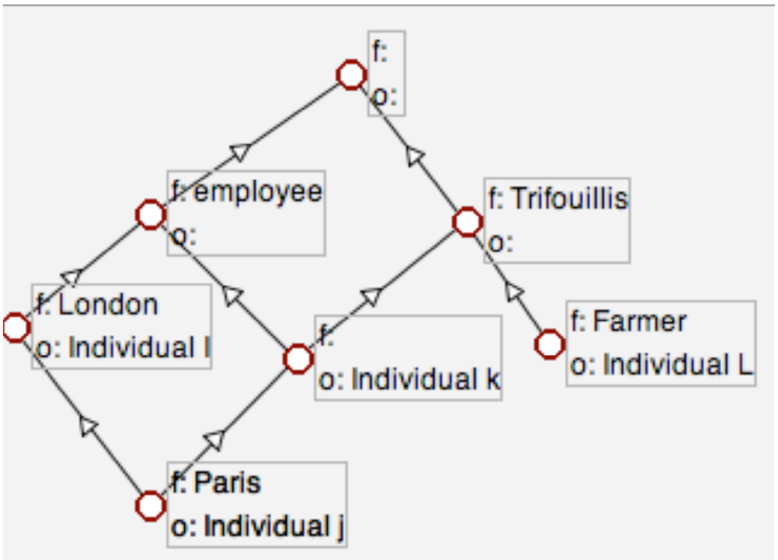
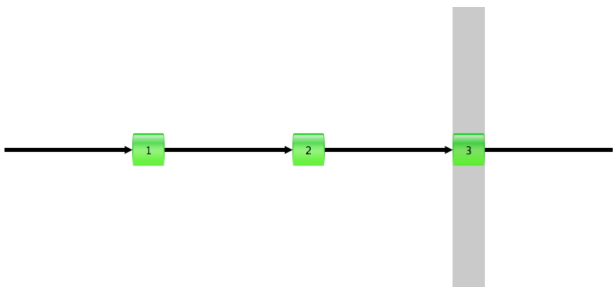
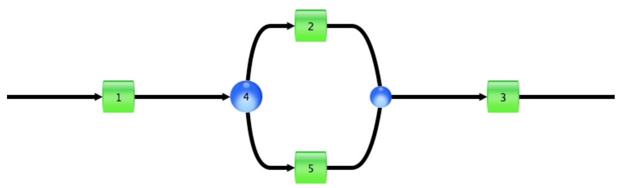
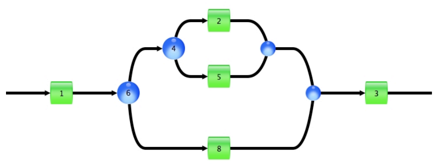


Figure 3*. The Galois lattice for categories of individuals (o) according to their features (f).*

According to the contextual categorization in Figure 3, individual i (employee in London), and individual l (farmer in Trifouilli) can be contacted for loan. An advantage of contextual categorization is to allow the rejection of individual k (employee in the countryside), while admitting individual j. Now, imagine that one wants to enlarge the client set. From facts (categories in the Galois lattice), a Generalized Galois Lattice (GCC) is realized through the building of complementary categories: ”not in big town”, “not farmer”, etc. But also “farmer not in the country” which is a counterfactual that helps for finding what can be a substitute category in large town of farmers: for instance a landlord that rent apartments, and so on. Then, one may apply the same procedure for this new kind of clients and test this procedure.







A1 : Take the water jug
C6 : Type of water?
running water
C4 : In a hurry?
no A2 : Pour water in jug
yes A5 : Pour hot water
mineral water
A8 : Take the bottle of mineral water
A3 : Fill the tank of the coffee machine

Figure 4. *Enrichment of know-how by integrating incrementally substitutes (here hot water) in a procedure formalized in the formalism of Contextual Graphs.*

Now, consider the following simple example represented in Figure 4. One generally uses cold water on the tap for coffee preparation (top-left of Figure 4). If we are late a morning, we are in a hurry and may wish to speed up some processes like the coffee preparation. A solution is to pour hot water in the jug because the principle of the coffee machine is to make water hot. Thus, we hope to speed up the process by finding this substitute of the cold water (top-right in Figure 4). This enrichment of know-how by integrating substitutes in procedure can be formalized in Contextual Graphs (CxGs) [Brézillon, 2005; 2007]. (In the bottom-left part of Figure 4 we have a new situation with a person that does not use tap water, but mineral water.) CxGs constitute a context-based representation of a task execution with all its variants. CxGs are oriented without circuits, with exactly one input and one output, and a general structure of spindle. A path (from the input to the output of the graph) represents a practice (or a procedure), a type of task execution with the application of selected methods. In Figure 4, the practice at the top-let is represented by the upper path of the CxG at the top-right of Figure 4. Practices are learned incrementally by the system. There are as many paths as practices and different solutions can be associated with the unique output [Brézillon, 2005]. We use CC and CxGs to implement new solutions: build counterfactuals as complementary categories, find substitutes from counterfactuals, test the implementation in CxGs.

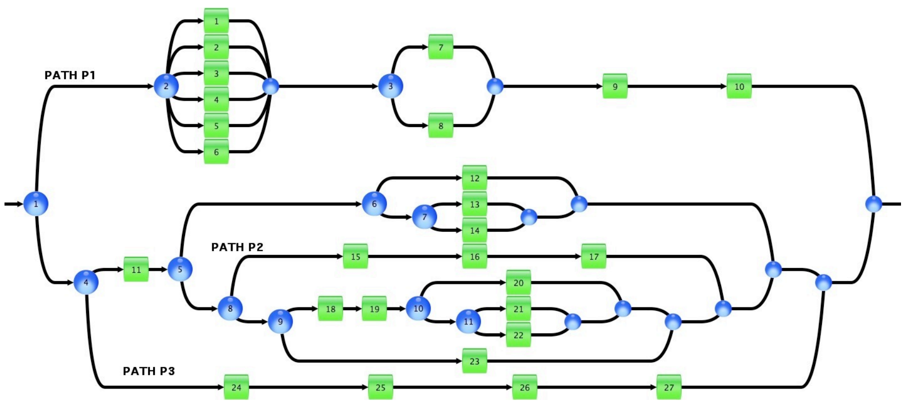
5. Modelling the search process

For considering the relevance of contextual graphs for modelling behaviours of users searching for information, we mapped ocular fixations sequences, scan paths, and contextual graph built from the search behaviour on a Web site (Brézillon et al., 2008). Eyes-tracking was used to determine which areas of the homepage the user considered and if the sequence of ocular fixations (see Figure 5) corresponds to one of the 18 paths described in the CxG.

Each path contains contextual elements existing in the homepage such as links or different areas of the homepage (text, picture…). Recording the sequence of fixations revealed if participants look at these elements before clicking. If so, we could conclude that this path allow to model this user behaviour.

This study concerns the behaviour of participants having to realize the following task: "*You wish to visit Arthur's Museum, What is the price of a single ticket*?" The name "*Arthur's Museum*" does not appear on the home page, and the answer was reachable from the home page in one click by two different links.

We used a twofold method. First, we proceeded in an exhaustive analysis of the Web site in order to determine all the possible paths to reach the answer starting from the home page. Figure 5 give the contextual graph of this task execution with all the possible paths including the path that corresponded to the prescribed behaviour. Second, we study a group of 9 participants by analyzing, in the one hand, their effective behaviours in a Contextual Graph, and, in the other hand, by recording the movement of their eyes by a eye-tracking technique.



C1 : Knowing where to look for?
Yes C2 : Which item?
Time schedule and prices A1 : Click on Time schedule and prices
Hangar_20 A2 : Click on Hangar_20
Event location A3 : Click on event location
Teacher entrance A4 : Click on teacher entrance
Agenda A5 : Click on agenda
Time Schedule A6 : Click on Time Schedule
C3 : Type of move on the page?
scrolling bar A7 : Go down to find "Atelier d'Arthur"
on item A8 : Click on item "Atelier d'Arthur"

A9 : Click on "More about..."
A10 : Read the price for individuals
No C4 : Is there a site map?
No A11 : Look for item "Atelier d'Arthur"
C5 : Is an item "Atelier d'Arthur" visible?
Yes C6 : Do you see a price?
Yes A12 : Read price for an individual
No C7 : Is there a link?
Yes A13 : Click on item "Atelier d'Arthur"
No A14 : Revise hypothesis for another place
No C8 : Is there the word "Atelier" (workshop)?
Yes A15 : Click on the item "events"
A16 : Identify the impasse (Wine workshop"
A17 : Modify the hypothesis
No C9 : Is there the word "event"?
Yes A18 : Click on the item "events"
A19 : Identify the impasse ("wine workshop")
C10 : Do you see a price?
Yes A20 : Read price for an individual
No C11 : Is there a link?
Yes A21 : Click on "Atelier d'Arthur"
No A22 : Revise hypothesis on location
No A23 : Revise hypothesis on location
Yes A24 : Click on item "site map"
A25 : Look for "Atelier d'Arthur"
A26 : Click on item "Ateier d'Arthur"
A27 : Look for individual price

Figure 5*. The different paths for addressing the question "What is the price for an individual?"*

Participants	P1	P2	P3
U1	0,38	-0,09	-0,09
U2	0,20	0,05	-0,10
U3	-0,61	0,98	-0,08
U4	-0,20	0,55	0,10
U5	-0,37	0,91	-0,05
U6	0,33	-0,13	-0,04
U7	-0,76	0,61	0,61
U8	-0,22	0,88	-0,42
U9	-0,22	0,76	-0,34

Table 1*. Correlation between three main paths (P1, P2, P3) in the contextual graph in Figure 5 and participants (Ui) eyes-tracking while searching a website information.*

As the web site was not user friendly, the information was difficult to find. Thus the task was quite a problem solving. Paths P2 and P3 correspond to paths used by participants that do not know where to find the information after the contextual Node P1 "*Do you get any idea of where to find the information: no*". Number of participant paths that matched P2 and P3 paths was of 8 on 9.

Figure 6*. Visual scan path of a participant searching for the information on the Web page.*

When there is a stop in eye movement. There is then a click on the fixed zone that corresponds to one step in the participant procedure. The participant procedure is then matched with paths in CxGs. Fixations on zones to be clicked not followed by a click means that the participant did not recognize it as a proper link. The participant procedure is then compared with paths in the CxG (see Figure 5). Fixations on zones to be clicked not followed by a clic means that the participant did not recognize it as a proper link.

As a result, path P3 in the contextual graph contains counterintuitive links. Participants supposed that the information "ticket price for the Arthur's Museum" was to be found in general information about tickets, or in specific information about a permanent museum activity. This was asserted by clicks on general information links or on links denoting localization (e.g. Hangar N° 20). As they do not find the information, among many possible links, participants choose to select "Events", that finally belong to the C2 path, and was a correct link to activate. We think that behaviour to be a counterfactual: "*if the information was not about a permanent activity*".

6. The contextual theory of substitute

In the constraint model developed by Richard [Richard, Poitrenaud & Tijus, 1993], most of the difficulties in problem solving emerge from misunderstanding the rules as expressed in the instructions and from the way in which the participant constructs the goal. The constraint model describes misunderstanding using a list of constraints in the form of "DO NOT (Action) ON (Object)". This list of constraints filters possible moves and allows predicting the moves a participant is going to perform. Feedback and above all, states for which the filter blocks all moves (impasses), are opportunities for deleting the implicit constraints. In this way, throughout the problem solving process, the participant continually adjusts his or her list of constraints and this entails adjusting his or her representation of the problem.

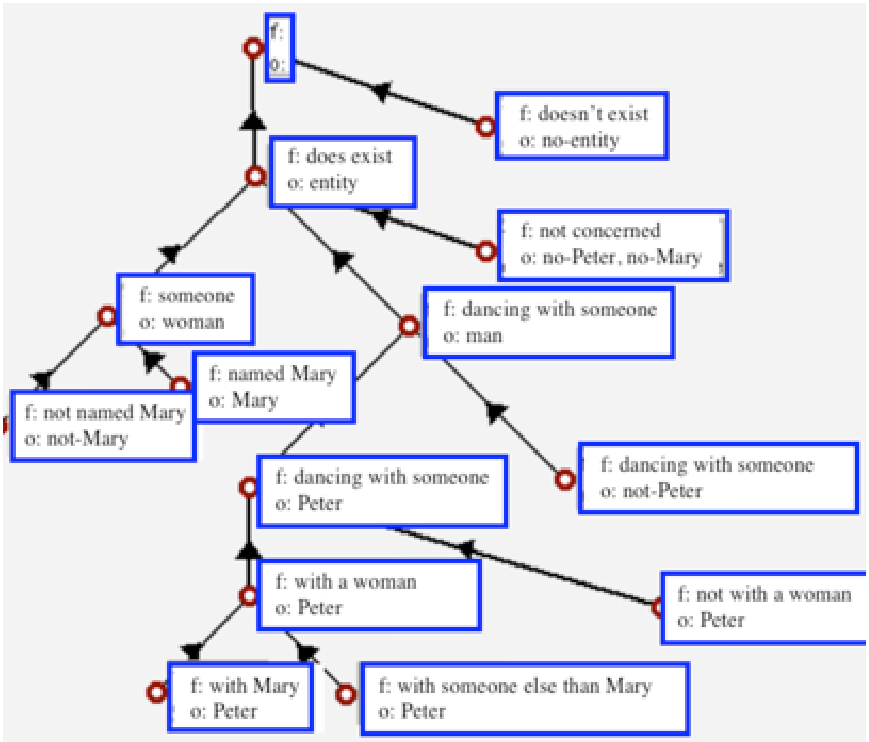


***Figure** 7. The network of contextual and of counterfactual categories built from "Peter and Mary are dancing."*

Deleting constraints that are responsible for errors, enlarge rationality and the field of thinking. However, this process does not inform which action is going to be selected. The level of decision-making corresponds to the level of the deleted constraint: its counterfactual category. For example, consider people that wear a black or a white hat and people that wear no hat. If you don't find, although present, an individual that wear a black hat, the search must be enlarged with people having a white hat, not to all the people.

Generalized contextual categorization (GCC) provides the set of counterfactual categories to be searched at the level of the category to be processed. Let's consider the following example. Starting with a fact: in a room, two persons are dancing and a question « *Is Peter dancing with Mary* ?». How many responses can you get? GCC considers all the relevant categories for understanding the sentence (people, people dancing, man, woman, Peter, Mary, and so on, with the counterfactual categories. This provides the following GCC network of hierarchical categories (Figure 7).

All the possible responses to the question "Is Peter dancing with Mary?" (e.g. Don't know these individuals, Peter is dancing with someone else, etc.,) can be derived by this GCC Process.

7. Conclusion

The role of counterfactuals in thinking has already been recognized [Fauconnier, 1985]. We propose an operational process of building solutions through the building of counterfactuals categories [Fakes] derived from actual categories [perceived as facts] to find substitutes that are integrated in known procedures. This whole process provides routes for solution, learning and finally clarity and more rational decision making.

8. References

[Brézillon, 2003] Brézillon, P. Representation of procedures and practices in contextual graphs. *The Knowledge Engineering Review, 18(2)*, (2003) 147-174

[Brézillon, 2005] Brezillon, P.: "Task-realization models in Contextual Graphs." In: Modeling and Using Context (CONTEXT-05), A. Dey, B. Kokinov, D. Leake, R. Turner (Eds.), Springer Verlag, LNCS 3554, pp. 55-68 (2005).

[Brézillon, 2007] Brézillon, P. (2007) Context Modeling: Task model and model of practices. In: Kokinov et al. (Eds.): Modeling and Using Context (CONTEXT-07), LNAI 4635, Springer Verlag, pp. 122-135.

[Brézillon and Pomerol, 1999] Brézillon, P. and Pomerol, J.-Ch. Contextual knowledge sharing and cooperation in intelligent assistant systems. *Le Travail Humain, 62(3), Paris: PUF*, (1999) pp 223-246.

[Brézillon and Tijus, 2006]. Brézillon, P. and Tijus, C. Rationality from fakes: a contextual approach. *Proceedings of the 9th European Meeting on Cybernetics and Systems Research*. Vienna, Austria, 2008 (To appear)

[Brézillon et al., 2008]. Brézillon, P., Tijus, C. and Léger, L. Increasing rationality of cognitive tasks with contextual graphs. *Proceedings of the 8th European Meeting on Cybernetics and Systems Research, Volume 2,* (pp. 359-364). Vienna, Austria, April 18-21, 2006 (Best Paper Awards)

[Fauconnier, 1985] Fauconnier, G. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA: MIT Press. (1985).

[Gentner and Toupin, 1986] Gentner, D., and Toupin, C.. Systematicity and surface similarity in the development of analogy. Cognitive Science, 10(3), 277-300. (1986).

[Gombrich, 1963] Gombrich, E. H. Mediations on a hobby horse. In: *Meditations on a Hobby Horse and Other Essays on the Theory of Art*. London (1963). (originally published in Aspects of Form, ed. L.L. Whyte, London, 1951).

[Kwon et al., 2002] Kwon, H., Im, I., and Van de Walle, B. Are you thinking what I am thinking – A comparison of decision makers' cognitive map by means of a new similarity measure. *Proceedings of the 35th Hawaii International Conference on System Sciences: 78*, 2002.

[Newell and Simon, 1972] Newell, A., and Simon, H.A. *Human Problem Solving*. Englewood Cliffs : Prentice-Hall. (1972).

[Richard et al., 1993] Richard, J.F., Poitrenaud, S and Tijus, C. Problem-solving restructuration: elimination of implicit constraints. *Cognitive Science, 17,* 497-529. (1993).

[Ripoll and Eynard, 2002] Ripoll, T., and Eynard, J. (2002). A Critical Analysis of Current Models of Analogy. *Proceedings of the 2002 Information Processing And Management of Uncertainty In Knowledge-Based Systems, IPMU* 2002.

[Roth-Berghofer, 2004] Roth-Berghofer, T. Explanations and Case-Based Reasoning. *Foundational Issues, Advances in Case-Based Reasoning* (ECCBR-2004)

[Tijus and Brezillon, 2006] Tijus, C. & Brezillon, P. Problem solving and Creativity for Decision Support Systems. In F. Adam, P. Brezillon, S. Carlsson, and P. Humphreys (eds), *Creativity and Innovation in Decision Making and Decision Support, Vol 1.*, (pp. 277-293), Ludic Publishing Ltd (2006)

[Tijus, 2003] Tijus, C. Métaphore et Analogie, les deux faces d'une seule pièce ? in C. Tijus (Ed.) *Métaphore et Analogie. Traité des Sciences Cognitives*. Paris: Hermes. (2003).

[Young et al., 2006] Young, W., Hesketh, B., and Neal, A. Using "War Stories" to Train for Adaptive Performance: Is it Better to Learn from Error or Success? *Applied Psychology, 55*, 282-302. (2006).

[Zamani and Richard, 2000] Zamani, M., and Richard, J.F. Object Goal Similarity and Analogical Transfer. *Memory and Cognition, 28*, 873-886. (2000).

[Zanga et al., 2004] Zanga, A., Richard, J.F., and Tijus, C. Implicit Learning in rule induction and problem solving. *Thinking & Reasoning, 10,* 55-83. (2004).

Centering Information Retrieval to the User

Bernd Ludwig* — **Stefan Mandl***

* *Chair for Artificial Intelligence, University of Erlangen-Nürnberg*
Am Weichselgarten 9, 91058 Erlangen (Germany)
{Stefan.Mandl,Bernd.Ludwig}@informatik.uni-erlangen.de

RÉSUMÉ.

ABSTRACT. *In this paper, we present a novel approach to text mining that helps to build intelligent user interfaces for recommender and information retrieval systems. The main problem for the user in information retrieval is that he must have almost perfect knowledge of the domain and the domain terminology. Our approach eases this burden by showing a way how to encode domain knowledge so that an information retrieval system can transform the user's way to talk about the domain in the expert's way to do that. After that transformation the system can search its data bases for appropriate information. We demonstrate the practicability of our approach in a case study on a TV recommender system.*

MOTS-CLÉS :

KEYWORDS:

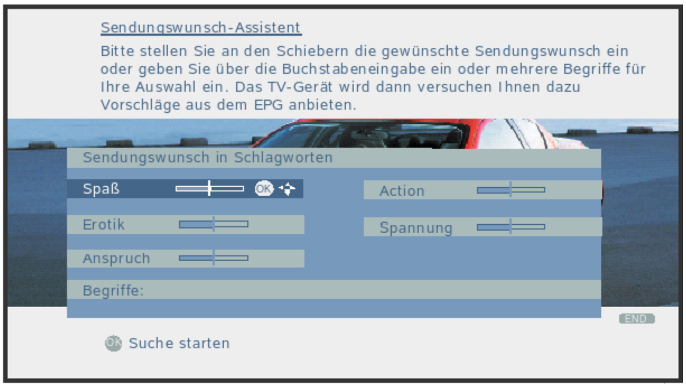


Figure 1. *A screen shot of the user interface implemented on our prototype TV set*

1. The User Perspective in Information Retrieval

Intelligent information retrieval is concerned with two orthogonal aspects :

– search *algorithms* that perform well for a particular retrieval task

– search *criteria* that are well suited to describe precisely a user's query and – in parallel – allow to compute good features for the search algorithm to compute optimal (with respect to a evaluation function) results.

Often, research is concentrated on the algorithms, while criteria are considered to be easy to find. For example, in (Satzger *et al.*, 2006) algorithms are presented for a preference based search in data bases, but the discussion does not focus the question how preferences for a particular user are entered, learned, or computed.

1.1. *Information Retrieval from Text Corpora*

However, if one takes a consultation system as an example that should advise the user how to layout a document with a picture in the upper right corner, and two columns of text floating around the picture and a footer and a header, the usual experience with any online help system is that the user has to know himself how a particular text processing system could be used for building complex layouts. In particular, he has to know the terminology of the system developer in order to enter the correct keywords for the explanations he is looking for.

The problem of unexperienced users not knowing the expert vocabulary is crucial for all user interfaces. If the user could formulate his queries in a "non-expert" manner this would require the user interface to be able to perform complex reasoning tasks. In this way it could infer the steps necessary to solve the user's problem. On the other hand, a user interface without these reasoning capabilities requires the user to have expert knowledge about the domain.

An important part of the domain knowledge consists in the vocabulary and language used to talk about domain relevant issues. The same observation is true for information retrieval systems in which there is a user centered point of view of talking about the domain and an expert centered point of view.

In concrete applications, there are many variations for the meaning of "user centered" and "expert centered". They depend mainly on the domain knowledge required to solve tasks in the domain.

1.2. *Analysis of the Design*

In order to set up a case study for an information retrieval system that allows the user to use his/her own vocabulary as far as possible, we analyzed the context of a TV recommender application, the requirements for its implementation, and the design of the information retrieval approach that helps to implement an easy-to-use recommender.

– **Context of Use**

In a Wizard-of-Oz study we collected data on how users select a TV programme when they can interact in natural language. Test persons were not obliged to any restrictions such as certain TV channels. There were no limitations in terms of language for the interaction with the wizard. The test persons should feel comfortable in a living room and choose a programme to watch.

– **Requirements**

The studies revealed the fact that test persons always "navigated" between several options they liked best. They asked questions on these options to get detailed information for their final decision. Furthermore, the study revealed that test persons talked about some distinct classes of mood and topics in order to filter out nice programmes. It became clear quickly that titles, genres, and broadcasting times played a minor role.

– **Design**

The analysis of the study resulted in a design that allows users to enter key words for a content based-search and to set sliders for the most important moods as they are common also in TV programme magazines.

1.3. *Other Approaches for TV Recommenders*

TV recommenders have already been built without analyzing textual descriptions about programmes. In (Blanco Fernández *et al.*, 2005; Blanco-Fernandez *et al.*, 2004), ontologies are applied to structure meta-information about programme. For a particular recommendation, it solves logical queries about the meta-information. However, the kind of meta-information used here tells much less about the programmes content than a natural language summary. In (Pigeau *et al.*, 2003), fuzzy reasoning is applied again on meta data.L Approaches incorporating reasoning about user preferences and collaborative filtering of such preferences also apply meta data.

Therefore, the contribution of our approach is the analysis of natural language summaries of TV programmes. The features extracted in this step can be used directly for classification or for getting additional meta data.

2. Structure of the Paper

The paper is focused on the algorithmic consequences of the design presented above. In the next section we discuss the concept of a switch from a expert centered perspective on the available information to a user centered one. We illustrate the concept in a case study on a TV recommender system. Next, we present a mathematical model of how the discussed switch can be made effective for retrieval or recommender systems relying on text mining approaches for data search. After that, we illustrate how the model presented up to now has been implemented in our TV recommender system. Finally, we present a user evaluation for the implemented system.

3. Case Study : TV Recommender Systems and the Real User

Via satellite, the Electronic Programme Guide (EPG) provides an enormous amount of information about TV programmes with natural language descriptions about the content of programmes. Viewers are overwhelmed by the huge number of channels and programmes when they select a programme to watch.

For the design and implementation of TV recommendation systems sophisticated user models such as (Ardissono *et al.*, 2004) are used. In order to allow for default reasoning, stereotypes for users are applied which are based on the analysis of the average user's lifestyle (see (Gena, 2001)). Much attention is paid on the issue on how to design an attractive, functional, and easy-to-use graphical interface between users and the recommendation system (see (van Barneveld *et al.*, 2004)). In order to increase the user's confidence in the system proposals, the generation of trust-worthy suggestions that take programmes watched earlier into account has been studied in detail (see (Buczak *et al.*, 2002)). For the implementation of the search, different approaches and theories of reasoning have been applied : (statistical) classifiers such as neural nets (Zhang *et al.*, 2002), fuzzy logic (Yager, 2003), similarity based reasoning (McSherry, 2002) – to name the most prominent ones. We use a text mining based approach as the necessary computations can be carried out quickly. Performance is an important factor as the system runs in real-time on an embedded platform used in a standard commercial TV set (see Fig. 1 for an example of how the GUI looks like).

In a user study (Nitschke *et al.*, 2003) conducted as part of the research project EMBASSI (see (Herfet *et al.*, 2001)), candidates were situated in front of a computer display that suggested an automatic recommendation system to be at work. The users were allowed to ask arbitrary questions about available TV programmes. A Wizard-of-Oz provided responses. The experiments showed that users express emotional attitudes they desire the programme to have, or even their own emotions hoping the system would come up with proposals that match their mood : *Liebe, Romantik (Love, romance) ; Entspannen (relax) ; Show, Witz (Show, fun)* In the domain of recommending TV programmes domain knowledge consists in know-how about the user's preferences, his current interests, the programmes available right now and an algorithm that computes matches between the programmes and the user's interests and preferences.

In the application of our case study, all knowledge is represented in terms of natural language descriptions. There is no classification system at hand which the system could use to restrict the search on formally given filter criteria. On the other hand, the user's interest is represented in natural language as well or in the setting of the four sliders that are used to characterize the mood of programmes the user is interested in.

4. User Centered versus Expert Centered Information

How to distinguish "user centered" and "expert centered" knowledge ? User centered knowledge can be represented as a four dimensional *ranking vector*

$$v = (\text{ranking for action} \quad \text{ranking for chill} \quad \text{ranking for fun} \quad \text{ranking for erotic})^T$$

as known from almost all TV magazines.

Expert knowledge entailed in the programme descriptions is hard to analyze completely by automatic means as the natural language understanding problem is not solved in general. However, text mining algorithms are a well-known technique for such tasks. In order to answer the question whether a text mining approach could be successful we tried to cluster TV programmes for German using a large sample of descriptions from the EPG data stream and the WEKA (see (Witten *et al.*, 2005)) implementation of the k-means clustering algorithm. The experiments showed that indeed the clustering could separate programmes that were also considered to be different by a human labeler :

Cluster	**Topics**
1	news, entertainment, movies
2	regional news
3	magazines

Consequently, this study leads to the hypothesis that the descriptions could be used as input for a text mining algorithm. So they would constitute the "expert centered" knowledge about the domain.

```
finden: 11.18 20.7
kai: 1.17 8.14 15.1
nehmen: 15.16 15.61 16.9 20.8
offenbar: 5.6 7.1 12.3 12.33
offenbaren: 12.3 12.5 12.48 14.22
opfer: 2.41 5.42 9.76 10.13 15.14 15.53 15.58 20.14 21.27 22.12 22.15
opfern: 2.41 5.42 10.14 20.14 21.26 22.12
schwierig: 9.38 9.53 9.53 10.18 10.55 12.4 12.35
suchen: 9.14 9.21 10.35 13.2
tot: 1.22 2.22 2.40 3.4 7.32 9.19 9.39 10.25
wald: 2.1
```

The matrix M for the first two words :

	finden	kai		finden	kai
1.17	$m_{0,0}$	$m_{0,1}$	15.1	$m_{3,0}$	$m_{3,1}$
8.14	$m_{1,0}$	$m_{1,1}$	20.7	$m_{4,0}$	$m_{4,1}$
11.18	$m_{2,0}$	$m_{2,1}$			

Figure 2. *Example for the mapping of words to topics*

5. The Need for Switching the Perspective

5.1. *Standard Text Mining*

However, standard text mining requires user centered knowledge and expert centered knowledge to be almost identical as the algorithms count the frequency of domain relevant words and build feature vectors the have a new dimension for each new domain relevant word. Matches are then computed by finding a vector (of a programme description) enclosing a small angle to another vector (of a user query). A feature vector for text mining counts the frequency in which a word occurs in a text (or computes some function on the frequency). So, if there a n different words in a lexicon for text mining, a feature vector looks like this :

$$f = (w_1\, w_2 \ldots w_i \ldots w_n)^T$$

w_i is the frequency of the i$^{\text{th}}$ word in the lexicon (or the result of a rating function applied on the frequency). If one has two such vectors, one can compute their scalar product (i.e angle between the vectors) as a measure of their similarity. However, in our case study, this approach does not work :

1) If the rankings are used as input, user queries and descriptions have different dimension and the value of a dimension in the query vector has a different meaning than that in the description vector.

2) If natural language is used as input, the vocabulary of user queries is so different from that in the descriptions that matches will be found rarely if at all.

5.2. *How to Switch the Perspective*

Therefore, the question is how to bring some sort of inference into the play, that gives good results in terms of precision and does not require computational resources beyond the limits of the embedded platform.

If one analyses the sample corpus of user queries it becomes obvious that users talk about topics and moods, not about concrete content. They ask for "something chilling and frightening" and not for a "movie in which a father kills his family one after another during the summer holiday". This observation leads to the conclusion that in our case study inference is *finding the topics (and emotions)* a movie is about and *finding the topics (and emotions)* the user talks about. After this generalization of concrete content which results in a list of topics, matches could eventually be computed comparing feature vectors on this higher level of abstraction.

5.3. *Case Study : Is Switching Practicable ?*

In order to validate this last hypothesis, we repeated the study whether programmes with different content could be separated by k-means clustering after assigning a list of topics to each description in the sample corpus (see Sect. "Finding Topics in Words"). The clusters were even better than in the first study :

Cluster	Topics	Cluster	Topics
1	travel, entertainment	4	infotainment
2	documentation, information	5	magazines
3	movies		

How to incorporate this approach in the mathematical model of text mining ? The mathematician's view is that a transformation from one vector space (that of different words in a text) to another one (that of different topics in a text) can be done by computing a matrix for the linear transformation.

6. Generalizing Text Mining by Abstraction

What should such a matrix look like ? With n different words and m different topics, we would like to have a matrix M of dimension $m \times n$ with :

$$(t_1 \dots t_i \dots t_m)^T = M \cdot (w_1 \dots w_i \dots w_n)^T$$

For an example of how the matrix M can look like, see Fig. 2. There are n different German base forms occurring in the description of a movie and m different topics encoded in the form `x.y` (The explanation for these codes will follow in Section "Finding Topics in Words" below). Each of these codes constitutes one of the m rows of M. What is the benefit of M ? If one multiplies M with a TF/IDF vector (Manning *et al.*, 1999) for the n different base forms, the result is a m-dimensional vector. Each dimension stands for one topic, and the value for each dimension is a weighted sum of the TF/IDF values. The weights are stored in the matrix M and relate words and topic on a quantitative basis : If we simply want to count the frequency of topics, after we have counted the frequency of words, M contains as elements :

$$m_{i,j} = \begin{cases} 1 & \text{word } j \text{ is in topic } i \\ 0 & \text{otherwise} \end{cases}$$

For the example in Fig. 2, the computation would be the following :

$$\begin{pmatrix} \text{tf-idf}(1.17) \\ \text{tf-idf}(8.14) \\ \text{tf-idf}(11.18) \\ \text{tf-idf}(15.1) \\ \text{tf-idf}(20.7) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \text{tf-idf}(finden) \\ \text{tf-idf}(kai) \end{pmatrix}$$

Now we are ready to compute the similarity s of two vectors v and w :

$$s = M \cdot \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix} \cdot M \cdot \begin{pmatrix} w_1 \\ ... \\ w_n \end{pmatrix} = \begin{pmatrix} t_1^v \\ ... \\ t_m^v \end{pmatrix} \cdot \begin{pmatrix} t_1^w \\ ... \\ t_m^w \end{pmatrix} = t^v \cdot t^w$$

t^v and t^w are the generalized feature vectors (topic vectors) for the both texts under consideration. This is the standard cosine similarity measure used in many classification tasks. The matrix M encodes the switch from user to expert knowledge : the entries of M relate words and topics statistically on the basis of a lexicon for German. For other applications, M could be changed and its entries can even be determined using machine learning techniques.

7. Iteration of the Abstraction Process

From the mathematical point of view t^v and t^w are again vectors and can be interpreted as feature vectors in an m dimensional vector space. Of course, one may apply another linear transformation to such a feature vector and repeat the abstraction process, i.e. generalize once again. What could this be good for ? Remember the case study above : in our recommender system, users can also change the values of four sliders for *action*, *chill*, *fun*, and *erotic*. Formally, this results in a ranking vector

$$r = \left(r_{\text{action}}\, r_{\text{chill}}\, r_{\text{fun}}\, r_{\text{erotic}}\right)^T$$

which has dimension 4. One can interpret each of these four dimensions as a cluster of topics (see Sect. “Drawing the Distinction ...”). Consequently, we can compute a feature vector in this space of moods from a feature vector in the space of topics by applying another matrix N of dimension $k \times m$. An element of the matrix is :

$$n_{i,j} = \begin{cases} r_j & \text{topic } j \text{ addresses mood } i \\ 0 & \text{otherwise} \end{cases}$$

Query		Topics		Recommendation	
Germ.	Engl.	ID	Description	Germ.	Engl.
Show	show	14.18	light music	Rock	rock
			Vinyl	vinyl	
			Pop	pop	
			Hit	pit	

Figure 3. *Relation between words and topics in a query and in a free text description*

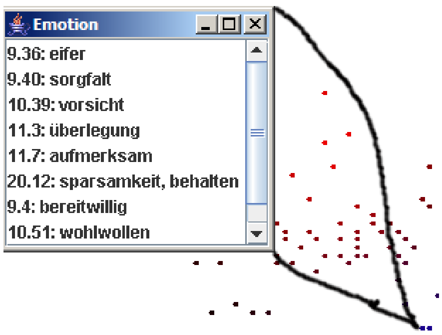


Figure 4. *From emotions to topics*

So, on this level of abstraction, the similarity s of two texts is :

$$s = N \cdot t^v \cdot N \cdot t^w = m^v \cdot m^w$$

As you can see, the same computations have to be carried out as for mapping from words to topics. Again, N can be constructed using any knowledge acquisition technique that computes a quantitative correlation between the dimensions of the input vector and those of the output vector.

8. The Implemented Recommender System

This section describes the system that has been implemented during our case study on generalized text mining. It also explains details on the knowledge sources used for the construction of the transformation matrices M and N described above.

8.1. *Finding Topics in Words*

The linguistic knowledge we exploit for finding the list of topics a German word is associates to is based on the DORNSEIFF lexicon for German. Like a thesaurus,

Engl. adj.	*German adj.*	*Dornseiffgroups*	*Activ*	*Eval*	*Angle*
Adventurous	Abenteuerlich	9.72,10.23,10.38	4.2	5.9	270.7
Affectionate	Herzlich,Liebevoll	4.50,10.51,15.33,10.49,10.52	4.7	5.4	52.3
Afraid	Besorgt,Bange	10.41,10.13,4.29,9.40	4.9	3.4	70.3
Aggressive	Aggressiv,Dynamisch, Energisch,Feindlich	10.30,10.55,18.32,9.35,5.35, 8.21,9.6,15.78,15.46	5.9	2.9	232.0
Agreeable	Angenehm,Annehmlich	9.54,10.10,10.52,12.13,11.46	4.3	5.2	5.0
Amazed	Erstaunt	10.29	5.9	5.5	152.0

Figure 5. *Some emotion words with V/A coordinates*

it groups words according to certain topics, i.e. in each group there are words (even of different word categories) that describe a particular aspect of a certain topic. The DORNSEIFF lexicon[1] is not a synonym lexicon, but a “topic” lexicon. Structured in a two-level hierarchy, the lexicon organizes topics in chapters (e.g. chapter 15 contains subtopics *social life*) and sub chapters (e.g. sub chapter 15.39 is the topic *reward*). If the meaning of a word is ambiguous, it is listed in more than one sub chapter. For the German words *finden (to find)* and *Kai (name of a boy or quay)*, the lexicon lists the following topics (cf. the example above) :

group id	**description**	**group id**	**description**
1.17	Ufer (the coast)	15.1	Vorname (first name)
8.14	Schiff (the ship)	20.7	Erwerb (the aquisition)
11.18	Entdeckung (the detection)		

For the construction of the matrix M this means :

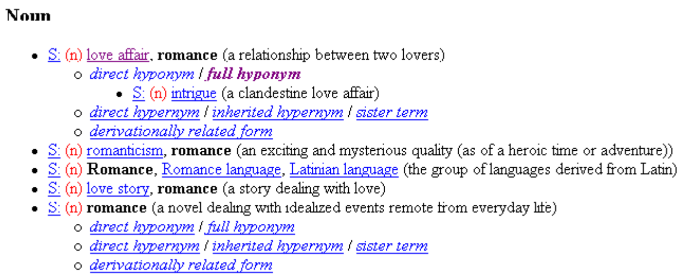
$$m_{1.17,\text{Kai}} = 1 \quad m_{8.14,\text{Kai}} = 1 \quad m_{11.18,\text{finden}} = 1 \quad m_{15.1,\text{Kai}} = 1$$
$$m_{20.7,\text{finden}} = 1$$

Fig. 3 illustrates how topics are used for finding good matches. In the example, the natural language user input can be generalized to a list of topics which is found in a number of words in a programme description. The higher the number of hits the better the programme matches the query.

8.2. *Detecting Emotions in Texts*

In order to keep the emotional aspect as a recommendation criterion, we use the valence/arousal model by WHISSEL (described in (Cowie *et al.*, 2001), pp. 39), which gives each emotion represented by an adjective a fixed position in a two dimensional coordinate system. Valence indicates emotion quality, hence whether it is considered

1. Interested readers can test the online version of the DORNSEIFF lexicon on `http ://wortschatz.uni-leipzig.de`.



Noun

- S: (n) love affair, **romance** (a relationship between two lovers)
    - *direct hyponym* / ***full hyponym***
        - S: (n) intrigue (a clandestine love affair)
    - *direct hypernym* / *inherited hypernym* / *sister term*
    - *derivationally related form*
- S: (n) romanticism, **romance** (an exciting and mysterious quality (as of a heroic time or adventure))
- S: (n) **Romance**, Romance language, Latinian language (the group of languages derived from Latin)
- S: (n) love story, **romance** (a story dealing with love)
- S: (n) **romance** (a novel dealing with idealized events remote from everyday life)
    - *direct hyponym* / *full hyponym*
    - *direct hypernym* / *inherited hypernym* / *sister term*
    - *derivationally related form*

Figure 6. *A* WORDNET *entry*

a positive or negative emotion. Arousal stands for the activation level an emotion possesses “i.e. the strength of the person’s disposition to take some action rather than none” ((Cowie *et al.*, 2001), p. 39). The table in Fig. 5 shows some English adjectives along with their German translation and their position in the two dimensional valence/arousal model (see columns *activ* and *eval*). Fig. 4 illustrates the results of WHISSEL’s approach graphically : all emotions are plotted in a two dimensional space. It also illustrates that a list of topics is associated to each emotion identified by WHISSEL. This relation between topics and emotions is computed in the following way : a topic belongs to an emotion if in the Dornseiff lexicon, a word out of the word family for the translation of an English emotional adjective is in this topic.

Furthermore, the valence/arousal model allows to localize the moods *fun*, *sadness*, *chill*, or *action* (see the four sliders in Fig. 1) in certain regions of the valence/arousal space. This (indirect) relationship between moods and vocabulary leads to the construction of the elements $n_{i,j}$ of matrix N and can be exploited for finding recommendations : if the user wants to see something chilling (r_{chill} is high) then the topics assigned to the emotions in the region for chill are good features to search for (as they deliver a high contribution in the computation of the scalar product if the appropriate values w_i are high). In our recommender system, the user can define these clusters during configuration. The next version of the system will learn these clusters indirectly from user feed back.

8.3. *Computing the Distance between Topics*

For the clarity of the presentation, one point has not been mentioned so far. In the implemented recommender system, the matrix M actually is not binary. The values for $m_{i,j}$ are computed applying another heuristic : in the field of word sense disambiguation, several metrics (see (Pedersen *et al.*, 2005)) are known to measure the semantic distance between two words. Some of these metrics use the WordNet dictionary (see Fig. 6 for an example entry) for the distance $\text{dist}(s,t)$ between the words s and t. To compute the distance between two words, one can count the number of hyperonym/hyponym steps necessary to move from s to t in the WordNet hierarchy. Fig. 6) shows the first steps starting with the word *love*. The steps are indicated by the links *direct hyperonym* and *direct hyponym* respectively.

We applied such a metric to compute the distance between each pair of topics in the Dornseiff lexicon. This results in a quadratic matrix S of dimension $m \times m$:

$$(s_{i,j})_{1 \leq i,j \leq m} = \text{dist}(i, j)$$

where $\text{dist}_{i,j}$ is computed as the distance between the topic names s and t for the pair of Dornseiff topics (s, t). The matrix S allows to rank those topics high that occur in the same text and and close to each other in terms of their semantic distance in WordNet. So, our recommender system uses a modified version of the matrix M which is calculated as $\tilde{M} = S \cdot M$.

8.4. *Computing Recommendations*

In order to compute proposals matching a user query, the recommender system computes the similarity value

$$s_{q,k} = (N \cdot \tilde{M} \cdot q) \cdot (N \cdot \tilde{M} \cdot k)$$

for a given query q and each programme description k which is broadcast in the time interval selected by the user. The description are ranked in descending order of $s_{q,k}$ as high values for the scalar product means that q and k are close together in the vector space, and a small angle between them indicates a good match.

9. Further Work : User Adaptivity

In a real application, the approach presented here depends heavily on the quality of the matrices described above. Individual values in the matrix N cells tell the recommender how important each emotion is for each of the moods *action*, *chill*, *fun*, and *erotic*. A value in matrix M tells which topics are relevant for an emotion.

Obviously, these values may vary from user to user. Therefore, to be user-adaptive the system must learn the values in the matrix cells from user feedback. Therefore, we are currently extending the user interface. It allows the user to rate a chosen programme. If the feedback is positive, the association between words, topics, and emotions is reinforced. In this way, the matrices M and N are adapted incrementally to the user's personal preferences.

10. Evaluation

As our developed system is considered to offer more comfort and flexibility to the users, the most important figure to evaluate is how good do proposals match the user query in the user's view.

answer	share	answer	share
very useful	36%	useless	7%
useful	26%	very useless	1%
fair	20%	no idea	10%

(a) Answers to Question 1

answer	share	answer	share
very good	20%	bad	10%
good	28%	very bad	5%
fair	22%	no idea	15%

(b) Answers to Question 2

answer	share	answer	share
yes	84%	do not know	7%
no	2%	no idea	7%

(c) Answers to Question 3

Figure 7. *Results of the User Evaluation*

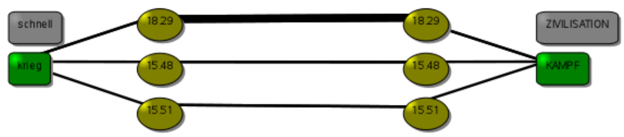


Figure 8. *Explanation of a Recommendation*

For a first evaluation, in a public presentation of the demonstrator system people of different sex, age, education and interest could test the system as long as they wanted to. They used the NL interface by typing in queries or playing with the sliders. People tested the system for about 15 minutes and entered around five queries in the average. The system responded by presenting a list of the programmes on air at the time of the query sorted according to how well each programme matched the query. Then the users filled in a questionnaire. 60 questionnaires have been evaluated. The first question was : *How useful do you consider the text input-and-search function of the system ?* It was answered as shown in Fig. 7(a). Test persons used the standard remote control of the TV set with one special key which invokes the GUI dialog in Fig. 1. In this way, they had nothing new to learn for the new feature of the TV set and could concentrate completely on the recommendations.

The second question was : *How appropriate do you consider the proposals you got from the system ?* The answers in this case are shown in Fi.g. 7(b). For deciding about the appropriateness of proposals, the candidates could read the description along with a graphical explanation (see Fig. 8 for an example) of what the system had computed.

The explanation shows all base forms in the EPG summary of a programme. It displays the words used for the recommendation in green, links them to the associated topics – for the user's query (left side) as well as for the summary (right side). The thickness of the lines between the topics on the left and on the right indicate the (relative) importance of each topic. The explanation and the full text supported the users in valuating the quality of the system's suggestions.

The last question *Would you recommend the system to a friend ?* (answers in Fig. 7(c)) demonstrates that users were content with the system.

11. Conclusions

In this paper, we presented a mathematical formulation of a generalized approach to text mining. "generalized" means that not only words serve as features, but that semantic abstractions can be made that allow to cluster the vocabulary into several categories that encode domain relevant knowledge.

We argued that this approach is suited to build recommender systems that analyze natural language documents as their data basis for computing proposals. From the user's point of view the main advantage of our approach is that the described process of abstraction can be used to switch from expert centered vocabulary to user centered vocabulary. This enables the user to get good results for his queries even without knowing the expert terminology of an application domain.

In a case study of a TV recommender system, we show that our approach is practicable and accepted by the users who took part in an evaluation of our system. From the computational point of view, the implemented approach is fast enough to perform in real-time even on embedded systems with low memory and CPU performance.

12. Bibliographie

Ardissono L., Gena C., Torasso P., Bellifemmine F., Difino A., Negro B., « User Modelling and Recommendation Techniques for Personalized Electronic Program Guides », *in* L. Ardissono, A. Kobsa, M. T. Maybury (eds), *Personalized Digital Television – Targeting Programs to Individual Viewers*, vol. 6 of *Human-Computer Interaction Series*, Springer, chapter 1, p. 3-26, 2004.

Blanco-Fernandez Y., Pazos-Arias J. J., Gil-Solla A., Ramos-Cabrer M., Barragans-Martinez B., Lopez-Nores M., « A Multi-Agent Open Architecture for a TV Recommender System : A Case Study Using a Bayesian Strategy », *ISMSE '04 : Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering*, IEEE Computer Society, Washington, DC, USA, p. 178-185, 2004.

Blanco Fernández Y., Pazos Arias J. J., Gil Solla A., Ramos Cabrer M., López Nores M., Barragáns Martínez A. B., « AVATAR : A Multi-Agent TV Recommender System Using MHP Applications », *IEEE International Conference on E-Technology, E-Commerce and E-Service (EEE)*, IEEE Computer Society Press, p. 660-665, mar, 2005.

Buczak A. L., Zimmerman J., Kurapati K., « Personalization : Improving Ease-of-Use, Trust, and Accuracy of a TV Show Recommender », *Proceedings of the TV'02 workshop on Personalization in TV*, Malaga (Spain), 2002.

Cowie R., Douglas-Cowie E., Tsapatsoulis N., Votsis G., Kollias S., Fellenz W., Taylor J., « Emotion Recognition in Human-Computer Interaction », *IEEE Signal Processing Magazine*, vol. 18, p. 32 - 80, January, 2001.

Gena C., « Designing TV Viewer Stereotypes for an Electronic Program Guide », *Proc. UM2001 Workshop on Personalization in Future TV (TV01)*, Sonthofen, July, 2001.

Herfet T., Kirste T., Schnaider M., « EMBASSI – Multimodal Assistance for Infotainment and Service Infrastructures », *Computers and Graphics*, vol. 25, n° 4, p. 581-592, August, 2001.

Manning C. D., Schuetze H., *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.

McSherry D., « A Generalised Approach to Similarity-Based Retrieval in Recommender Systems », *Artif. Intell. Rev.*, vol. 18, n° 3-4, p. 309-341, 2002.

Nitschke J., Hellenschmidt M., « Design and Evaluation of Adaptive Assistance for the Selection of Movies », *Proceedings of IMC 2003 "Assistance, Mobility, Applications"*, Rostock, June, 2003.

Pedersen T., Patwardhan S., Michelizzi J., « Wordnet : :Similarity - measuring the relatedness of concepts », *Proceedings of the Nineteenth National Conference on Artificial Intelligence*, p. 1024-1025, 2005.

Pigeau A., Raschia G., Gelgon M., Mouaddib N., R., Saint-Paul, « A fuzzy Linguistic Summarization Technique for TV Recommender Systems », *The 12th IEEE International Conference on Fuzzy Systems*, p. 743-748, 2003.

Satzger B., Endres M., Kießling W., « A Preference-Based Recommender System », *7th International Conference on Electronic Commerce and Web Technologies (EC-Web '06/Dexa 2006)*, Lecture Notes in Computer Science, Cracow (Poland), September, 2006. ISSN : 0302-9743.

van Barneveld J., van Setten M., « Designing Usable Interfaces for TV Recommender Systems », *in* L. Ardissono, A. Kobsa, M. T. Maybury (eds), *Personalized Digital Television – Targeting Programs to Individual Viewers*, vol. 6 of *Human-Computer Interaction Series*, Springer, chapter 1, 2004.

Witten I. H., Frank E., *Data Mining : Practical machine learning tools and techniques*, 2 edn, Morgan Kaufmann, San Francisco, 2005.

Yager R. R., « Fuzzy logic methods in recommender systems », *Fuzzy Sets Syst.*, vol. 136, n° 2, p. 133-149, 2003.

Zhang T., Iyengar V. S., « Recommender Systems Using Linear Classifiers », *Journal of Machine Learning Research*p. 313-334, 2002.

“Methodological Aspects of HCP” session

02 – 03:30 pm

Thursday, June 12, 2008

Analysis of monotonicity properties of new normalized rule interestingness measures

Roman Słowiński *[a] — Salvatore Greco — Izabela Szczęch ***

** Institute of Computing Science, Poznan University of Technology,*
Piotrowo 2, 60–965 Poznań, Poland

Roman.Slowinski@cs.put.poznan.pl

Izabela.Szczech@cs.put.poznan.pl

[a] Systems Research Institute, Polish Academy of Sciences,
01–447 Warsaw, Poland

*** Faculty of Economics University of Catania,*
Corso Italia, 55, 95129 Catania, Italy

salgreco@ unicit.it

ABSTRACT. *The paper considers interestingness measures for evaluation of relevance and usefulness of "if…, then…" rules induced from data. We propose a way to normalize three popular measures: rule interest function of Piatetsky-Shapiro, gain measure of Fukuda et al. and dependency factor used by Popper and Pawlak. The normalization transforms the measures to the interval [–1, 1], whose bounds correspond to maximal Bayesian confirmation and disconfirmation, respectively, and thus make them more meaningful. The new normalized measures are analyzed with respect to a valuable property M of monotonic dependency on the number of objects in the dataset satisfying or not the premise or the conclusion of the rule. The obtained results have a practical application as they lead to efficiency gains while searching for the best rules.*

KEY WORDS: *Data mining, Normalization, Rule Interest Function, Gain Measure, Dependency Factor, Monotonicity property M*

1. Introduction

Rules are popular patterns induced from data using various techniques of data mining. In this active research area, rule evaluation has been considered by many authors from different perspectives. To guide the data analyst identifying valuable rules, various quantitative measures of interestingness (attractiveness measures) have been proposed and studied (e.g. support, anti-support, Bayesian confirmation measures) (Hilderman *et al*., 2000). They all reflect some different characteristics of rules.

The problem of choosing an appropriate interestingness measure for a certain application is difficult because the number and variety of measures proposed in the literature is so wide. Therefore, there naturally arises a need to analyze theoretical properties of measures. Properties of measures group them unveiling relationships between them, and are helpful in choosing an appropriate measure for a particular application (Lenca *et al*., 2008). Some of the properties have also very practical applications.

In this paper, we focus on three commonly known and used measures: rule interest function (Piatetsky-Shapiro 1991), gain measure (Fukuda *et al*., 1996) and dependency factor advocated by (Popper 1959) and (Pawlak2004). We propose a way to normalize these measures to the interval [-1, 1], whose bounds correspond to maximal Bayesian confirmation and disconfirmation, respectively, and thus make them more meaningful. We analyze the new normalized measures with respect to a valuable property M, introduced by (Greco *et al*. 2004), of monotonic dependency of the measure on the number of objects satisfying or not the premise or the conclusion of the rule. Moreover, on the basis of satisfying the property M, we draw some practical conclusions about very particular relationship between these measures and two other simple but meaningful measures of rule support and anti-support.

The paper is organized as follows. In section 2, there are preliminaries on rules and their quantitative description. Next, in section 3, we analyze the normalized rule interest function, gain measure and dependency factor with respect to property M. Section 4 presents practical application of the obtained results. The paper ends with conclusions.

2. Preliminaries

Let us consider discovering rules from a sample of larger reality given in a form of a data table. Formally, a *data table* is a pair $S = (U, A)$, where U is a nonempty finite set of objects, called *universe*, and A is a nonempty finite set of *attributes*. For every attribute $a \in A$, let us denote by V_a the domain of a. By $a(x)$ we will denote the value of attribute $a \in A$ for an object $x \in U$. A *rule* induced from a data table S is denoted by $\phi \to \psi$ (read as "*if* ϕ, *then* ψ"), where ϕ and ψ are built up from elementary conditions using logical operator $\wedge$ (and). The *elementary conditions* of a rule are defined as ($a(x)$ *rel* v)) where *rel* is a relational operator from the set {=, <, ≤, ≥, >} and v is a constant belonging to V_a. The antecedent ϕ of a rule is also referred to as *premise* or *condition*. The consequent ψ of a rule is also called *conclusion*, *decision* or *hypothesis*. Therefore, a rule can be seen as a consequence relation (see critical discussion in (Greco *et al*., 2004) about interpretation of rules as logical implications) between premise and conclusion. The rules mined from data may be either *decision* or *association* rules, depending on whether the division of A into condition and decision attributes has been fixed or not.

2.1. *Support and Anti-support Measures of Rules*

One of the most popular measures used to identify frequently occurring association rules in sets of items from data table S is the support. The *support* of condition ϕ, denoted as $sup(\phi)$, is equal to the number of objects in U having property ϕ. The support of rule $\phi \rightarrow \psi$ (also simply referred to as support), denoted as $sup(\phi \rightarrow \psi)$, is equal to the number of objects in U having both property ϕ and ψ; for those objects, both premise ϕ and conclusion ψ evaluate to true.

Anti-support of a rule $\phi \rightarrow \psi$ (also simply referred to as anti-support), denoted as *anti-sup*$(\phi \rightarrow \psi)$, is equal to the number of objects in U having property ϕ but not having property ψ. Thus, anti–support is the number of counter-examples, i.e. objects for which the premise ϕ evaluates to true, but which miss the property ψ. Note that anti-support can also be regarded as $sup(\phi \rightarrow \neg\psi)$. Thus, it is considered as a cost-type criterion, which means that the smaller the value of anti-support, the more desirable the rule is.

2.2. *Piatetsky-Shapiro's Rule Interest Function, Gain and Dependency Factor*

The *rule interest* function RI introduced by (Piatetsky-Shapiro 1991) is used to quantify the correlation between premise and conclusion. It is defined by the following formula:

$$RI = sup(\phi \rightarrow \psi) - \frac{sup(\phi)\, sup(\psi)}{|U|}. \qquad [1]$$

For rule $\phi \rightarrow \psi$, when RI=0, then ϕ and ψ are statistically independent and thus, such a rule should be considered as uninteresting. When $RI > 0$ ($RI < 0$), then there is a positive (negative) correlation between ϕ and ψ (Hilderman *et al*., 2000).

The *gain* measure of (Fukuda *et al*., 1996) is defined in the following manner:

$$gain(\phi \rightarrow \psi) = sup(\phi \rightarrow \psi) - \Theta\, sup(\phi). \qquad [2]$$

where Θ is a fractional constant between 0 and 1. Note that, for a fixed value of Θ = *sup*(ψ)/|*U*|, the *gain* measure becomes identical to the above *RI*.

The *dependency factor* considered in (Pawlak 2004) and also advocated by (Popper 1959), is defined in the following manner:

$$\eta(\phi \to \psi) = \frac{\dfrac{sup(\phi \to \psi)}{sup(\phi)} - \dfrac{sup(\psi)}{|U|}}{\dfrac{sup(\phi \to \psi)}{sup(\phi)} + \dfrac{sup(\psi)}{|U|}}. \qquad [3]$$

The *dependency factor* expresses a degree of dependency, and can be seen as a counterpart of correlation coefficient used in statistics. When ϕ and ψ are independent on each other, then η(ϕ→ψ)=0. If −1< η(ϕ→ψ) <0, then ϕ and ψ are negatively dependent, and if 0< η(ϕ→ψ) <1, then ϕ and ψ are positively dependent.

2.3. *Normalization of interestingness measures*

Among widely studied and applied interestingness measures there is also a group of Bayesian confirmation measures which quantify the degree to which the premise provides "support for or against" the conclusion (Fitelson 2001). Thus, formally, a measure *c*(ϕ→ψ) can be regarded as Bayesian measure of confirmation if it satisfies the following definition:

$$c(\phi \to \psi)\begin{cases} > 0 & if \;\; Pr(\psi \mid \phi) > Pr(\psi), \\ = 0 & if \;\; Pr(\psi \mid \phi) = Pr(\psi), \\ < 0 & if \;\; Pr(\psi \mid \phi) < Pr(\psi). \end{cases} \qquad [4]$$

Under the "closed world assumption" adopted in inductive reasoning, and because *U* is a finite set, it is legitimate to estimate probabilities Pr(ϕ) and Pr(ψ) in terms of frequencies *sup*(ϕ)/|*U*| and *sup*(ψ)/|*U*|, respectively. In consequence, we can define the conditional probability as Pr(ψ|ϕ) = Pr(ϕ∧ψ)/Pr(ϕ), and it can be regarded as *sup*(ϕ→ψ)/*sup*(ϕ). Thus, the above condition can be re–written as:

$$c(\phi \to \psi)\begin{cases} >0 & if \quad \dfrac{sup(\phi \to \psi)}{sup(\phi)} > \mathrm{sup}(\psi)/|U|, \\ =0 & if \quad \dfrac{sup(\phi \to \psi)}{sup(\phi)} = \mathrm{sup}(\psi)/|U|, \\ <0 & if \quad \dfrac{sup(\phi \to \psi)}{sup(\phi)} < \mathrm{sup}(\psi)/|U|. \end{cases} \qquad [5]$$

Since *RI*, *gain* measure (iff $\Theta = sup(\psi)/|U|$) and *dependency factor* are Bayesian confirmation measures (Szczęch 2007), we propose to normalize them, so that they would distinguish between two completely different situations: situation α in which confirmation occurs (i.e. when $sup(\phi\to\psi)/sup(\phi) \geq sup(\psi)/|U|$) and situation β in which disconfirmation occurs (i.e. when $sup(\phi\to\psi)/sup(\phi) < sup(\psi)/|U|$). Inspired by (Crupi *et al*. 2008), who have analyzed a group of normalized Bayesian confirmation measures, we propose to normalize *RI*, *gain* and *dependency factor* by dividing them by the maximum value they obtain in case of confirmation, and by the absolute minimum value they obtain in case of disconfirmation. In this way, we will obtain confirmation measures taking values from the interval [−1, 1]. The issue of normalizing measures keeps gaining significance in the literature and has also been taken up in (Diatta *et al.*, 2007).

There are many approaches to determining those maximum and minimum values, which eventually lead to different normalizations. In this paper, we present the normalization, inspired by the approach of (Nicod 1923), for which we consider only cases in which there is the evidence, while we ignore cases where there is no evidence. For example, in case of "*all ravens are black*", the evidence is "*raven*" and the hypothesis is "*black*". In this situation, a *black raven* supports the conclusion that *all ravens are black*, *non-black ravens* are against this conclusion, and everything which is *not a raven* can be ignored.

For the clarity of presentation of the normalized measures, the following notation shall be used from now on:

$$a = sup(\phi \to \psi),\ b = sup(\neg\phi \to \psi),\ c = sup(\phi \to \neg\psi),\ d = sup(\neg\phi \to \neg\psi),$$
$$a + c = sup(\phi),\ a + b = sup(\psi),\ b + d = sup(\neg\phi),$$
$$c + d = sup(\neg\psi),\ a + b + c + d = |U|.$$

We assume that set U is not empty, so that at least one of a, b, c, d is strictly positive, and that any value in the denominator of any ratio is different from zero.

Based on the above notation, the *rule interest* function [1] can be expressed as:

$$RI = a - \frac{(a+b)(a+c)}{a+b+c+d} \qquad [6]$$

and thus the normalized *RI* should take the following form:

$$RI_{norm} = \begin{cases} \dfrac{a - \dfrac{(a+b)(a+c)}{a+b+c+d}}{a + c - \dfrac{(a+b+c)(a+c)}{a+b+c+d}} & \text{in case of confirmation} \\[2ex] \dfrac{a - \dfrac{(a+b)(a+c)}{a+b+c+d}}{\dfrac{b(a+c)}{a+b+c+d}} & \text{in case of disconfirmation} \end{cases} \qquad [7]$$

In case of the *gain* measure, its definition [2] can be expressed as:

$$gain = a - \Theta(a+c) \qquad [8]$$

and the normalized *gain* should be defined as:

$$gain_{norm} = \begin{cases} \dfrac{a - \Theta(a+c)}{(a+c)(1-\Theta)} & \text{in case of confirmation} \\[2ex] \dfrac{a - \Theta(a+c)}{(a+c)\Theta} & \text{in case of disconfirmation} \end{cases} \qquad [9]$$

The *dependency factor* [3] takes the following form in the applied notation:

$$\eta = \frac{\dfrac{a}{a+c} - \dfrac{a+b}{a+b+c+d}}{\dfrac{a}{a+c} + \dfrac{a+b}{a+b+c+d}} \qquad [10]$$

which determines the following definition of the normalized *dependency factor*:

$$\eta_{norm} = \begin{cases} \dfrac{\dfrac{a}{a+c} - \dfrac{a+b}{a+b+c+d}}{\dfrac{a}{a+c} + \dfrac{a+b}{a+b+c+d}} \times \dfrac{1 + \dfrac{a+b+c}{a+b+c+d}}{1 - \dfrac{a+b+c}{a+b+c+d}} & \textit{in case of confirmation} \\[2ex] \dfrac{\dfrac{a}{a+c} - \dfrac{a+b}{a+b+c+d}}{\dfrac{a}{a+c} + \dfrac{a+b}{a+b+c+d}} & \textit{in case of disconfirmation.} \end{cases} \qquad [11]$$

We have also considered, however not included in the further analysis in this paper, two other normalization approaches called “Bayesian” and “likelihoodist” (Fitelson 2007). The first one is related to the idea that the evidence confirms the hypothesis, if the hypothesis is more frequent with the evidence rather than with ¬evidence, and in this context, analogously, the evidence disconfirms the hypothesis, if ¬hypothesis is more frequent with the evidence rather than with ¬evidence. The second approach is related to the idea that the evidence confirms the hypothesis, if the evidence is more frequent with the hypothesis rather than with ¬hypothesis, and in this context, analogously, the evidence disconfirms the hypothesis, if the evidence is more frequent with ¬hypothesis rather than with the hypothesis.

2.4. *Property M of monotonicity*

Greco, Pawlak and Słowiński have proposed in (Greco *et al.*, 2004) property M of monotonic dependency of an interestingness measure on the number of objects satisfying or not the premise or the conclusion of a rule. Formally, an interestingness measure F satisfies the property M if:

$$F\left[\sup(\phi \to \psi), \sup(\neg\phi \to \psi), \sup(\phi \to \neg\psi), \sup(\neg\phi \to \neg\psi)\right] \qquad [12]$$

is a function non-decreasing with respect to $\sup(\phi \to \psi)$ *and* $\sup(\neg\phi \to \neg\psi)$, and non-increasing with respect to $\sup(\neg\phi \to \psi)$ *and* $\sup(\phi \to \neg\psi)$.

The property M with respect to $sup(\phi \to \psi)$ (or, analogously, with respect to $sup(\neg\phi \to \neg\psi)$) means that any evidence in which ϕ and ψ (or, analogously, neither ϕ nor ψ) hold together increases (or at least does not decrease) the credibility of the rule $\phi \to \psi$. On the other hand, the property of monotonicity with respect to $\sup(\neg\phi \to \psi)$ (or, analogously, with respect to $\sup(\phi \to \neg\psi)$) means that any evidence in which ϕ does not hold and ψ holds (or, analogously, ϕ holds and ψ does not hold) decreases (or at least does not increase) the credibility of the rule $\phi \to \psi$.

Let us present the interpretation of property M on the following example used in (Hempel 1945). Let us consider a rule $\phi \to \psi$: *if x is a raven, then x is black*. In this case, ϕ stands for the property of being a raven and ψ is the property of being black. If an attractiveness measure $F(\phi \to \psi)$ possesses the property M, then:

- the more black ravens there are in the dataset, the more credible is the rule, and thus $F(\phi \to \psi)$ obtains greater (or at least not smaller) values,
- with the increase of the number of non-black non-ravens $F(\phi \to \psi)$ also obtains greater (or at least not smaller) values,

- the more black non-ravens appear in the dataset, the less credible becomes the rule and thus, $F(\phi\rightarrow\psi)$ obtains smaller (or at least not greater) values,
- the more non-black ravens are the dataset, the less credible is the rule and thus, $F(\phi\rightarrow\psi)$ obtains smaller (or at least not greater) values.

2.5. *Support–Anti-support Pareto-optimal border*

Let us denote by $\preceq_{s\sim a}$ a partial preorder given by the dominance relation on a set X of rules with the same conclusion, taking into account two interestingness measures *support* and *anti-support*, i.e. given a set X and two rules r_1, $r_2 \in X$, $r_1 \preceq_{s\sim a} r_2$ if and only if

$$sup(r_1) \leq sup(r_2) \wedge anti\text{-}sup(r_1) \geq anti\text{-}sup(r_2).$$

Recall that a *partial preorder* on a set X is a binary relation R on X that is reflexive and transitive. The partial preorder $\preceq_{s\sim a}$ can be decomposed into its asymmetric part $\prec_{s\sim a}$ and its symmetric part $\sim_{s\sim a}$ in the following manner: given a set X and two rules r_1, $r_2 \in X$, $r_1 \prec_{s\sim a} r_2$ if and only if

[13]

$$sup(r_1) \leq sup(r_2) \wedge anti\text{-}sup(r_1) > anti\text{-}sup(r_2), \text{ or}$$
$$sup(r_1) < sup(r_2) \wedge anti\text{-}sup(r_1) \geq anti\text{-}sup(r_2),$$

moreover, $r_1 \sim_{s\sim a} r_2$ if and only if

[14]

$$sup(r_1) = sup(r_2) \wedge anti\text{-}sup(r_1) = anti\text{-}sup(r_2).$$

If for a rule $r \in X$ there does not exist any rule $r' \in X$, such that $r \prec_{s\sim a} r'$, then r is said to be *non–dominated* (i.e. *Pareto-optimal*) with respect to support and anti-support. A set of all non-dominated rules with respect to these measures is also referred to as a *support–anti-support Pareto-optimal border*. In other words, it is the set of rules such that there is no other rule with the same conclusion, having a greater support and a smaller anti-support.

The approach to evaluation of a set of rules with the same conclusion in terms of two interestingness measures being rule support and anti-support was proposed and presented in detail in (Brzezińska *et al.*, 2007), and also studied in (Słowiński *et al.*, 2006). The idea of combining those two dimensions came as a result of looking for a set of rules that would include all rules optimal with respect to any interestingness measure with the desirable property M.

Theorem 1. (Brzezińska *et al.*, 2007) When considering rules with the same conclusion, rules that are optimal with respect to any interestingness measure that has the property M must reside on the support–anti-support Pareto-optimal border.

The above theorem states that the best rules according to any interestingness measures with M are in the set of non-dominated rules (i.e. objectively, the best) with respect to support–anti-support. This valuable and practical result allows to identify a set of rules containing most interesting (optimal) rules according to any interestingness measures with the property M, simply by solving an optimized rule mining problem with respect to rule support and anti-support.

3. Analysis of normalized measures with respect to property M

In order to prove that a normalized measure has property M, we need to show that it is non-decreasing with respect to *a* and *d* and non-increasing with respect to *b* and *c* both in case of confirmation and disconfirmation.

Theorem 2. The normalized *rule interest* function has the property M.

Proof: We will only present the proof that the normalized *RI* is non-decreasing with respect to *a* in case of confirmation, and omit the other proofs as they are analogous. Through simple mathematical transformation, we obtain the following form of the normalized *RI* in case of confirmation (for simplicity denoted by $RI_{\text{norm+}}$):

$$RI_{norm+} = \frac{ad - bc}{ad + cd} \,. \qquad [15]$$

$RI_{\text{norm+}}$ will be non-deceasing with *a* if and only if an increase of *a* by $\Delta>0$ will not result in a decrease of $RI_{\text{norm+}}$. Simple algebraic transformations show that:

$$\frac{(a+\Delta)d - bc}{(a+\Delta)d + cd} - \frac{ad - bc}{ad + cd} = \frac{cd\Delta + bc\Delta}{d(a+\Delta+c)} > 0 \,. \quad \square \qquad [16]$$

Theorem 3. The normalized *gain* measure has the property M.

Proof: Analogous to the proof of Theorem 2.

Theorem 4. The normalized *dependency factor* does not have property M.

Proof: The normalized *dependency factor* in case of confirmation does possess property M (the proof is analogous to proof of Theorem 2), however, we can prove by a counterexample that the normalized *dependency factor* in case of disconfirmation (for simplicity denoted as η_{norm-}) does not have property M: Let us consider case α, in which *a*=7, *b*=2*, c*=3, *d*=3, and case α', in which *a* increases to 8 and *b*, *c*, *d* remain unchanged. The normalized *dependency factor* in case of disconfirmation does not have property M as such increase of *a* results in the decrease of the measure: $\eta_{norm-}(\phi \to \psi) = 0.0769 > 0.0756 = \eta'_{norm-}(\phi \to \psi)$. □

4. Consequences for the user - practical application of the results

In the previous section, we have proved that the normalized measures of *rule interest* function and *gain* possess the property M, while the normalized *dependency factor* does not have this property. These results are of practical value as they show that rules optimal with respect to normalized *RI* or *gain* reside on the Pareto-optimal border with respect to support and anti-support (when considering rules with the same conclusion). Moreover, they allow potential efficiency gains as:

- rules optimal with respect to normalized *RI* or *gain* can be found in the support–anti-support Pareto-optimal set instead of searching the set of all rules,
- rule evaluation can be narrowed down to mining only the support–anti-support Pareto-optimal set instead of conducting rule evaluation separately with respect to normalized *RI, gain*, or any other measure with property M, as we are sure that rules optimal according to normalized *RI*, *gain*, or any other measure with property M, are in that Pareto set.

To illustrate practical application of the above theoretical results, we have conducted several computational experiments analyzing rules optimal with respect to normalized *RI* in case of confirmation. In Figure 1, there is an exemplary diagram from those experiments presenting induced rules in the perspective of rule support and anti-support. For a real life dataset containing information about technical state of buses, a set of all possible rules for which the premise confirms the conclusion was generated. A set of rules with the same conclusion was then isolated and rules non-dominated with respect to support and anti-support were found (those rules form the support–anti-support Pareto-optimal border). The support–anti-support Pareto-optimal border is indicated in Figure 1 by circles connected by a line. In the generated set of rules, by empty red squares we have distinguished rules optimal according to the normalized *RI* in case of confirmation (these are rules with $sup(\phi \to \psi)=49$ and $anti\text{-}sup(\phi \to \psi)=2$, or with $sup(\phi \to \psi)=50$ and $anti\text{-}sup(\phi \to \psi)=4$). The diagram shows that, indeed, rules optimal with respect to the normalized *rule interest* function in case of confirmation lie on the support–anti-support Pareto-optimal border.

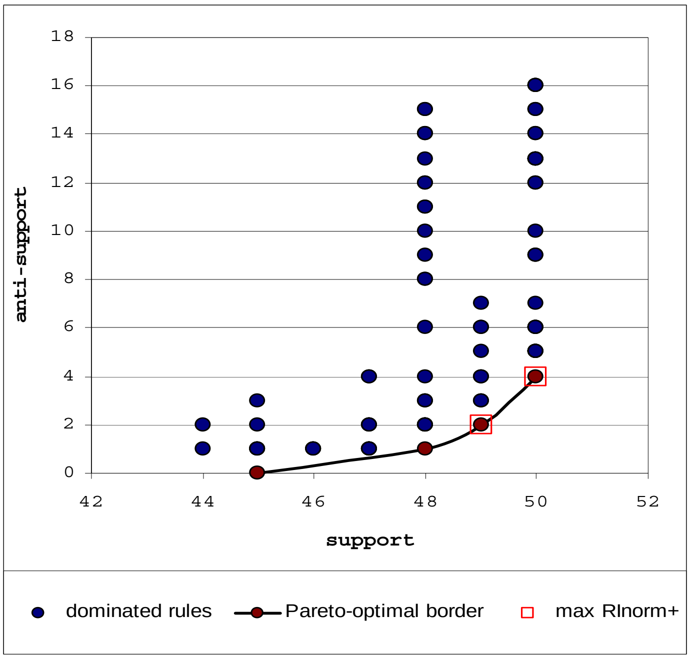


Figure 1. *Support–anti-support Pareto-optimal border*

5. Conclusions

In this paper, we have considered three popular measures: *rule interest* function, *gain* measure and *dependency factor*. The normalization consists in dividing the measures by the maximum value they obtain in case of confirmation, and by the absolute minimum value they obtain in case of disconfirmation. In this way, while keeping the interval of variation [-1, 1], the measures behave differently in case of confirmation and in case of disconfirmation.

A theoretical analysis of the new normalized measures with respect to valuable property M has been conducted. It has been proved that the normalized measure *RI* and *gain* satisfy property M, while the normalized *dependency factor* does not possess this property. The possession of property M implies that rules optimal with respect to the normalized *RI* and *gain* will be found on the support–anti-support Pareto-optimal border (when considering rules with the same conclusion). These results have also been illustrated on an exemplary dataset. It is, therefore, legitimate to conclude that rule evaluation can be narrowed down to mining only the support–anti-support Pareto-optimal set instead of conducting rule evaluation separately with respect to normalized *RI, gain*, or any other measure with property M, as we are sure that rules optimal with respect to normalized *RI*, *gain*, or any other measure with property M, are in that Pareto set.

6. References

Brzezińska I., Greco S., Słowiński R., "Mining Pareto–optimal rules with respect to support and anti-support", *Engineering Applications of Artificial Intelligence*, vol. 20 no. 5. 2007, p. 587-600.

Crupi V., Tentori K., Gonzalez M., "On Bayesian measures of evidential support: Theoretical and empirical issues", *Philosophy of Science* 2008.

Diatta J., Ralambondrainy H., Totohasina A., "Towards a Unifying Probabilistic Implicative Normalized Quality Measure for Association Rules", *Quality Measures in Data Mining Series: Studies in Computational Intelligence,* vol. 43, 2007.

Fitelson B., *Studies in Bayesian Confirmation Theory*. Ph.D. Thesis, University of Wisconsin, Madison 2001.

Fitelson B., "Likelihoodism, Bayesianism, and Relational Confirmation", *Synthese*, vol. 156 no. 3, 2007, p. 473-489. Springer Netherlands.

Fukuda T., Morimoto Y., Morishita S., Tokuyama T., „Data Mining using Two-Dimensional Optimized Association Rules: Schemes, Algorithms, and Visualization" *Proc. of the 1996 ACM SIGMOD Int'l Conference on Management of Data*, 1996, p. 13-23.

Greco S., Pawlak Z., Słowiński R., „Can Bayesian confirmation measures be useful for rough set decision rules?" *Engineering Applications of Artificial Intelligence*, vol. 17, 2004, p345–361.

Hempel C.G., "Studies in the logic of confirmation (I)" *Mind* vol. 54, 1945, p. 1-26.

Hilderman R.J., Hamilton H.J., "Applying Objective Interestingness Measures in Data Mining Systems", *Fourth European Symposium on Principles of Data Mining and Knowledge Discovery (PKDD-2000),* Lyon, France, Springer Verlag, September 2000 p. 432-439.

Lenca Ph., Meyer P., Vaillant B., Lallich S., "On selecting interestingness measures for association rules: User oriented description and multiple criteria decision aid", *European Journal of Operational Research*, Vol. 184, Issue 2, 2008, p. 610-626.

Nicod J., *Le problème de la logique de l'induction*, Alcan, Paris, 1923.

Pawlak Z., "Some issues on rough sets", *Transactions on Rough Sets I,* LNCS, 2004, p. 1-58.

Piatetsky-Shapiro G., "Discovery, analysis and presentation of strong rules", *Knowledge Discovery in Databases*, AAAI/MIT Press, 1991, pp. 229-248.

Popper K.R., *The Logic of Scientific Discovery*. Hutchinson, London 1959.

Słowiński R., Brzezińska I., Greco S., „Application of Bayesian confirmation measures for mining rules from support-confidence Pareto-optimal set", *Invited paper in Artificial Intelligence and Soft Computing*. LNAI 4029, Springer, Berlin, 2006, pp. 1018-1026.

Szczęch I., *Multicriteria Attractiveness Evaluation of Decision and Association Rules*, Ph.D. Thesis, Poznań University of Technology, Poznań, Poland, 2007.

Towards a human centred methodology for dynamic allocation of functions

Th. Coye de Brunélis* — P. Le Blaye*

**ONERA/DCSD - The French Aerospace Lab*
BA 701
13 611 SALON CEDEX AIR
France

Thierry.coye_de_brunelis@onera.fr
Patrick.le_blaye@onera.fr

ABSTRACT: Few methods already exist to guide the design of a system based on Dynamic Function Allocation (DFA). At best they are incomplete and mostly related to the classical -static- Function Allocation (FA) process (Cook et al., 1999, Dearden et al., 2000) while the existing methods for the design of adaptive interfaces do not take into account all the issues raised by static and dynamic FA (Rothrock et al., 2002). In this paper, we review the existing approaches and issues regarding static and dynamic FA then we propose a preliminary methodology for the design and the evaluation of a system implementing DFA. Finally we link the methodological steps to a possible structure for DFA implementation, in an attempt to integrate in a single process the different necessary methodological approaches previously considered separately.

KEY WORDS: Dynamic Function Allocation; System Design; Human-centred automation

1. Introduction

The Function Allocation (FA) is a well identified step of the design of complex high risk interactive systems that deals with the fundamental issue of allocating function between the human and the technical system. However the classical (static) function allocation faces some recurrent issues, associated with high levels of automation: unbalanced workload, loss or degradation of situation awareness, loss or degradation of skills (Hoc, 2000).

As other researchers (Rouse, 1988) we consider dynamic function allocation (DFA) as a promising way to improve overall system performance. DFA refers to the variable distribution of functions in real time between the system and the operator(s) to achieve optimal system performance (Cook *et al*, 1999). The field of DFA is an active area of research. Existing research on DFA includes discussion regarding critical human performances issues related to DFA, investigation of specific issues and phenomena thought to be relevant to DFA and empirical evaluation of DFA (Scallen & Hancock, 2001).

In addition, adaptive interfaces are an active field of research in the human-computer interaction community, addressing many of the issues raised by DFA from a different but complementary point of view. Keeble and Macredie (2000) define an adaptive interface as “one where the appearance, function or content of the interface can be changed by the interface (or the underlying application) itself in response to the user’s interaction with it”. In this way, this field of research only addresses the interface of the system with the operator, while the field of FA focuses more deeply on the functions of the technical system. Nevertheless, the interface is an essential component of the DFA implementation and a consensus is that it is to a large extent responsible for the usability of the system (Palmer *et al*, 1995 ; Older *et al*., 1997).

In the present paper, the limitations of static FA and the expected benefits of DFA will be first presented. Second, the issues associated with the implementation of DFA will be considered with the aim to derive some guidelines for the design of a system using DFA. Third, a preliminary methodology and a generic structure for DFA implementation will be proposed.

2. From classical function allocation to dynamic function allocation

2.1. The classical steps of function allocation in interactive system design

FA is a well identified step in the design of an interactive system (EN ISO 11064-1, EN ISO 13407) that seek to allocate functions to human or system or both. According to Dearden *et* al. (2000), “allocation of function is an early stage of the design of a human-machine system. The input to allocation of function is a specification of the functions that the human-machine system must deliver within its intended working context. The output from allocation of function is a specification, at an appropriate level of abstraction, of the functionality of the automated sub-systems that will be required. The goal of allocation of function is to design a system for which the performance (including considerations of safety and reliability) is high; the tasks of the operator are achievable and appropriate to the operator's role; and the development of the system is technically and economically feasible".

Following this approach, Dearden *et al*. proposed a method for FA that starts with an identification of the functions required for the intended use of the system. It then identifies the functions which must be kept allocated to the operator, based on the role of the operator and legal and technical constraints. The next steps proceed by examining a collection of scenarios. Each scenario may involve many functions that must be delivered concurrently. The scenario helps to expose the complexity of the working context that is being designed, in an approach similar to scenario-based design (Carroll, 1995). The method considers the relative merits of the set of design options for the functions relevant to each scenario, in terms of workload, situation awareness and performance.

This FA process results in partial automation or different levels of automation (LoA) for each function, so that human and automation perform a variable part of the task to achieve each function. Through their definition Dearden *et al*. stressed that technical and economical feasibility of the system are essential dimensions to be taken into account, together with other characteristics such as safety and reliability, especially in the case of high risk systems (Johnson *et al*., 2000).

Some other methods for FA can be found in the literature; whatever the method, the design of static FA raises some issues regarding the interaction with the human operator.

2.2. Ironies of automation and limitations of static FA

Many issues have been identified and can happen when high levels of automation are used (Hoc, 2000; Parasuraman, 1997). Bainbridge (1987) talked of “ironies of automation” to describe paradoxical system designs where automation is used to improve common human-machine performances but eventually creates more unsafe conditions. Many studies have been devoted to the failures in coupling human and automation, especially within the aviation domain where such failures may have obvious consequences on safety (Parasuraman & Mouloua, 1996). The general term used to describe the “automation induced effects on human operators” covers many different aspects such as the nature of functions to be automated (stage of information processing), their level of automation, the design of the interface of the automated system, the human-machine dialog, and the global performance of the system… Although FA is primarily concerned with the level of automation, its success also highly depends on, for instance, the design of the interface. As such it is strongly involved in the 3 induced effects described hereafter (Hoc, 2000).

<u>Unbalanced (mental) workload:</u> Operators possess a flexible but ultimately limited attentional capacity. An excessive demand on resources imposed by the

task(s) typically results in performance degradation. Automation can sometimes induce extreme levels of workload, either too low or too high. When the operator faces long periods with reduced workload because of high automation, a loss of vigilance may appear and put the operator out-of-the loop. On the contrary, when the operator suddenly has to take up a critical situation because the system hands off, peaks of workload may appear and lead to operator's performance degradation.

Situation awareness (SA) degradation: SA refers to the level of awareness that an actor has of the current situation which he/she is placed in and its projection in the future (Endsley, 1995). Automation which assists operators reduces the occasions for them to interact with the system. So, because of these limited interactions, operators may experience difficulties to update their knowledge of the system states and thus to maintain their SA. The resulting degradation of the SA – described as an out-of-the-loop syndrome- partakes of the operator's incapacity to handle critical situations as he/she does not know what happens and does not have enough time to acquire the necessary information.

Loss or degradation of skills and expertise: Human loss of skills and expertise is the consequence of designing machines that play autonomous roles, either performing low-level functions (decision implementation) or high-level functions (diagnosis, decision-making) (Hoc, 2000). Bainbridge stressed these "ironies of automation" as aiming to assist the operator and at the same time reducing the occasions to maintain him/her skills. When operators must take up manual control, they are likely to reach a poor performance.

FA faces two limitations. The first one is linked to the fact that levels of automation are static by definition. Functions and tasks are allocated for the whole system life and human-machine interactions could be limited leading to automation induced effects on human operators. The second one is the possible incompatibility of the different solutions generated by the FA process for the different missions the system has to fulfil. How to deal with these limitations?

DFA appears as a possible solution to reduce classical automation induced effects and to support human-computer interactions including supervision activity. On one hand the dynamic reallocation of functions and tasks during process could help to solve or to reduce the human-automation interaction induced effects of unbalanced workload, SA degradation and skills degradation. In this case, DFA can be seen as an intra-mission solution. On the other hand DFA could permit to switch from a specific mission allocation to another specific mission allocation. DFA is then an inter-mission allocation solution.

2.3. Dynamic Function Allocation, a human centered automation for more efficient design

Given a system that includes some combinations of human and/or automated components or agents, DFA refers to the process of redistributing tasks or functions amongst those agents with the goal of improving overall system performance (Cook *et* al, 1999).

DFA can be seen as a complementary process of static FA in the design process. As suggested by the Dearden method, DFA is a step beyond in complex system design, and can not be considered separately from FA. Static FA is the necessary foundation supporting DFA implementation. The previous FA leitmotiv “who does what” becomes with DFA “who does what and when” (Figure 1).

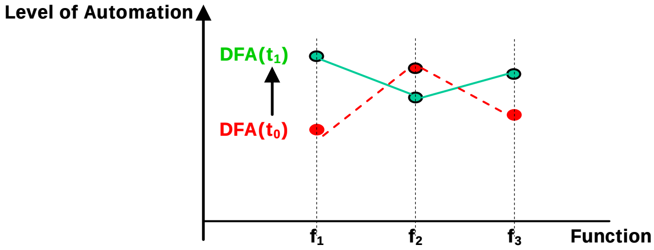


Figure 1. *The modification of the levels of automation through DFA*

The human centered approach to the design of automation explicitly considers the impacts of the introduction of automation on the humans in the system and on the overall behaviour of the system, from the beginning and continuously throughout the design process. The goal of human centered automation is to support human efforts rather than to replace them. The human-centred automation theory proposed by Billings (1991, 1997) states that complex systems should be designed to keep the human in the loop with the goal of improving overall system performance. This goal could be achieved by maintaining a balanced workload, supporting the operator’s achievement of SA and preventing skill degradation through meaningful involvement (Endsley, 1995; Endsley & Kaber, 1999). As such, we consider that DFA should actually lies within human centered design.

Interest in DFA has increased within the recent past as a result of hypothesized benefits. Though the expected benefits of DFA are encouraging, some specific issues emerge regarding it use and implementation. We hereafter discuss these issues and propose an approach to address these issues during DFA implementation.

3. Specific issues related to DFA

3.1. Current implantation of DFA

The concept of DFA is not a new one. Human-to-human reallocation has been around ever since people began working in teams. DFA amongst system hardware elements is also relatively common, as when a redundant system component takes over for another component that is overloaded or has failed. Even human-to-automation DFA, the newest possibility on the horizon, has been discussed and investigated for over 20 years (Rouse *et al.*, 1986). DFA is already implemented in some systems such as for instance in commercial aircraft where an automatic reversion of the autopilot modes can appear if some thresholds are overshot. The pilot can also switch the autopilot in various modes by him/herself. Cases of clumsy automation have been reported with such DFA implementation where the pilot looses SA regarding the conditions of autopilot reversion. Pilot then has to supervise a degraded situation with little time and a limited understanding of the system state and behaviour. Our proposed design methodology aims to face these automation induced problems.

DFA promises to provide a means of achieving a dynamic, maximally efficient blending of human and automated resources to effectively accomplish complex objectives. Many issues appear when considering the DFA implementation in a complex dynamic system. Who is supposed to make the decisions concerning when and how FA must be altered? Who is in charge of the detection of the adaptation need? Which DFA criterions should be taken into account to trigger adaptation? Who elaborates the adaptation change proposals? Which tasks or functions could be shared (what)? Who has the final authority to decide of the triggering? Who performs the DFA triggering?

3.2. Authority sharing for DFA (who)

3.2.1. Human vs. technical system for triggering DFA

When reviewing the literature, the issue of authority sharing (to decide on the allocation changes) appears as opposing the human operator to the technical system. If the triggering is performed by the operator -the authority remains to the operator- the allocation is called *explicit*. If the triggering is made by the technical system -the authority remains to the system- the allocation is said *implicit* (Millot & Lemoine, 1998; Cook *et al.*, 1999). The dynamic reallocation of function is made through a part of the system called the "allocator core" (Figure 2).

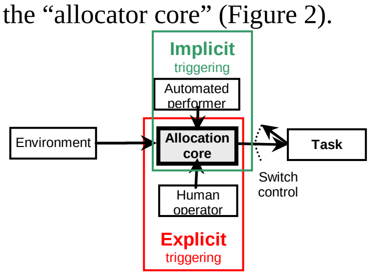


Figure 2. *Explicit vs. implicit DFA triggering (adapted from Wickens, 1992)*

One of the major advantages to keep the operator responsible for DFA triggering is that the operator remains aware of the reasons leading to DFA triggering which do lead to minimize automation surprises. On another hand the decision making process may create an over workload and interact strongly with the main task to be performed particularly in case of a critical event or of a degraded system state.

The major advantage of an adaptation triggered by the system is that the operator is totally unloaded of the triggering task which may avoid an inappropriate overload (especially during critical situations). Furthermore a well designed system could anticipate the triggering and perform *ad hoc* adaptation as required by the situation. From another hand, if the operator is not informed about the proposal process, an out-of-the-loop effect could appear.

In high risk domains such as aviation, the application of implicit DFA to functions central for the operator's role is questionable. As the human operator remains responsible for the ultimate safety of the system, a hierarchy *de facto* exists between the operator and the system (Millot & Lemoine, 1998). Moreover, the automation may not be totally reliable given the context. As a consequence, the operator has to monitor the system carefully which may produce burden on the operator in addition to the original tasks. Also, human-centered automation claims that the human operator must be maintained as the final authority and must be able to exercise decision concerning when and how function allocation must be changed (Billings, 1997, Woods *et al.*, 1989, Palmer *et al.*, 1995). With regard to the Sheridan & Verplanck (1978) scale, the human- centered automation principle is violated when the LoA may be positioned at level 7 and more.

On another hand, Rouse (1988) argues that "when an aid is most needed, it is likely that humans will have few resources to devote to interacting with the aid. Put simply, if a person had the resources to instruct and monitor an aid, he/she would probably be able to perform the task with little or no aiding". This statement pleads in favour of an implicit allocation. Furthermore there are cases where good reasons exist for automation to trigger DFA (e.g. safety nets in aviation or in nuclear power plants when a precise and immediate reaction is required in response of an emergency).

As states Inagaki (2003), "a clear-cut answer is hard to get for the decision authority issue".

3.2.2. Adaptation change proposal

3.2.2.1. A link to cooperation

DFA involves human-machine cooperation. Although there is no consensus on a clear definition of cooperation (Debernard, 2006), the research field of cooperation provides a foundation to consider the cooperation activities involved in the triggering of DFA. Among others, Hoc (1998) offers a descriptive framework of cooperation at a high level of abstraction that is interesting as a guideline for design or evaluation but is not sufficient to specify DFA implementation.

Campbell *et al*. (1997) argue that before making a decision regarding allocating functions or tasks to another agent an experienced agent would (or should):

1. **Assess his own state** with regard to workload and capability to perform the activity,
2. **Estimate the cost of allocating the function** such as the increase in communications, monitoring and interactions,
3. **Make a determination** regarding the ability of the second agent to perform the activity, and,
4. **Make judgments** regarding the trust and confidence he has in the agent successfully performing the activity.

This decomposition highlights the different steps needed for a human agent to make a decision to allocate functions to another agent (human or technical). By the way, it suggests that in case of an already high loaded situation the triggering process may create an operator's overload, which has been identified as an issue of implicit DFA.

3.2.2.2. Application to implicit DFA

Based on the previous decomposition, we propose hereafter a human-like decomposition of the steps the allocator core of an implicit DFA implementation should perform to allocate the functions to the human or to the technical system. The previous decomposition indeed does not specify the need for the allocator to identify which functions or part of functions should be allocated (after step 1). Next it does not show that the allocator may have to define one or more proposal(s) regarding the appropriate level of automation and then to submit it to the operator (need for a feedback). As a result, the allocator should be able to:

1. **Assess the situation** with regard to capability of the human and technical agents to perform the activity,
2. **Assess which function(s)** or part of function(s) the system could take over (*e.g.* in case of operator overload) or hand off (*e.g.* in case of operator underload) for the handling of the situation,
3. **Estimate the cost** of allocating the function to the operator,
4. **Make a determination** regarding the ability of the operator to perform the activity, including the information required,
5. **Assess at what time** the reallocation can be performed,
6. **Make some triggering proposals** and propose them to the operator in case of operator final authority for triggering decision.

3.3. Which tasks or functions should be shared (what)?

The answer to this question closely depends on the benefits expected from the DFA. In chapter 2 we discussed the limitations of classical FA and the potential enhancements DFA could bring: regulation of workload (overloading or underloading), maintaining SA or skills, improvement of overall system performance… All those goals are linked together and acting on one may impact the others. Once again trades off are required. The choice of the functions or tasks to delegate or to share with another agent should be based on the priority set between these possible goals (workload, SA, skills or performance).

The technical feasibility of the function to be delegated and moreover, its criticality with regard to the role and responsibility of the operator are also key parameters to be taken into account. For instance in case of a high workload of the user during a critical event, the tasks or functions to be delegated will be preferentially secondary tasks, unessential for the situation handling.

3.4. DFA criterions for an appropriate triggering (when)

The question “When triggering DFA?” takes on 2 different forms.

On one hand, the question relates to the appropriate time to trigger the allocation during the user’s activity. The McFarlane’s taxonomy of human interruption can be used as a starting point (McFarlane & Latorella, 2002) to define the most appropriate conditions and time when the DFA should occur, with the main goal to minimize the interferences with the normal behaviour of the user.

On the other hand, the question refers to the type of indicator on which to base the allocation triggering during the system process. To this end, 4 types of criterions can be taken into account. Kaber & Endsley (2003) propose 4 types of DFA criteria based on a synthesis of the literature as follow:

- **Critical events**: DFAs triggered by occurrence of events critically impacting system goals (*e.g.* malfunction),
- **Performance measurement**: DFAs triggered by degradation in human monitoring performance below a criterion measure (Parasuraman, 1993),
- **Psychophysiological assessment**: real time assessment of operator workload using for example physiological measures as electroencephalogram or heart-rate variability as a basis for decision to automate (e.g. Byrne & Parasuraman, 1996),
- **Behaviour modelling**: DFAs occur to human and computer to achieve predetermined pattern of overall system functioning (Rouse *et al*., 1986).

These criteria may be implemented in an allocator core through *production rules* such as “If criteria C is detected, then function *F* must be handed over to the automation, if the function was dealt with by the human at that time point” (Inagaki, 2003).

The question when the dynamic reallocation should occur calls for a discussion regarding how the trigger (whether time or criteria based) may be detected.

3.5. How may the need for triggering be detected?

We advocate a human-centred approach to this question. The choice and use of the criteria (or combination of criteria) to implicitly trigger the DFA should ideally be consistent with what the user would do in an explicit condition.

Among the different types of criteria, the critical-event criteria may be the most straightforward to implement, if the critical events are defined properly. No investigations are needed regarding how human cognition or behaviour could be modelled, what parameters must be measured to infer the human state, and how. However the detection of these critical events may require a specific technology or availability of information in the system. A possible limit to that it doesn't take the operator's state into account directly; as stated by Inagaki (2003). The subjective workload under critical event may significantly differ among operators. Also, environmental and system events may be taken into account: we hypothesize that current non critical system events can be used to re-involve or maintain human operator in the loop during long under loaded phases.

A psychophysiological measurement-based logic can trigger DFA by explicitly reflecting the mental status of an operator at a specific circumstance. However, a first limitation is that sensing devices are highly dependant on the individual and too sensitive to local fluctuations in the operator's workload or physiological states. Also, they're often expensive and not every operator may welcome situations in which he is monitored by sensing devices all the time. And finally, performance measurement occurs "after the fact", that is after a point in time when adaptation may be needed (Inagaki, 2003; Scerbo *et al*, 2001). Furthermore as Campbell *et al*. (1997) stated it is generally not sufficient to base function reallocation decisions solely on an assessment of workload change. Some other measures of situational and agent state may be required.

It may be possible to extract leading indicators from good behaviour models; for instance some attempts have been made to model a procedural activity using formal tools such as Petri nets, and to follow the actual activity in real time through this type of model. Some allocation changes can then be triggered as planned within the model. However it is obviously not always easy to develop models that appropriately represent the reality and address all the relevant possible situations.

As a conclusion, an efficient DFA implementation should probably not be based on one single type of criteria: a current research direction is to combine the different criteria.

3.6. Guidelines for the implementation of DFA

Different DFA guidelines already exist (Höök, 2000; Palmer et al, 1995; Older et al., 1997). Those guidelines are powerful trend indicators for designers providing them a high level vision of properties a DFA interface should have. We just expose hereafter 3 of the 5 usability challenges summarised by Jameson (2003):

<u>Predictability and transparency:</u> Predictability refers to the extend to which a user can predict the effect of his/her actions and transparency refers to the extent to which he/she can understand system actions and/or have a clear picture of how the system works. Those properties are closely linked to the user's SA regarding the system functioning and the user's trust regarding the system.

Controllability: It refers to the extend to which the user can perform or prevent particular actions or states of the system if he/she has the goal of doing so. This property impacts the way the user has the feeling of controlling the system more than predictability and transparency can do.

Unobtrusiveness: refers to the extend to which the system places demands on the user's attention which reduce the user's ability to concentrate on his/her primary tasks. In particular, it is directly linked to the manner the user's activity may be interrupted. This property may deeply impact the user's performance to achieve his/her functions.

4. Towards a DFA methodology

In order to guide the DFA process and to integrate it in a more general system design process, a preliminary methodology is currently under development in order to provide designers with a step by step efficient process.

This methodology integrates the approach described in the previous chapter. It is based on an already existing U-model (Lepreux *et al.*, 2002) and is compliant with the reference standards (*e.g.* ISO 13 407). An aim is also to establish a link between the cognitive and technical approaches as suggested by Millot (1998).

A first top-down phase leads to the system specification. A preliminary step is to identify the functions required to perform the intended mission(s). The analysis of existing and reference situations is also performed using for instance hierarchical task analysis (HTA). Then the possible static and dynamic function allocation for different nominal and critical scenarios are identified (e.g. Dearden, 2000). The *a priori* modelling of the tasks, environment, system and human operator prepares the computational implementation of DFA. The FA process, the models and the choice of a human-machine cooperation mode result in the specification of the human-computer interface.

The aim of the second bottom-up phase is to analyse and evaluate the actual human-machine interaction in order to refine the design if needed. The performance analysis coupled to a cognitive activity analysis result in the evaluation with regard to the design objectives. The models constructed *a priori* during the top-down phase or the human-computer interface may then be refined.

This iterative design process shall be applied successively to concept principles, to simulation means then to prototype and finally to the industrial system.

5. A general structure for DFA implementation

The methodological steps previously considered lead to identify the different components which should be integrated for DFA implementation:

A. A user model (computational representation of user's expertise, goals, behaviour, performance, given the system context),
B. A tasks/functions model (functions, procedure, specificity such as criticality, difficulty, duration and articulation of the sub-tasks, possible disruption)
C. A system model (automated functions, dynamic behaviour),
D. An environment model (physical and meteorological conditions, interactions with other socio-technical systems),
E. The allocator core (in charge of the appropriate DFA triggering proposal). Itself composed of three sub models:
   a. An interaction model (computational prediction of current and future operator and system's states),
   b. A dialogue model (computational representation of user – machine dialogue)
   c. The user interface for the submission of proposal to the user or selection and validation of the chosen proposal by the user (displays and commands)

Some high level structures exist for DFA or adaptive interfaces (Benyon and Murray, 1993; Rothrock *et al.*, 2002). Our aim is to develop a more detailed generic structure consistent with our methodological approach for design and evaluation. This structure would encompass all the components required between the human operator and the system.

6. Conclusion

A preliminary human centred methodology is proposed for the design of a system involving dynamic function allocation.

General considerations on the basic question "who does what and when" supporting the DFA has lead to identify specific issues and methodological steps to be considered for the implementation of DFA. Our aim is to establish a link between the iterative process to be followed during design and the necessary components of a system supporting DFA.

7. References

Bainbridge, L. (1987) Ironies of automation. In J. Rasmussen, K.D. Duncan and J. Leplat (Eds.), *New technology and human error*, pages 271-284. Chiester, UK : Wiley, 1987.

Benyon, D.R. and Murray, D.M. (1993). Applying user modelling to human-computer interaction design, *Artificial Intelligence Review*, 6, 43-69.

Billings, C. E. (1991). Human-centered aircraft automation: a concept and guidelines. Technical report number 103885. Moffet field, CA: NASA Ames Research Center.

Billings, C. E. (1997). Aviation Automation: The Search for a Human-Centered Approach. Mahw h, NJ : Lawrence Erlbaum Associates.

Byrne, E.A. and Parasuraman, R. (1996). Psychophysiology and adaptive automation, *Biological Psychology*, 42: 249-268

Campbell, G., Cannon-Bowers, J., Glenn, F., Zachary, W., Laughery, R., & Klein, G. (1997). *Dynamic function allocation in the SC-21 Manning Initiative Program.* Orlando, SC-21/ONR S&T Manning Affordability Initiative.

Carroll, J.M. (1995). *Scenario-based design: envisioning work and technology in system development*. New York: Wiley.

Cook, C., Corbridge, C, Morgan, C. & Turpin, E. (1999). People in control: an international conference on human interfaces in control rooms, Cockpits and command centres, 21-23 june 1999, Conference publication No. 463, IEE, 1999.

Dearden, A., Harrison, M. & Wright, P. (2000). Allocation of function: scenarios, context and the economics of effort. Int. J. Human-Computer Studies, 52, 289-318.

Debernard, S. (2006). Coopération Homme-Machine et Répartition Dynamique de Tâches – Application au Contrôle de Trafic Aérien, Mémoire d'Habilitation à Diriger des Recherches, Université de Valenciennes et du Hainaut Cambrésis.

EN ISO 11064-1:2000, Ergonomic design of control centres – Part 1 : Principles for the design of control centres. European committee for standardization, Brussels.

EN ISO 13407:1999, Human-centred design processes for interactive systems. European committee for standardization, Brussels.

Endsley, M.R. (1995) Toward a Theory of Situation Awareness in Dynamics Systems, *Human factors*, 37-1:32-64.

Endsley, M.R. & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42, 462-492.

Hoc, J.-M. (1998). How can we evaluate the quality of human-machine cooperation? COOP'98, 3rd Int. Conf. on the Design of Cooperative Systems, Cannes, France.

Hoc, J.-M. (2000). From human-machine interaction to human-machine cooperation, *Ergonomics*, 43(7), 833-843.

Inagaki, T. (2003) Adaptive automation: Sharing and Trading of Control. In E. Hollnagel (Ed.), *Hdbk of Cog. Task Design*, pages 147-170. London, UK: LEA

Johnson, P., Harrison, M. & Wright, P. (2005) Development of a representation to support function allocation in a systems engineering context – Deliverable 2 : An enhanced function allocation method based upon the York and Sheffield methods, Heslington : University of York.

Kaber, D.B. & Endsley, M.R. (2003). The effect of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task, *Theoretical issues in ergonomics science*, 1-40

Keeble, R.J. & Macredie, R.D. (2000). Assistant agents for the world wide web intelligent interface design challenges. *Interacting with computer*, 12(4), 357-381.

Lepreux, S., Abed, M., Kolski, C., Jung, S., & Legendre, M. (2001). Vers un système d'aide à l'évaluation de la qualité des réseaux ferroviaires. Proc. of the 2ème conférence Annuelle d'Ingénierie Système, 26-28 juin, 2001, pp. 193-202, Toulouse, France.

McFarlane, D.C. & Latorella, K.A. (2002). The scope and importance of human interruption in human-computer interaction design, *Human Computer Interaction*, 17, 1-61.

Millot, P. (1999). Systèmes Homme-Machine et Automatique, JDA'99, Nancy, 21-23 septembre 1999.

Millot, P. & Lemoine, M.P. (1998). An attempt for generic concepts toward Human-Machine Cooperation. IEEE SMC'98, San Diego, USA, October.

Older, M. T., Waterson, P. E. & Clegg, C. W. (1997). A Critical Assessment of Task Allocation Methods and their Applicability. Ergonomics, 40 (2), 151-171.

Palmer M. T., Rogers W. H., Press H. N., Latorella K. A. & Abbott T.S. (1995). *A crew-centered flight deck design philosophy for high-speed civil transport (HSCT) aircraft*, NASA TM 109171, Hampton, Virginia : Langley Research Center.

Parasuraman, R. (1993). Effects of adaptive function allocation on human performance. In D.J. Garland and J.A. Wise (Eds.), *Human factors and advanced aviation technologies*, p. 147-158.

Parasuraman, R. (1997). Human and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*, 230-253.

Parasuraman, R. and Mouloua, M. (1996) *Automation and human performance: theories and applications*. Mahwah, Lawrence Erlbaum.

Rothrock, L., Koubek, R., Fuchs, F., Haas, M. & Salvendy, G. (2002). Review and reappraisal of adaptive interfaces : toward biologically-inspired paradigms, *Theoretical issues in ergonomics science*, 3(1), 47-84.

Rouse, W.B. (1988). Adaptive aiding for human/computer control, *Human Factors*, 30, 431-443.

Rouse, W.B., Geddes, N.D. and Curry, R.E. (1986). An architecture for intelligent interfaces: outline of an approach to supporting operators of complex systems, *Human Computer Interaction*, 3:87-122.

Scallen, S.F. & Hancock, P.A. (2001). Implementing adaptive function allocation, *International journal of aviation psychology*, 11(2), 197-221.

Scerbo, M.W. (1996). Theoretical perspectives on adaptive automation, In R. Parasuraman & M. Mouloua (Eds), *Human performance in automated systems : Theory and applications.* Mahwah, NJ : Lawrence Erlbaum Associates, 37-64

Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R. Di Nocero, F., and Prinzel III, (2001). *The efficacy of psychophysiological measures for implementing adaptive technology*. NASA TP-2001-211018. Hampton, VA.

Woods, D.D. (1989). The effect of automation on human's role : Experience from non-aviation industries. In S. Norman & H. Orlady (Eds.), *Flight deck automation: Promises and realities* (NASA CR-10036, pp. 61-85).

Linguistic forms as markers of interactions among decision processes

Josiane Caron-Pargue

University of Poitiers
Department of Psychology
97 Avenue du Recteur Pineau
F-86022 POITIERS CEDEX

jcaron@univ-poitiers.fr

ABSTRACT. *The verbal protocol of a 10-year-old solving four consecutive times 4-D Hanoi Towers is analysed in terms of Culioli's enunciative operations. Notably locations, and modal terms are cognitively interpreted as markers of micro-processes intervening in the reconstruction of knowledge within the situation. In this paper, modal verbs are introduced as markers of modified attentional focusing, at different levels of control of the representation distributed between internal and external spaces. Results allow to identify several on line reorganisations of more or less complex attentional focusing processes, several steps of their progressive decontextualization, distinguishing automatic vs controlled decontextualisation. They provide explanations for the generalization of constraints and for the resolution of impasses. These results may be extended to other situations, and to collaborative agents.*

RESUME: *Le protocole verbal d'un enfant de 10 ans, simultané à la résolution d'une tour de Hanoï à quatre disques, est analysé à partir des opérations énonciatives de Culioli. Notamment, les repérages et les termes modaux sont interprétés en termes de marqueurs de micro-processus intervenant dans la reconstruction des connaissances dans la situation. Dans ce papier, les verbes modaux sont les marqueurs de modifications dans la mise en perspective des informations distribuées entre les espaces interne et externe. Les résultats montrent plusieurs réorganisations on line des processus attentionnels, plusieurs étapes, automatiques vs conscientes, de leur décontextualisation. Ils fournissent des explications pour la généralisation de contraintes et pour la résolution d'impasses. Enfin, ils peuvent être généralisés à d'autres situations, notamment au problème de l'interaction entre agents.*

KEY WORDS: *problem solving, decision processes, verbal reports, modal terms, distributed representation.*

MOTS CLES: *résolution de problème, processus de décision, protocoles verbaux, termes modaux, représentation distribuée.*

1. Introduction

A number of authors agree that reorganizations take place in the actualization of knowledge, contextualizing and decontextualizing it within the situation by means of semiotization (cf. Clancey 01 for a discussion). The challenge is to characterize the micro-processes underlying these reorganizations. At the cognitive level, mechanisms of internalization and externalization may interact through distributed representations (Piaget 76; Zhang 00); at the linguistic level, the construction of contextualized utterances from oriented predicative relations (i.e. propositions to which truth values have not yet been assigned) can be modelled by enunciative operations (Culioli 95). Then, a cognitive interpretation of these operations can bring information about reorganizations of knowledge.

The purpose of this paper is to track interactions among decision processes, marked by two kinds of enunciative operations, cognitively interpreted, namely locations and modal verbs, through the analysis of the verbal protocol of a 10-year-old , solving a 4-D Hanoi Tower[1] (cf. Annex 1). Our hypothesis is that linguistic forms will allow the identification of contextual constraints and of their progressive generalizations, thus providing some insights on learning processes.

Contextual constraints have already been identified for the Hanoi Towers (Klahr *et al.*, 1981; VanLehn 91; Richard 90; Fireman 96), but the processes underlying their generalizations have not been explained. The common approach relies on literal meaning; it does not account for the construction of planning, and of the external representation, reduced to physical environment (Clancey 91).

Finally, the choice of Hanoi Towers may appear as a very specific problem. But the enunciative operations involved in our analyses transform this problem into a much more complex one. Relying on an interactive view of language, in production and in comprehension, they allow an extension of the method to other tasks. And the problem space can be decomposed among the different agents (Zhang, 1998), therefore an extension to the cases of collaborative agents can be conceived.

2. The cognitive interpretation of enunciative operations

The presence of modal terms in an utterance marks that there is a gap between the situation and the utterer. The utterer, disengaged from the situation, has to find a way, a strategy, in order to come back in it, and to construct a stabilized and asserted representation. Modal terms mark an access to other possibilities, and the

[1] *The 4-D Hanoi Tower task:* Three vertical pegs named A, B, C, are aligned from right to left on a wooden board. Four coloured disks of decreasing size stand on peg A at the initial state (pink for disc 1, the smallest one; green for disc 2; yellow for disc 3; black for disc 4, the biggest one). The goal is to move all the disks from A to C, under two constraints: only one disk must be moved at a time, and a disk may not be placed on the top of a smaller one.

introduction of modulations in order to reinforce or reorganize the utterer's primary representation (Culioli, 1995). On the cognitive level, modal terms can be interpreted as the markers of a controlled access to memory, allowing to distinguish controlled vs automatic decision processes (Caron-Pargue *et al.*, 1996, 2000). Notably, modal verbs such as *can*, *want*, *have to*, and some interrogative forms, often in the verbalization of an action, mark a detachment from the situation. They mark difficulties, and an attempt to gather non focused information for reorganizing the planning. Whereas the absence of modal term marks an activity already well known in the situation, and an automatic planning[2].

Locations between the verbalizations of two successive moves (Bégoin-Augereau *et al.*, 2007) imply automatic entailments from the locator toward the locatum (e.g. in [1], the disks-location 2=>*1*, characterized by the repetition *the green disk*; in [2], the move-location *m(3)=>m(1)*, marked by the repetition *on C*), and an attentional focusing, with the locator (respectively *2* and *m(3)*) primarily in the foreground. They provide explanations for already known constraints, e.g. in [3], the location *3=>1* explains the necessity to pass from Richard's constraint *do not jump over a peg (empty or not)* to the constraint *do not jump over a peg with a disk* (Richard 90). They constitute the elements automatically stored in memory, reused at once in similar contexts, which can simplify the constraints of memorization. Furthermore, the 9th move where the solution path leaves the optimal strategy (Welsh, 1991), e.g. the impasse *2 on 4 on C*, can be explained by the location *4=>2*.

[*1.2]* I take the green disk I put it on C (43/ 1/ 2) [1]
[1.3] so I take back the pink disk I put it on the green disk (43/ - / 21)

[1.8] I take the yellow disk I put it on C (421/ - / 3) [2]
[1.9] I take the pink disk I put it on C (42/ - / 31)

[2.10] so that I can take the green disk on column C (4/ 1/ 32) [3]
[2.11] then I take the pink disk I put it on the green disk on column C (4/ - / 321)

[2.1] so I take the pink disk in order to put it on B so that after I can put the green disk on C [4]
[2.2] so I take the green disk and I put it on C (43 / 1/ 2)
[2.3] Then I take the pink disk and I put it on the green disk (43/ - / 21)

In addition, interactions resulting from internalization and externalization, between internal vs external representational spaces (Piaget 76; Zhang 00) can be specified in terms of locations. In fact, starting terms[3] mark the distinction between internal and external decision processes (Bégoin-Augereau *et al.*, 2007), a narrowed access to abstraction, and the categorization of external occurrences, re-inscribed as contextual categories at the internal level. Thus, opposite locations can be constructed, e.g. in [4]: first *2=>1*, primarily at the external level, is categorized and reinscribed at the internal level by starting term *2*; the other *1=>2* is an internal

[2] Connectives such as *in order to*, without modal verb, mark an automatic planning.
[3] A starting term is marked by an anaphora, e.g. in [2], *the yellow disk*, referring to disk 3, replaced by the anaphora *it* in *I put it on the disc C*, is a starting term.

location[4], so that the equivalence *1<=>2* constitutes an elementary class, i.e. an internal elementary chunk of disks, having the status of condition for mental moves.

Our assumption is that an interaction between modal verbs and locations can be conceived as follows. Modal verbs will introduce a modulation within attentional focusing constructed by locations. If the modal verb bears on the locator (e.g. in [3], the modal verb *can* bears on the move of disk *2*, locator, in the disk-location *2=>1*), then, the attentional focus on *2* will be reinforced, which constructs several degrees of attentional focusing in the foreground. If the modal verb bears on the locatum (e.g. in [4], *can* bears on the move of the locatum in the move-location *m(1)=>m(2)*, and in the internal disk location *1=>2*, for which *m(1)* and *1* are focused), then it will shift the location into the opposite one (*m(2)=>m(1)* and *2=>1*), which constructs local equivalences associated to the initial focused locators (*m(1)<=>m(2)* and *1<=>2*, respectively associated to *m(1)* and *1*, for this example).

3. Results and conclusion

The detail of the modelling is presented in Annex 1. The chosen unit of analysis was the verbalization of a move. Furthermore the verbalization in terms of “take” or/and “put” (Richard 90) allows to differentiate goals, quoted as follows: *get off*, when verbalization is reduced to “take”, when the place to which the disk has to go is not verbalized; *put* when reduced to “put”; *move* when both “take” and “put” are verbalized. Besides, “general goals”, marked by the absence of naming disks and peg, are considered as generic ones, applied several times, when necessary; d_i, marked by disk *i* not involved in a location but renaming a peg, means that *i* is not directly intended in the planning, but in its background.

Trial 1. This trial shows several reconstructions of the same external locations, which mark an automatic generalisation of constraints, (a) for *2=>1* reconstructing pyramid *12*: on *C* at the beginning; on *4* still on *A* before the first move of *4*, with a reinforced focus on *2*; and on disk *4* on *B*; (b) for *m(2)=>m(1)* changing the flat position of pyramid *12* just before it is reconstructed on disk *3*; reconstructing it: on peg B, when disk 4 is already on *C*; on C for ending the puzzle.

Two ways for constructing equivalences can be shown: (a) by means of modal general goal *get off* (for *2<=>1*, in each of the above contexts; for *3<=>1* starting to move pyramid 12 for the second time, at the beginning); (b) *m(2)<=>m(1)*, equally focused on *m(1)* and on *m(2)*, for constructing pyramid *12* on *4* on *C*; it results from the articulation of m*(3)=>m(2)* with *m(1)* (giving rise to *m(3)=>m(1)* and *m(2)=>m(1)*), added to a refocalization on *m(1)*. Besides, the location *m(1)=>m(4)*, putting *m(4)* in the background of *m(1)*, and the goal d_4, explain that pyramid *123* is first moved to C, with 4 not being in attentional focusing.

---

[4] The criterion of the recognition of the internal location is the extension of the verbalization of a move, e.g. *so that I can put the green disk on C,* in [4].

Trial 2. A first feature is the construction of elementary contextual classes: (a) *move 1<=>move 2*, in [4], with external locations *2=>1* and *m(2)=>m(1)*, both categorized by starting term *2*, with a reinforced focus on *2* marked by the modal goal *move 2 on C*, both reconstructed as internal locations; (b) *get off 1<=>get off 2*, in a very similar way, in [6], with internal and external locations *2=>1*.

Another feature consists in the management of attentional focusing for reconstructing pyramid *123* as pyramid *23* (cf Annex 1, [2.4]-[2.11]). (a) The flat position of disks *2* and *3* is modified (cf. fig 1): first, by a focus bearing on *3* to move *2* and *1* in its background (fig.1a). That is marked by: the external equivalence *m(3)<=>m(1)*, associated to *m(3)* for the modal goal *put (disk 1) on B*; the external location *m(1)=>m(2)*; the internal location 3=>2. Second, by a focus bearing on *2* for moving disks *1* and *3* in its background: *1* marked by external location 2=>1, *3* by internal equivalences *2<=>1*, and *2<=>3*, associated to disk *2* for modal goals *get off 1* and *get off 3* (cf. fig. 1b). Then, the pyramid *12* is reconstructed on disk *3*, with a reinforced focus on disk *2*, marked by the modal goal *move 2 on C*, and by external locations *2=>1* and *m(2)=>m(1)*. A similar reconstruction of pyramid *123* can be seen in Annex 1 (in [2.16]-[2.19]), just before the first move of disk *4*, in which the focus on *1* is also reinforced.

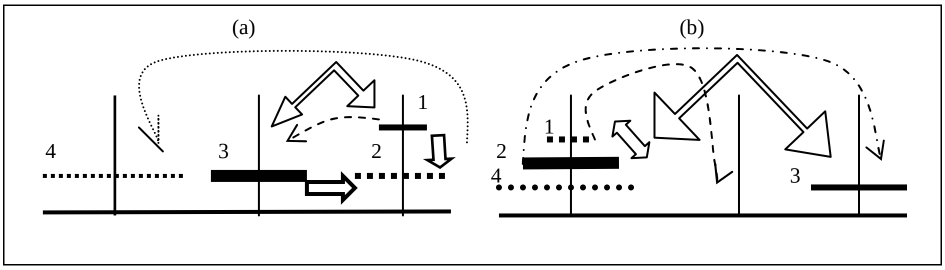


Fig.1. *Reconstruction of pyramid 123.*

Finally, modifications in attentional focusing explain the impasse *2* on *4* on *C*, and its resolution. A reinforced focus on *2*, with in its background the internal equivalence *1<=>4*, allows first to move *4* to *B*. Then, the equivalence *4<=>2*, associated to *4*, reinforced with modal goals (*get off 4* and *move 2 on 4*) constructs *2 on 4 on B*, which is not an impasse, but allows to move *4* from *B* to *C*. This focus, reasserted immediately after, marked by locations *4=>2* and *2=>1*, leads to the impasse *2 on 4 on C* (cf Annex 1, in [2.20]-[2.24]). The resolution of the impasse is done at the internal level with a return to a focus on *m(3)* entailing the internal equivalence *1<=>2*, equally associated to *1* or *2*, and the external location *2=>1*.

Trial 3: Articulations among goals *get off* and *put* within one move characterize this trial: both being modal goals (Annex 1, in [3.2], [3.31],[3.34]), or modal and automatic goals (in [3.12], [3.30]).

But a main feature concerns the quasi-achievement of the automatic move of *1* endowed with the function of directing other moves. It was already started in trial 2 with locations and equivalences, associated to *1*, and *m(1)*, notably 1<=>2, at the internal level, articulated with the automatic goal *put (1) on B* and the modal goal *move 2 on C*. The specificity of trial 3 is to extend it to the articulation among locations and modal goals: for the move of *4* to *B*, in [3.5]-[3.9], with *1<=>2* and *1<=>4*, in internal space, and modal goals *move 1 on 2* and *get off 4*. Then, in [3.12], *m(1)<=>m(3)*, associated to *m(1)*, to modal goal *get off 3*, and to automatic one *put (3) on 4*, starts a readjustment of the place to which pyramid *123* is going. Finally, in [3.32], the reconstruction of internal location *1=>2*, occurs with the external location *m(1)=>m(2)*, without any specific goal, and starts the move of pyramid *(123)* allowing disk *4* to go to *C*.

Trial 4. A simplification of attentional focusing gives place to automaticity. It is marked[5] by a decrease in the number of locations and modal goals, and an increase of the number of automatic and intended goals. Whereas a return to attentional focusing allows to avoid the impasse '*2 on 4*', marked by an accumulation of locations, and a reinforced focus on *4*, for modal goal *get off 4*.

Furthermore, a controlled generalization of constraints occurs three times, marked by interjections. The automatic move *1=>2* is articulated with immediate goals, but not yet with the main goal, as shown by several consecutive changes of the flat position of disks *1* and *2*. (a) Interjections *no* arise within the move of 1 without any interruption of the automatic move of *1* directing *2*: in [4.5], the focus is reinforced on *2*, and finally *1* is placed to let *2* on *3*, in accordance with trial *3*, in which *3<=>2*, associated to *3*, allowed *4* to go from *B* to *C*; but here, in trial *4*, disk *4* is still on *A*, and *no* marks that 4 on *B* can be avoided; in [4.7], 1<=>2, associated to *1*, is reconstructed; then *no* marks the recognition and anticipation of the wrong ongoing position of 12 on disk 4, and the break of it, entailing one more change of the flat position of 12 in [4.9]-[4.10]. (b) The interjection *oh yes*, in [4.14], marks the recognition of how to spread pyramid 12 in order to get disk 4 on peg C.

In conclusion, tracking interactions among automatic and controlled decision processes, respectively marked by enunciative operations of locations and by modal verbs (Culioli 95), allows to differentiate automatic vs controlled micro-processes of decontextualization. These processes intervene in the construction of mental moves; and could be considered as steps constructing classical rules, e.g. Vanlehn's rules[6],

[5] Number of modal goals (7, 15, 16, 4); of automatic goals (0, 11, 13, 8); of intended disks (2, 2, 2, 7); of internal locations (0,11, 5, 2); of external locations (12, 16, 20, 13), along trials.
[6] *Vanlehn's 1blk*. If the goal is to move a disk from one peg to another, and there is a single disk blocking the move, then get the blocking disk to the peg that is not involved in the move.

in order to establish links with classical approaches. The interactions among decision processes aim at structuring and reorganizing perceptual processes, by means of several complex levels of control which require further study.

These effects rely on a distinction between disk-locations and move-locations, specific to the Hanoi Towers, but may be extended to other tasks for each kind of objects and actions involved in the situation, and also between different agents. The interactions can be studied in the different problem spaces (Zhang, 1998).

5. References.

Bégoin-Augereau S., Caron-Pargue J., "Linguistic forms as markers of automatic entailments in a problem solving task", *2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology – Workshops*, Fremont CA, November 2-5 2007, p. 60-63.

Caron-Pargue J., Caron J., "Les interjections comme marqueurs du fonctionnement cognitif", *Cahiers de Praxématique*, vol. 34, 2000, p. 51-76.

Caron-Pargue J., Fièvre M.D., "Psycholinguistic analysis of a 10-years-old's verbal protocol", *Antwerp Papers in Linguistics*, vol. 85, 1996, p. 45-74.

Clancey W. J., "Situated cognition: stepping out of representational flatland", *A.I. Communications. The European Journal on Artificial Intelligence*, vol. 4, 1991, p. 109-112.

Clancey W. J., "Is abstraction a kind of idea or how conceptualization works? ", *Cognitive Science Quarterly*, vol. 1 no. 3-4, 2001, p. 389-421.

Culioli A., *Cognition and representation in linguistic theory*, Amsterdam, J. Benjamins, 1995.

Fireman G., "Developing a plan for solving a problem: a representational shift", *Cognitive Development*, vol. 11, 1996, p. 107-122.

Piaget J., *The grasp of consciousness: action and concept in the young child*, Cambridge MA, Harvard University Press, 1976, (Translation of *La prise de conscience*, Paris, PUF, 1974)

Richard J.-F., *Les activités mentales*, Paris, Armand Colin, 1990.

VanLehn K., "Rule acquisition events in the discovery of problem solving strategies", *Cognitive Science*, vol. 17, 1991, p. 497-529.

Welsh M., "Rule-guided behavior and self monitoring on the tower of Hanoi disk-transfer task", *Cognitive Development*, vol. *6*, 1991, p. 59-76.

Zhang J., "A distributed representation approach to group problem solving", *Journal of the American Society of Information Science*, vol. 49 no 9, 1998, p. 801-809.

Zhang J., "External representations in complex information processing tasks", In A. Kent (Ed.), *Encyclopedia of Library and Information Science*, vol. 68, 2000, p. 164-180, New York, Marcel Dekker, Inc.

Annex 1. ***Protocol analysis: the on line storage of information in memory.***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
Trial 1							
[1.1] -	I take the pink disk I put it on disk B #2	432 / 1 / -	<1>		p_B		
[1.2] -	I take the green disk I put it on on C	43 / 1 / 2	<2>	< 2=>1 >		2	< 2<=>1 >
[1.2bis] -	then er #4 I can take back				< get off (di) >		
[1.3] -	so I take back the pink disk I put it on the green disk	43 / - / 21	<1>,2		< get off (1) >		
[1.4.] -	I take the yellow disk I put it on B #2)	4 / 3 / 21	<3>	< 3=>1 >, < m(3)=>m(2) >		3, m(3)	< 3<=>1 >
[1.5] -	er #5 I take the pink disk I put it on the yellow disk	4 / 31 / 2	<1>,3		< get off (12) >		
[1.6] -	I take the green disk I put it on er B		<2>	< 2=>1 >		2; F(+2)	< 2<=>1 >
[1.6.bis] -	Experimenter: you are not allowed	(violation)					
[1.6.ter] -	I put it on A	42 / 31 / -					
[1.7] -	I take the pink disk I put it on the green disk	421 / 3 / -	<1>,2		1		
[1.8] -	I take the yellow disk I put it on C	421 / - / 3	<3>	< m(3)=>m(1) >	< get off (12) >	m(3)	
[1.9] -	I take the pink disk I put it on C	42 / - / 31	<1>				
[1.10] -	I take the green disk I put it on B	4 / 2 / 31	<2>	< m(2)=>m(1) >		m(2);	
[1.11] -	I take the pink disk I put it on B #3	4 / 21 / 3	<1>				
[1.12] -	er #3 I take the pink disk I put it on A #3	41 / 2 / 3	<1>		< get off (12) >		
[1.13] -	I take the green disk I put it on C	41 / - / 32	<2>	< m(2)=>m(1) >		m(2)	
[1.14] -	I take the pink disk I put it on B	4 / 1 / 32	<1>				
[1.15] -	I take the disk B I put it on disk C	4 / - / 321	<1>	< m(1)=>m(4) >	< get off (12) >	m(1)	
[1.16] -	I take the black disk I put it on B	- / 4 / 321	<4>				
[1.17] -	I take the pink disk I put it on A	1 / 4 / 32	<1>				
[1.18] -	I take the green disk I put it on the black disk	1 / 42 / 3	<2>,4	< 2=>1 >	(d_4),< get off (1) >	2	< 2<=>1 >
[1.19] -	I take the pink disk I put it on the green disk on B	- / 421 / 3	<1>,2				
[1.20] -	er #2 then I take the yellow disk I put it on A	3 / 421 / -	<3>				
[1.21] -	I take the pink disk I put it on C	3 / 42 / 1	<1>				
[1.22] -	I take the green disk I put it on on A	32 / 4 / 1	<2>				
[1.23] -	I take the pink disk I put it on B	321 / 4 / -	<1>				
[1.24] -	I take the black disk I put it on C #2	321 / - / 4	<4>				
[1.25] -	Then I take the pink disk I put it on B	32 / 1 / 4	<1>				
[1.26] -	I take the green disk I put it on C	3 / 1 / 42	<2>				
[1.27] -	er #5 I take the green disk I put it on the yellow one	32 / 1 / 4	<2>,3				
[1.28] -	I take the disk B I put it on C	32 / - / 41	<1>	< m(1)=>m(2) >			

Annex 1. ***(continued)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
[1.29] -	I take the green disk I put it on B #3	3 / 2 / 41	<2>	< m(2)=>m(1) >		m(2)	m(2)<=>m(1)
[1.30] -	I take the pink disk I put it on B	3 / 21 / 4	<1>				
[1.31] -	I take the yellow disk I put it on C	- / 21 / 43	<3>				
[1.32] -	I take the pink disk I put it on A	1 / 2 / 43	<1>				
[1.33] -	I take the green disk I put it on C	1 / - / 432	<2>	< m(2)=>m(1) >		m(2)	
[1.34] -	I take the pink disk I put it on C	- / - / 4321	<1>				
Trial 2							
[2.1] -	So I take the pink disk in order to put it on B	432 / 1 / -	<1>	<1> => <2>	put on B	<1>	<1> <=> <2>
[2.1bis]	so that after I can put the green disk on C		2		< move 2 on C >		
[2.2] -	so I take the green disk and I put it on C #3	43 / 1 / 2	<2>	< 2 => 1 >, < m(2)=>m(1) > m<1> =>m<2>		2, m(2); F(+2), F(+m(2)) m<1>; m<1> <=> m<2>	
[2.3] -	then I ... I take the pink disk and I put it on the green disk	43 / - / 21	<1>,(2)				
[2.3 bis]	on column C in order to take out the yellow disk		(3)	<1> => <3>	get off 3	<1>	
[2.4] -	I take the yellow disk I put it on column B	4 / 3 / 21	<3>	< m(3)=>m(1) >,<3> => <2>		m(3), **<3>**, m(3)<=>m(1)	
[2.5] -	I take the pink disk to put it on the black disk er #2 on disk B I want to say	4 / 31 / 2	<1>,4	< m(1)=>m(2) > <1> => <2>	put on 4 < put (1) on B >	m(1) <1>	 <1> <=> <2>
[2.6] -	then I take the green disk to put it on the black disk	42 / 31 / -	<2>,4	< 2=>1 >	put on 4	2	2<=>1
[2.6 bis]	so that after I can put the pink disk		1	<2> => <1>	< get off 1 >	<2>	<2> <=> <1>
[2.6 ter]	and so that I can put the yellow disk #2		3	<2> => <3>	< get off 3 >	<2>	<2> <=> <3>
[2.7] -	so I take the pink disk and I put it on the green disk	421 / 3 / -	<1>, 2				
[2.8] -	I take the disk er the yellow disk I put it on column C #2	421 / - / 3	<3>				
[2.9] -	then I take the pink yellow disk I put it on column B	42 / 1 / 3	<1>				
[2.10] -	so that I can put the green disk on column C #2	4 / 1 / 32	2	2=>1	< move 2 on C >	2; F(+2)	
[2.11] -	then I I take the pink disk I put it on the green disk on column C	4 / - / 321	<1>,2	m(2)=>m(1) < m(1)=>m(4) >, <1> => <4>		m(2); F(+m(2)) m(1)	
[2.11bis] -	so that I can take out the black disk		4		< get off 4 >	<1>	<1> <=> <4>
[2.12] -	I take the black disk I put it on column C #2	- / 4 / 321	<4>	< 4=>2 >		4; F(+4)	4<=>2
[2.13] -	then I take the pink disk to put it on column A #2	1 / 4 / 32	<1>		put on A		
[2.14] -	so that I can put the green disk on the black one #2	1 / 42 / 3	2,4	2=>1	< move 2 on 4 >	2; F(+2)	

Annex 1. ***(continued)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
[2.15] -	then I take the pink disk to put it to put it on the column and on the green disk	- / 421 / 3	<1>,2		put on 2		
[2.16] -	then I take the disk yellow to put it on A on column A	3 / 421 / -	<3>	< m(3)=>m(2) >	put on A	m(3)	m(3)<=>m(2)
[2.17] -	so that after I can put the pink disk on C on column C	3 / 42 / 1	<1>	< m(1)=>m(4) >	< move 1 on C >	m(1); F(+m(1))	
[2.18] -	so that I can put the green disk on column A	32 / 4 / 1	2	m(2)=>m(1)	< move 2 on A >	m(2);	m(1)<=>m(4)
[2.19] -	then I take the pink disk I put it on A	321 / 4 / -	<1>			F(+m(2))	
[2.20] -	so that I can put the black disk on C	321 / - / 4	4	4=>2	< move 4 on C >	4; F(+4)	
[2.21] -	then I take the pink disk I put it on ...the... the column	32 / 1 / 4	<1>				
[2.22] -	so that I can take out the green disk er I put it on the black one #2	3 / 1 / 42	<2>,(4)	< 2=>1 >	d_4, < get off 2 >	2;	F(+2)
[2.23] -	I take the disk I take the pink disk to put it on the green disk	3 / - / 421	<1>,(2)		put on 2		
[2.24] -	and I can take out the the yellow disk and I put it on column B	- / 3 / 421	<3>	m<3> =>m<1>	< get off 3 >	m<3>	
[2.25] -	then I take the disk er	1 / 3 / 42	1	<1> => <2>	get off 12	<1>	<1> <=> <2>
[2.25bis] -	then I take the pink disk in order to free the tower in order to remove the green disk		(<2>)		get off 2		
[2.25ter] -	where do I put it on the yellow disk #2		(3)		< put on 3 >		
[2.26] -	I take the pink disk I put it on the green one	1 / 32 / 4	<1>,(2)	<2> => <1>		<2>	<2> <=> <1>
[2.27] -	Then then I take the pink disk #4 I put it on the column	- / 32 / 41	<1>				
[2.28] -	And I take the green disk to put it on column A	2 / 3 / 41	<2>	< 2=>1 >	put on A	2	
[2.28bis] -	Then so that I can				< move (d_i) >		
2.29] -	I take the pink disk and I put it on the green one	21 / 3 / 4	<1>,(2)				
[2.30] -	And I can put the yellow disk on the column on the black disk	21 / - / 43	3,(4)		d_4, < move 3 on 4 >		
[2.31] -	Then I take the pink disk on the I put it on B	2 / 1 / 43	<1>				
[2.32] -	Then I take the green disk I put it on on the column	- / 1 / 432	<2>				
[2.33] -	And I take the pink disk and I put it on column C	- / - / 4321	<1>				

Annex 1. ***(continued)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
Trial 3							
[3.1] -	I take the pink disk I put it on column C	432 / - / 1	<1>				
[3.1bis] -	So that I can say				< move (d_i on p_i)		
[3.2] -	so that I can take the green disk	43 / 2 / 1	<2>	< m(2)=>m(1) >	< get off 2 >		
[3.2bis] -	where do I put it on column B				< put on B (on 2) >	m(2);	F(+m(2))
[3.3] -	Then I take the pink disk I put it on the disk						
	and on column B	43 / 21 / -	<1>				
[3.4] -	Then I take the disk the yellow disk						
	where do I put it on column C	4 / 21 / 3	<3>		< put on C >		
[3.5] -	Then I take I take the pink disk to put it on column A	41 / 2 / 3	<1>	<1> => <2>	put on A	<1>	<1> <=> <2>
[3.6] -	I take disk B in order to put it on column C	41 / - / 32	<2>	< m(2)=>m(4) >, <2> => <1>	put on C	m(2)	
[3.6bis] -	in order to take the pink disk		1		get off 1	<2>	<2> <=> <1>
[3.7] -	And so that I can put the pink disk on green disk	4 / - / 321	1,(2)	1 => <4>	d_2,< move 1 on 2 >	1	1<=> <4>
[3.7bis] -	so that I can take out the black disk		4		< get off 4 >		
[3.8] -	I take the black disk I put it on column B	- / 4 / 321	<4>	< 4=>1 >		4;	F(+ 4)
[3.9] -	I take I take the pink disk I put it on the black disk	- / 41 / 32	<1>	< m(1)=>m(3) >		m(1)	
[3.10] -	So that I can take the green disk and put it on column A	2 / 41 / 3	<2>	< 2=>1 >, < m(2)=>m(1) >	< get off 2 >	2, m(2);	F(+2)
[3.11] -	Then I take the the pink disk and I put it	21 / 4 / 3	<1>				
	on the green column on column A		(2)				
[3.12] -	So that I can take the yellow disk to put it on the black disk	21 / 43 / -	<3>,(4)		< get off 3 >, put on 4		
[3.13] -	#2 er then I take the pink disk I put it on column B	2 / 431 / -	<1>				
[3.14] -	I take I take the green disk to put it on column C	- / 431 / 2	<2>		put on C		
[3.15] -	Then I take I take the pink disk I put it on column A	1 / 43 / 2	<1>	< m(1)=>m(3) >		m(1);	m(1)<=>m(3)
[3.16] -	Then I take the disk						
	I'll rather take the pink disk on on column C	- / 43 / 21	1				
[3.17] -	So that I can take the yellow disk I put it on column A	3 / 4 / 21	<3>	< 3=>1 >	< move 3 on A >	3;	F(+3)
[3.18] -	I take the disk er #2 the pink disk I put it						
	on the yellow disk	31 / 4 / 2	<1>,(3)				
[3.19] -	So that I can put after the green disk on the column						
	on the black disk	31 / 42 / -	2,(4)	< 2=>1 >	d_4, < move 2 on 4 >	2;	F(+2)

Annex 1. ***(continued)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
[3.20] -	I take the disk er the pink disk to put it on column B	3 / 421 / -	<1>		put on B		
[3.20bis] –	in order to put it on the green disk		2		(put on 2)		
[3.21] -	I take the yellow disk to put it on column C	- / 421 / 3	3		put on C		
[3.22] -	Then I take then I take the pink disk to put it						
	on column A	1 / 42 / 3	<1>		put on A		
[3.23] -	Then I take the green disk to put it on column C	1 / 4 / 32	<2>	< m(2)=>m(1) >	put on C	m(2)	
[3.24] -	Then I take the pink disk to put it on column C	- / 4 / 321	<1>		put on C		
[3.25] -	Then I take the black disk I do not take the black disk						
	I take the pink disk I rather put it on column A	1 / 4 / 32	<1>,(4)	< m(1)=>m(2) >, <1> => <2>		m(1), <1>	
[3.26] -	Then I take the green disk I put it on the black disk						
	to column A	1 / 42 / 3	<2>,(4)	< 2=>1 >		2	
[3.27] -	Then I take the pink disk I put it on the green disk						
	to column B	- / 421 / 3	<1>,(2)				
[3.28] -	I take the yellow disk I put it to column A	3 / 421 / -	<3>	< 3=>2 >, < m(3)=>m(2) >		3, m(3)	3<=>2
[3.29] -	Then I take the pink disk on C	3 / 42 / 1	1				
[3.30] -	So that I can take the green disk		(2)		< get off 2 >		
	in order to put it on the yellow disk to column A	32 / 4 / 1	<2>,(3)	2=>1, m(2)=>m(1)	put on 3 to A	2, m(2);	F(+2),F(+m(2))
[3.31] -	So that after I can take back the pink disk				< get off 1 >		
	and put it on the green disk on column A	321 / 4 / -	<1>,(2)		< put on 2 on A >		
[3.32] -	Then I can take the black disk and I put it on the column	321 / - / 4	<4>		< get off 4 >		
[3.33] -	Then I take my pink disk I put it on column C	32 / - / 41	<1>				
[3.34] -	So that I can take after the green disk						
	where do I put it on the column the B	3 / 2 / 41	<2>	< 2=>1 >, < m(2)=>m(1) >	< get off 2>	2; F(+2)	
					< put on B>	m(2); F(m(2))	
[3.35] -	Then I take back my pink disk I put it						
	on the green disk on B	3 / 21 / 4	<1>,(2)				
[3.36] -	Then I take back I take my yellow disk to put it						
	on column C		<3>		put on C, d_4		
[3.36bis]-	so that there are the black one and the yellow one	- / 21 / 43	(3),(4)		put on 43		
[3.37] -	Then I take the pink disk I put it on column A	1 / 2 / 43	<1>				

Annex 1. ***(continue)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
[3.38] -	Then I take the green disk and I put it on on column C	<u>1 / - / 432</u>	<2>	< 2=>1 >, < m(2)=>m(1) >		2, m(2)	
[3.39] -	And then I take the pink disk and I put it						
	on ... on the green disk and on column C	<u>- / - / 4321</u>	<1>				
Trial 4.							
[4.1] -	So I take the pink disk I put it on column C	<u>432 / - / 1</u>	<1>				
[4.2] -	then I take the green disk I put it on column B	<u>43 / 2 / 1</u>	<2>	< 2=>1 >, < m(2)=>m(1) >		2, m(2)	
[4.3] -	then I take the column er I take the pink disk in order	<u>43 / 21 / -</u>	<1>		put on 2		
	to put it on the green disk and on column B #2		2		put on B		
[4.4] -	then I take the yellow disk in order to put it to column C	<u>4 / 21 / 3</u>	<3>		put on C		
[4.5] -	then I then I take the pink disk in order to put it on						
	no on the black disk	<u>41 / 2 / 3</u>	<1>,(4)		d_4 , put on 4		*[no]*
[4.6] -	then I take the green disk in order to put it on the disk	<u>41 / - / 32</u>	<2>	<2> => <1>	put on 3	<2>; F(+(2))	
[4.7]	so that after I can put er				< move (d_i on p_i) >		
[4.7bis]	so that after I can take back the green disk		(2)		< get off 2 >		
[4.7ter] -	no so that after I take the pink disk er I put it						*[no]*
	on column B	<u>4 / 1 / 32</u>	<1>	<1> => <2>	move 1 on B	<1>	<1> <=> <2>
[4.8] -	and then I take the ... the green disk I put it						
	on the black disk to column A	<u>42 / 1 / 3</u>	<2>,(4)		d_4		
[4.9] –	and then I take the pink disk and I put it on column C						
	with the yellow one	<u>42 / - / 31</u>	<1>,(3)		d_3		
[4.10] -	then I take the green disk and I put it on column B	<u>4 / 2 / 31</u>	<2>	< 2=>1 >, < m(2)=>m(1) >		2, m(2)	
[4.11] -	so that I take I take the pink disk I put it on er						
	on column B on the green disk	<u>4 / 21 / 3</u>	<1>,(2)		move 1 on 2 on B		
[4.12] -	then I take the yellow disk I put it back on the black disk	<u>43 / 21 / -</u>	<3>,(4)		d_4		
[4.13] -	then I ... I put er #2 I take the pink disk I put it on						
	on er column C	<u>43 / 2 / 1</u>	<1>				

Annex 1. ***(continued)***

Lines	Verbalization	States	Disks	Locations	Goals	Focalization	Units
[4.14] -	then I er #3 oh yes						*[oh yes]*
	then I take the green disk I put it on the yellow disk	432 / - / 1	<2>		d_3		
[4.15] -	after I take the pink disk and I put it on column B	432 / 1 / -	<1>				
[4.16] -	then I take the green disk I put it on column C	43 / 1 / 2	<2>	< m(2)=>m(1) >		m(2)	
[4.17] -	then I take the pink disk I put it on column C	43 / - / 21	<1>				
[4.18] -	so that I can get off the yellow disk	4 / 3 / 21	3		< get off 3 >		
[4.19] -	then I take the pink disk I put it on the black disk	41 / 3 / 2	<1>,(4)		d_4		
[4.20] -	I take the green disk in order to take it						
	on the yellow disk er#3	41 / 32 / -	<2>,(3)		d_3 , put on 3		
[4.21] -	I take the pink disk	4 / 321 / -	1				
[4.22] -	so that after I can take the black disk and I put it	- / 321 / 4	<4>	< 4=>1 >	< get off 4 >	4; F(+4)	
	to column C			< m(4)=>m(1) >		m(4)	
[4.23] -	then I take the pink disk I put it on the black one						
	to column C	- / 32 / 41	<1>,(4)	< m(1)=>m(3) >		m(1)	
[4.24] -	I take the green disk I put it to column A	2 / 3 / 41	<2>	< 2=>1 >, < m(2)=>m(1) >		2, m(2)	
[4.25] -	then I take the pink disk I put it on column A						
	on the green one	21 / 3 / 4	<1>,(2)				
[4.26] -	er then I take the disk er yellow on column C						
	with the black one	21 / - / 43	3,(4)		d_4		
[4.27] -	then I take the pink disk I put it on column B	2 / 1 / 43	<1>				
[4.28] -	then I take the green disk and I put it on column C	- / 1 / 432	<2>	< m(2)=>m(1) >		m(2)	
[4.29] -	then I take the pink disk and I put it on column C	- / - / 4321	<1>				

<i> for i varying from 1 to 4 means that i is a starting term; (i) means that disk i is not moved; x=>y is an external location and <x> =><y>an internal one; < x=>y >means that x=>y is categorized by a starting term; <goal> is a modal goal, the other goals are automatic ones. F(+i) means that the focus on i is reinforced by the modal goal.

The French version of the verbal protocol is published in the appendix of Caron-Pargue and Fièvre (1996). The translation follows the French text as closely as possible; hence, some formulations may look rather awkward in English.

“Societal Aspects of HCP” session

04:00 – 05:30 pm

Thursday, June 10, 2008

Evaluating the representativeness of parties and coalitions after German parliamentary elections 2005

Andranik Tangian

Hans Boeckler Foundation, 40476 Duesseldorf, Germany
andranik-tangian@boeckler.de

ABSTRACT: *The representativeness of five leading German parties and coalitions is estimated with regard to party manifestos and votes received at the parliamentary elections 2005. The party manifestos are converted into Yes/No answers to 95 questions (Relax the protection against dismissals? Close nuclear power plants? etc.). On each question, every party represents its adherents as well as those of the parties with the same position. Therefore, a party usually represents a larger group than that of its adherents.*

The index of popularity of a party is defined to be the percentage of the electorate represented, averaged on all the 95 questions. The index of universality of a party is the frequency of representing a majority of electors. The questions are considered either unweighted, or weighted by an expert, or weighted by the number of GOOGLE-results for given keywords (the more important the question, the more documents in the Internet). The weighting however plays a negligible role because the party answers are backed up by the party `ideology' which determines a high intra-question correlations.

The indicators show that the SPD (Social Democratic Party) is the most popular and universal party, contrary to its second rating in elections. It looks that social democratic values are predominating in Germany, but a number of their adherents, having been disappointed by the previous social democratic governance, voted either for conservators, or for the left party. However, an insufficient representativeness of conservators forced them to cooperate with social democrats and to build together the governing coalition.

KEY WORDS: *Parliamentary election, fractions, coalitions, theory of voting, indices of popularity and universality.*

Cultural Elements in Internet Software Localization

Valentina Dagiene — Tatjana Jevsikova

Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania
dagiene@ktl.mii.lt
tatjanaj@ktl.mii.lt

ABSTRACT. *Due to wide usage of Internet across the world, it is important to provide a user with the virtual environment that does not contradict to natural cultural environment he or she belongs to. Software localization is mainly referred to as language translation. However, there are many other important elements depending on country and people who will use localized software. In this paper, the main cultural elements, observed in internet software, are discussed and classified. Major attention is paid to cultural elements, which are not included in existing locale models.*

KEY WORDS: *cultural diversity, cultural aspects of software, software localization, deep localization, locale, internationalization.*

1. Introduction

Internet and its resources are widely used in various scopes of life; it becomes an important part of human culture. Internet provides electronic services for everyday use and is used to bring people together in order to create a dialog. Due to wide usage of Internet across the world, it is very important to provide user with the virtual environment that does not contradict to natural cultural environment he or she belongs to. Therefore, Internet software localization is extremely significant.

The web (World Wide Web) is usually used as a synonymous of the Internet. However, structurally the web is a convenient way to navigate through most part of the Internet resources. Internet environment can be viewed as content (websites) and software (either client-side, or server-side) that is used to access this content. The web is accessed by means of browsers, such as Mozilla Firefox, Microsoft Internet Explorer, Opera, etc. Other Internet services (email, chats, conferences, distant learning, etc.), can be also accessed using appropriate clients (they operate on the client side) or web-based applications (they operate on the server side and are accessed by the browser too). The term "Internet software" here is used as a general term to address: 1) software, used to access Internet resources, 2) web-based applications.

Software localization is mainly referred to as language translation. However, there are many other important elements, depending on language, country and people who will use localized software. Let us name them language-driven and community-driven components. In this paper, we are going to deal with the main features of localization in respect of both these groups of cultural elements.

2. Structure of Cultural Elements in Software

Culture has many different meanings. Various definitions of culture reflect differing theories for understanding, or criteria for evaluating human activity. Anthropologists and other behavioural scientists usually define culture as "the full range of learned human behaviour patterns". A cultural environment is able to provide an individual with an emotional space in which set of beliefs, values, and behaviours can be commonly shared by all the members within the same society or ethnic group (Ember, 1977). Culture provides the context in which the world is understood. It provides rules for behaviour, communication, interaction and understanding (Evers, 2001).

According to Vaske and Grantham (Vaske, Grantham, 1990), culture is: a) generally adaptive (adjustable to the particular conditions of both physical and social environments); b) mostly integrated (the elements and features which make up the culture are mostly adjusted to or consistent with each other); c) always changing (because of adapting to certain cultural events or integrating with other cultures).

There is a relation between user's culture and software usability (e.g. Qingxin Shi, 2007). We can also state that software can influence culture (this especially applies to Internet software, web-based applications and websites - a kind of software that is extremely popular in use nowadays). A software engineer should understand culture characteristics and cultural differences to create software, adaptable for many cultures.

Software localization should be considered from the beginning of the software development process. Only properly developed (i.e. internationalized) software can be localized well. Cultural elements in software play more and more important role due to increasing software migration from one cultural environment to another (translation to various languages and usage outside the origin country). The cost of software development can be considerably reduced by working internationally: the original product is developed once to be adaptable and translatable into many languages, i.e. localized to many languages. However, translation is only one part of localization. Even getting software texts appear in target language is rather interpretation, not just translation. In this paper, we will concentrate not on translation issues, but on the other part of localization: adaptation of cultural elements.

General requirements to localized products are exactly expressed in the statement that “they should look and feel as if they were developed in the country, the language and the culture they are localized for“ (Schäler, 2002). However, the question about user friendliness of localized software remains. How could it be maintained in localized software on the level of original software?

Classification of culture has been a topic of studies by Hofstede (Hofstede, 1991), Trompenaar (Trompenaar, 1997), Hall (Hall, Hall, 1990). The cultural dimensions identified by Hofstede are most cited and offer possibility to structure culture according to the five concepts: Power Distance; Individualism vs. Collectivism; Masculinity vs. Feminity; Uncertainty Avoidance and Long-term vs. Short-term Orientation. A table that assigns scores for the dimensions to each of the 72 countries that were surveyed was produced as a result of the study. Trompernaar and Hall suggest respectively seven- and four-dimensional culture model. These are categories that organize general cultural data. Speaking about software, we can look at software elements that are based on culture and cultural conventions.

Basing on the analysis of the main localization-related standards (they will be mentioned below), locale models and more than 10 year experience in software localization, we suggest classifying the most important software elements for successful software portability to different cultural environments into language-driven (those that depend on language used by the culture) and community-driven (those that depend on other cultural traditions and conventions) (Figure 1).

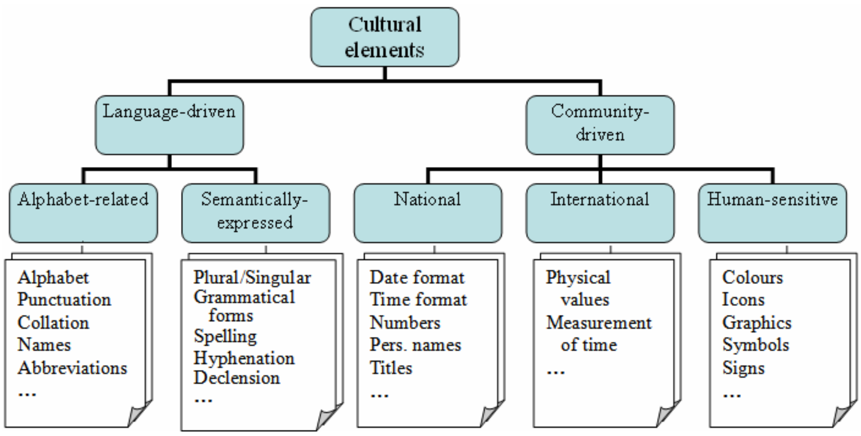


Figure 1. *Classification of software cultural elements*

The classification presented here is aimed to help researchers to evaluate the level of internationalization of the original software, check the user-friendliness of localized software, software developers to develop better-internationalized software, localizers to adapt more cultural elements to the target culture and detect internationalization bugs. We believe that it can help to pay more attention to various cultural elements in internet software and raise quality of its localizations.

Language-driven elements (see section 4) do not refer just to translation from original language to the target language. They are also about correct usage of alphabet in all parts of the software, applying correct collation rules, composition of names and abbreviations (alphabet-related issues), as well as changing forms of the words, composing one sentence from several strings, spelling support, etc. (semantically-expressed issues). Community-driven elements are based on traditions of particular community, e.g. usage of date and time formats, numbers formats, measurement system, colours, signs, etc. They may be divided into national (that are accepted on the national level), international (that are accepted on the international level) and human-sensitive (they depend on deep cultural habits).

Some language-driven and community-driven cultural elements are possible to formalise, e.g. alphabet, basic collation rules, date and time format, number format, calendar, phone number representation format, personal name representation format, etc. Some of these elements are included in a formal locale definition (see section 3). Human-sensitive elements can be considered as a reflections of a “real culture” that has been classified by Hofstede, and are difficult or impossible to formally express, e.g. icons and graphics, colour scheme, hand signals, symbols, usage of sounds and videos, historical data, privacy elements, legal data, examples.

3. Locale Data

International standard on procedures for registration of cultural elements (ISO/IEC 15897) defines locale as “the definition of the subset of a user’s information technology environment that depends on language, territory, or other cultural customs”. Locale is usually identified by the language, using two-letter language code (ISO 639-1), and by territory, using two-letter territory code (ISO 3166-1). One country may have several locales (e.g. Canada has two locales because of two national languages, English and French), and at the same time different countries speaking the same language may belong to different locales (e.g. both the USA and the UK are English speaking, but they belong to different locales).

POSIX (Portable Operating System Interface for Computer Environments) standard was one of the first to define basic locale data. Presented locale model coincides with C programming language locale model, defined in the language standard. POSIX locale model (ISO/IEC 9945-2) has six main categories, defining:

– Character classification and case conversion (uppercase and lowercase characters, digits, punctuation, spaces, character conversion and other attributes).

– Collation order (collation of character sequences, collation weights, ordering by weights, equivalence classes, one-to-many mapping).

– Monetary formatting (rules and symbols that are used to format monetary numeric information).

– Numeric, non-monetary formatting (rules and symbols that are used to format non-monetary numeric information: decimal delimiter, the symbol that shall be used as a separator for groups of digits to the left of the decimal delimiter, the size of each group of digits).

– Date and time formats (abbreviated and full weekday names, abbreviated and full month names, appropriate date and time representation, 12-hour or 24-hour clock, AM and PM strings (in the case of 12-hour clock), how years are counted and displayed for each era in a locale).

– Formats of informative and diagnostic messages and interactive responses (format and values used by various utilities for affirmative and negative responses).

These are minimum locale elements. Some new elements were added in later locale models. Set of Formal Definitions of Cultural Conventions (FDCC) defined in ISO/IEC 14652 adds new categories to POSIX and some POSIX categories are extended. New categories include: 1) format of postal addresses; 2) information on measurement system; 3) format of writing personal names; 4) format for telephone numbers and other telephone information.

International standard on procedures for registration of cultural elements (ISO/IEC 15897) specifies the procedures to be followed in preparing, publishing and maintaining a register of cultural specifications for computer use, including freeform narrative cultural elements specifications, POSIX locales, charmaps and repertoiremaps (according to ISO/IEC 9945-2), using machine-parsable form (e.g. SGML). Narrative cultural specification is used to provide more cultural information than POSIX locale enables. First six clauses of narrative specification coincide with POSIX locale categories (see above). Next clauses include additional information, e.g. national or cultural Information Technology terminology; personal naming rules; inflection; hyphenation; spelling; numbering; coding of national entities; identification of persons and organizations; electronic mail addresses; payment account numbers; keyboard layout; man-machine dialogue.

The standard mentioned above extends POSIX locale elements, adds a lot of necessary locale elements. Unfortunately, not enough locale data are collected in the repository; therefore, software developers cannot benefit from using it in their products.

By far the largest and most extensive standard repository of locale data is Unicode CLDR (Common Locale Data Repository). It provides key building blocks for software to support the world's languages. This data is used by a wide spectrum of companies for their software internationalization and localization (Unicode, 2007). The main difference from the registry described above, is that CLDR has a lot of data collected (as for 2007, more than 100 different locales): 1) dates and time formats; 2) number and currency formats; 3) measurement system; 4) collation specification (sorting, searching, matching); 5) translated names for languages, territories, scripts, timezones, and currencies; 6) script and characters used by a language. CLDR uses an XML format: Locale Data Markup Language (LDML). CLDR project provides tools as well to view and export locale data to POSIX compatible format, Java Resource Bundles, OpenOffice.org locale format.

As we can see, there is no unified way to represent locale data in a formal way. Each locale model adds additional elements to existing POSIX locale categories and extends POSIX categories. In addition to locale models mentioned here, various programming languages and environments have their own locale models (e.g. Java, PHP, GNU gettext). Locale models are not compatible with each other, usually only compatibility with POSIX locale is maintained, but POSIX has too limited set of locale elements. We could also see that formal locale data definition includes some language alphabet-related and community-driven national and international elements from the above classification. Other elements will be discussed in the next sections.

4. Language-Driven Elements

Some elements are not included in locale definition, due to deeper dependences on language differences. Unless characters, used by the language, are specified in locale definition, we can see some places in internet software, where usage of national characters is neglected. One of such areas is names of various objects: files, folders, domains, logins, persons, passwords, etc. (Jevsikova, 2006). Another area is semantically expressed language-related elements, e.g. matching of grammatical forms, composition of text strings according to the context.

4.1. Alphabets and Names

Names (identifiers) are not only used by computers, but also by humans. Names in a native language and script are easier to devise, memorize, guess, understand, manipulate, correct, and identify with that script and language (Dürst, 2003). In some older software existing requirement to use in names only 26 letters (English alphabet) is too restrictive even for most languages using Latin script, because their alphabets have some extra letters (e.g. å, š, ž, …), and some English letters are not used (mostly q, w, and x).

Login names. In many web-based applications (e.g. virtual learning environments, email clients, instant messengers, etc.) user authorization is performed. The user must provide system with his login name as one of the components of authorization data. However, most of the websites and software do not support usage of national characters in user login name: only underscores, numbers, and basic Latin alphabet letters are usually accepted. Some systems use such a name not only for internal identification, but also for addressing the user in the system. Therefore, it is natural that user should choose name, composed from his/her native language alphabet letters. Unless no technical problems to use international characters exist, developers still avoid implementing this feature in their products.

Personal names. Practically all the software allows using all letters of alphabet in personal names. However, in telecommunications many users avoid using their native alphabet and write their names with spelling errors. Our observations show that the number of incorrectly written names (avoiding letters with diacritics) of Skype users varies from 10% to 90% depending on the language. Such a great “illiteracy” may be caused by previous experience with outdated software or influence of present restriction on login names.

Passwords. Password in internet software that uses authorization usually may be composed from letters and digits. However, many programs still restrict the set of letters to ASCII characters. It is unnatural for languages that alphabets do not coincide with ASCII character set to restrict using some or all own letters and offer to use foreign letters instead. The restriction of the character set available for password reduces its security as well.

File names. Operating systems allow naming files by natural names. Letters, digits and some other textual symbols, such as dot, hyphen, etc., reasonably excluding those, which have special predefined meaning in naming the path, such as colon, slashes. Anybody may choose file names in any language reflecting their content. The same is valid for directory (folder) names.

There are some different cases using such files: 1) for storing documents in local computer, 2) exchanging documents between computers by removable storage devices, 3) sending documents as parts of e-mail messages or as their attachments, or directly by instant messengers, 4) storing web pages or other web content on a server, 5) using inside applications.

1. There are no problems for storing documents on a local computer. Thus, we may create a directory tree with names of its components from natural language vocabulary. Application, which does not accept such files, should be considered as obsolete.

2. Exchanging documents between computers by removable storage devices works well as long as the same 8-bit encoding is used in both computers. It works in nation-wide context extended to neighbouring countries, using languages with the same encoding. The Unicode would give a worldwide extension.

3. Native names of documents sent as parts of e-mail messages, as attachments, or by instant messengers, are not widely used yet because of restrictions of e-mailers of near past and misinterpreting the rules of e-mail protocols requiring that names must contain only ASCII characters. Native names before sending are encoded in

UTF-8 code by %FF sequences without non-ASCII letters and after received are decoded back. The sender and recipient of e-mail see only native names. The method works worldwide. Unfortunately, still not all programs accept native names.

4. The problem of storing web pages or other web content on a server in files with native names is theoretically solved. Native names must be encoded in the same way as e-mail attachments. However, the problems with practical realization remain.

5. File and directory names inside application, which are invisible for user, may remain not translated. However, those that are shown to user, such as containing email folders, must be localized. References to such files go from hardcoded parts of application. Dynamic file name converter is required. This is a task for software developers or localizers.

Domain names. User, working with internet software, operates with domain names. For a long time only 26 letters (English alphabet) could be used in internet domain names. In 2003, the document on using international characters in domain names was issued (Falstrom et al, 2003). According to it, the domain name (represented in Unicode) is converted to ASCII string, so called Punycode (to be compatible with existing DNS systems), and before showing it to user, it is converted to Unicode characters again. The recent versions of popular web browsers support internationalized domain names. However, many organisations still avoid using them. Possible reasons are increased phishing attacks using homographs when implementing such names worldwide and that some applications do not work with internationalized domain names, used in links.

4.2. *Semantically-Expressed Elements*

Matching of plural and singular forms. Internet (and other) software usually have dialogs where next to the number (either entered by the user, or presented by the program) an appropriate object name is shown. However, application rarely chooses the correct form of object name according to the number, for example, "1 object", "2 objects", "3 objects". In English, two forms of noun are used (plural and singular); in other languages, three or more forms are used (in Lithuanian, Polish, Russian and many other languages there are three forms: one singular and two plurals). There are languages with four (Maltese, Slovenian) and six forms (Arabic). Application should include the number and separate strings for all plural and singular forms in localizable sources and dynamically match them.

Grammatical forms. In inflective languages (e.g. Lithuanian, Finnish, Polish) names (or other words) in dialog windows may appear in various cases. It is not reasonable to require that generators of all forms for arbitrary names of all languages be included into original software. However, a placeholder for such generator should be foreseen. For example, in practically all logins to web-based system vocative case is required, e.g. English “Hello, Jonas” will be “Sveikas, Jonai” in Lithuanian.

Many problems during localization cause sentences or phrases, composed of several separate strings, as well as some parameters used inside the strings. Such a parameterization can be suitable for one language, but unacceptable for other languages, as strings order, word forms, usage of capital letters inside the sentence is different among the languages and cultures. The usage of parameter should be commented via localization comments, so that localizer could see the context.

Spellchecking is a necessary function of internet software. Spellchecking component can be used for not only preparing documents using office applications, but also to check spelling during internet chats, while filling in webpage forms, etc. As web mail and SMS sending from a webpage becomes more and more popular, checking spelling of written email message or SMS is one of webpage forms spellchecking applications. Therefore, this is a necessary component of web browser, chat client, mail client, etc. Spelling rules can differ a lot in various languages and writing systems. They can be rather difficult to implement, due to exceptions from existing formal rules. Over last years, open source spellchecking tools have been developed. Their quality is constantly increasing, and they are free. Moreover, open source spellchecking tools can be adapted freely adapted to specific culture and software needs.

Hyphenation is closely connected with line breaking, as it is usually used to find optimal place for line break. It is also connected with spellchecking due to similar grammatical and morphological analysis algorithms used for splitting into syllables as well for spellchecking. Line breaking place is detected in two steps: 1) identifying possible place for line break (it can be marked by control symbols, indicating possible break place (e.g. U+000D, U+2028 and others) or forbidding from break (e.g. U+00A0)); 2) text formatting algorithm chooses optimal line break place. Different software implements line breaking in different ways. Unicode standard provides the specification for common, suitable for various cases line breaking algorithm.

5. Implementation of Some Cultural Elements

Besides language-driven cultural elements, discussed in the previous section, internet software usually has basic national and international community-driven elements (some of them are defined in locale) (Figure 2).

Calendar. Calendar is usual component of web-based applications (e.g. content management system, blog system, virtual learning environment) and other software. Various calendars are used in the world. Gregorian calendar is widely used, but even the same calendar in different locales can be used in different ways. For instance, first day of the week and holidays can differ.

Date and time. Time measurement is based on time zone (identified as an offset from Greenwich Mean Time). The offset depends not only on locale (locale region can embrace several time zones); it is related with geographical position and can change depending of daylight saving time considerations. There are short and long date representation formats, which are different in different locales, e.g. mm/dd/yy, yyyy.mm.dd, yyyy-mm-dd for short date formats. Long date formats are more complex, year, month and day numbers are combined with language strings in different ways. Some locales use other months names cases (not nominative case) when they are used in long date format. 12- and 24-hour time formats and various hour and minutes separators (e.g. full stop, colon) are used, e.g. in the UK 12 hours format is used (e.g. 11:30 AM, 5:40 PM), in many other European locales 24 hours format is used (e.g. 19:45). Universal date and time representation, which can be used for internal date and time implementation and information exchange is defined by International standard ISO 8601.

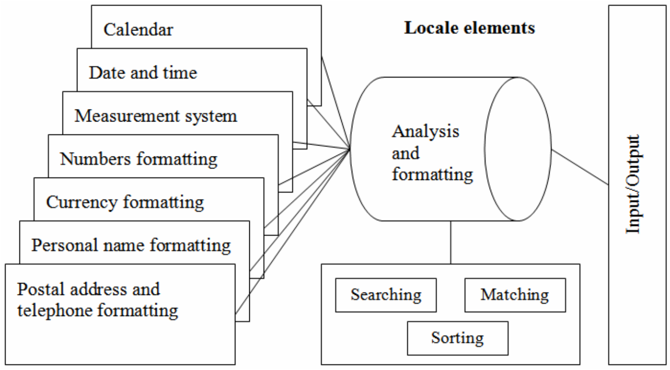


Figure 2. *Implementation of locale elements in internet software*

Measuring system indicates, which system (or systems) of measurement is used in the locale. For instance, many European locales use metric system, while in the USA U.S. custom units are still in use.

Numbers formatting. International standard defines two decimal delimiters: comma and point (full stop). In the USA, decimal separator is point, while in many European countries comma is used. Thousands (digit groups) delimiters also vary, e.g. in the USA, thousands delimiter is point, in most European countries it is comma or blank space. Incorrect decimal separator in particular language can be confused with thousand separator and lead to wrong understanding of number value.

Currency formatting. This includes national rules and symbols that are used to represent monetary amounts and currency symbols. They typically appear in web forms, online shops, etc.

Personal name formatting. Person’s first and last names are represented differently in different cultures (e.g. "First Last", "Last First", "Last, First"). In internet-based applications, such pairs of names usually appear in greeting when starting the session, as well as external identifier of user in the system, visible to other users. So, the order of name and surname should be included in the localizable part of source either as a separate parameter indicating the order, or via including parameters that have different names.

Postal address and telephone number format usually appears in internet applications, e.g. in user profile data, forms, form filler options. If telephone number is composed of several separate fields, there should be a possibility to change fields order and their number. Number of fields for postal address should be appropriate for the locale (e.g. software should not show field “State”, if locale does not have such an administrative entity).

Analysis and formatting. Information, presented in universal (numeric) format, during formatting is converted into human-readable form, suitable for target culture. During analysis process, an opposite task is performed (Figure 2). This component also provides message formatting, when locale-specific elements are included into text string. Case changes, plural and singular forms, order of words (see section 4.2.), should also be considered here.

Collation. Sorting, searching and matching are an important function of software. Sorting is related with searching, as this function is applied faster on a set of sorted elements. The collation elements can also be used for string matching, so that a proper native-language match is produced. For example, "ß" will properly match against "ss". Even one and the same locale can have several collation methods for different elements (e.g. dictionary elements sorting and telephone book elements sorting can use different rules). In non-alphabetic scripts (e.g. Eastern Asia’s ideographic scripts) collation can be phonetic or based on character representation properties. Some characters may have more than one canonical equivalents (e.g. when combination sequences are used) and their collation is context-dependant (e.g. in some locales H < Z, but CH > CZ). Unicode technical standard UTS 10 defines multi-level international collation algorithm, comparing 1) base characters; 2) accents; 3) case; 4) punctuation; 5) combining characters. This is a general algorithm, in particular locale not all collation levels are necessarily to be used, so they can be omitted and collation sequences may differ. International standard ISO 12199:2000 provides differences in alphabetical sorting of different Latin languages.

Unless all the elements discussed above are very important for true adaptation of software, we can notice that implementation of these elements is still not enough for internet software to become naturally integrated into users’ cultural environment.

6. Human-Sensitive Elements

Seeking true localization, proper attention should be paid to a composition of elements discussed above and such cultural elements in software as icons, graphics, colour scheme, hand signals, symbols, usage of sounds and videos, historical data, examples. We call them here Human-sensitive elements. They are not easy to unambiguously identify, usually not defined by national or international standards (normative documents) and depend on deep cultural habits, country or its historical unit's cultural conventions. They can also depend on individual persons and should be adaptable to person's habits. It is very difficult to systemize and express them in a formal way.

Human-sensitive cultural elements are more difficult to implement in internet software than in autonomously running software. This is due to two main reasons: they are deeply "grown" into the program, and internet software has more links with other software. Human-sensitive cultural elements in internet software should be: a) flexibly adaptable to software and other cultural components; b) flexibly fitting to each other; c) flexibly chosen by the user (multiple choices).

Some developers in order to "ease" localization process (i.e. to reduce localization costs) try to use less text and present information graphically (e.g. by usage of icons). However, observations have shown that even if graphic elements do not include text, they are interpreted differently by the users from different cultures. This is confirmed, e.g. by study of the USA, Germany and China users (Auer, Dick, 2007). Metaphors in software interface provide different associations (Evers, 2001), colours also have different meanings.

The term "culturability" has been suggested, combining the words "culture" and "usability" (Barber, Badre, 1998). The authors of this term have inspected hundreds of websites, originating in different countries and languages, and identified specific culture and genre design elements by using "cultural markers" – elements that are most prevalent, and possibly preferred within a particular cultural group, – such as colours, icons, metaphors, grouping methods, flags, sounds, fonts, shapes, links, preferences for text vs. graphics, spatial orientation, etc., to facilitate user performance. The research confirmed that the patterns are emerging and reflect cultural practices and preferences in websites, influenced both by country of origin and genre.

Kondratova and co-authors (Kondratova et al, 2005) in their research use a Cultural Web Spider, designed to extract information on culture specific webpage design elements (cultural markers) from the HTML and CSS code of websites for a particular country domain, that could help to create a cultural interface design "look and feel" prototyping tool.

Many researchers confirm an importance of the cultural dimensions, set by Hofstede and Trompenaar for developing global software and localizing internet software. For instance, during website localization, some images, browsing scheme should be changed, according to the target culture. In some cases, software functions should also be changed. Marcus and Gould (Marcus, Gould, 2001) analyse effect of cultural dimensions on websites and portals and points out possible implications of Hofstede's dimensions for UI components. Later, a grouped and ranked list of cultural dimensions that could form a decision making tool kit in a localization process was published (Marcus, 2004).

Mohd Isa and Md Noor (Wan Mohd Isa, Md Noor, 2007) summarise Hofstede's, Hall's and Trompenaar's defined cultural dimensions and create recommendations for a website navigation scheme and content presentation.

Recent research on incorporation cultural dimensions into global software includes attempts to create culturally adaptive software, applying AI mechanisms (Reinecke, Berstein, 2007). It is also proposed (Reinecke et al, 2007) to incorporate culture into a usermodel in order to implement adaptable personalization mechanisms, assigning Hofstede's value for each cultural dimension according to user's birthplace, country of current and former residence, languages, sex, age, political orientation and education level.

7. Discussion and Conclusion

Development of internationalized software and localization, so that localized product would be accepted naturally for the target users, is a very complex process. In this research, we have structured, identified and analysed the main cultural elements, which are important for deep localization of internet software, into two main categories and five subcategories. While the developed list of cultural elements is limited, we hope that it can help to pay more attention to the complex set of cultural elements while designing, localizing and testing localized or intended to localize internet software.

Existing shortcomings in software internationalization can be explained by the lack of categories included in formal locale definitions, and lack of compatibility of different locale models. Our analysis has shown that special attention during internet software development should be paid not only for a generalized set of elements, defined in existing locale models, but also to the ability to use international characters in object names (names of logins, files, domains, passwords); an ability to include a component for language's grammatical forms generation; usage of parameters in localizable strings should be reduced due to different rules of words and phrases composition in different cultures. Another trend for future work could be some formalization of human-sensitive elements, used in software.

Bibliography/References

Auer S., Dick E. "When does a difference make a difference? A snapshot on global icon comprehensibility". In: Jacko, J.A. (Ed.) *Human-computer interaction, Pt 2, Proc. of 12th Intern. Conf. on Human-Computer Interaction. LNCS*, 4551, 2007, p. 3-12.

Barber W., Badre A. "Culturability: The Merging of Culture and Usability". *Our Global Community. Conference Proceedings*, 1998.

Dürst M.J. “Internationalization of XML – Past, Present, Future”. *XML Conference&Exposition. Pensylvania, December 7-12*, 2003.

Ember C.R., Ember M. *Anthropology*. Prentice-Hall, Englewood Cliffs, NJ, 1977.

Evers V. “Cross-Cultural Understanding of Graphical Elements on the DirectED Website”. In Smith, A (Ed) *Proc. of Annual Workshop on Cultural Issues on HCI*, Putteridge Bury, University of Luton, 2001.

Evers V. Cultural aspects of user interface understanding. Doctoral dissertation. Institute of Educational Technology, The Open University, 2001.

Faltstrom P., Hoffman P., Costello, A. Internationalizing Domain Names in Applications. Network Working Group, Request for Comments: 3490, 2003.

Hall E., Hall M. *Understanding Culture Differences*. Maine, Intercultural Press, 1990.

Hofstede G. *Cultures and Organization*. McGraw Hill, New York, 1991.

Jevsikova T. “Internationalization and Localization of Web-based Learning Environment”. In: R. Mittermeir (Ed.) *Informatics Education – the Bridge Between Using and Understanding Computers. LNCS*, 4226, 2006, p. 310-319.

Kondratova I., Goldfarb I., Gervais R., Fournier, L. “Culturally Appropriate Web Interface Design: Web Crawler Study”. *Proceedings of the 8th IASTED CATE Conference*, 2005.

Marcus A. A Practical Set of Culture Dimensions for Global User-Interface Development. AM+A, 2004.

Marcus A., Gould E.W. Cultural dimensions and global Web design. AM+A, 2001.

Qingxin Shi. “Cultural Usability: The Effects of Culture on Usability Testing”. *Human-Computer Interaction – INTERACT 2007*, *LNCS*, 4663, 2007, p. 611-616.

Reinecke K., Bernstein A. “Culturally Adaptive Software: Moving Beyond Internationalization”. In: Aykin, N. (Ed.) *Usability and Internationalization, pt II, proc. Global and Local User Interfaces. LNCS*, 4560, 2007, p. 201-210.

Reinecke K., Reif G., Bernstein A. “Cultural User Modeling With CUMO: An Approach to Overcome the Personalization Bootstrapping Problem”. *Cultural Heritage on Semantic Web Workshop of ISWC*, 2007, p. 83-89.

Schäler R. “The cultural dimensions in software localization”. *Localisation Focus*, vol. 1, isssue 2, 2002, Localization Research Centre, Limerick.

Trompenaars F. *Riding the waves of culture – Understanding cultural diversity in business*. Nicholas Brealey Publ., London, 1997.

Unicode Inc. *Unicode CLDR Project: Common Locale Data Repository*, 2007.

Vaske J., Grantham C.E. *Socializing the Human-computer Environment*. Ablex, Norwood, NJ, 1990.

Wan Mohd Isa W. A. R., Md Noor N. L. “Incorporating the cultural dimensions into the theoretical framework of website information architecture”. In: Aykin, N. (Ed.) *Usability and Internationalization, pt 1, proc. Global and Local User Interfaces. LNCS*, 4559, 2007, p. 212-221.

Contributors Index

TELECOM Bretagne ISBN: 978-2-908849-22-6

HCP-2008 – Human Centered Processes

Part I: Main Conference

Today the human world and the digital information world form a complex changing environment. Any combination of human actors and software systems must behave like a process able to maintain its integrity, efficiency and adaptability in such dynamical environments. Humans engage in interaction and cooperation in many different contexts that often coincide in the same space-time interval. We need a better understanding and implementation of effective solutions to support humans in their adaptive, interactive and cooperative processes. The dynamics of cognitive processes and cross-learning are of much interest, especially when decision makers are associated with decision support systems: the advent of a paradigm for human-machine sustainability.

The papers met in this volume correspond to the communications done during the Third International Conference on Human Centered Processes (HCP-2008), which held in Delft (The Netherlands) during June 2008. HCP-08 is in line with the series of Human Centered Processes conferences held previously at Brest (France, 1999) and in Luxembourg (2003) under the umbrella of the EURO Working Group Human Centered Processes. HCP-2008 is a platform with a plenary interdisciplinary conference and workshops in order to foster a shared interest across disciplines and stimulate within-discipline progress and information exchange.

The third conference (2008) focuses on the human actor and software agent collaboration in safety and time critical systems-of-systems. This initial focus included topics ranging from: architectural requirements for such systems, cognition-based approaches for awareness and decision support, context awareness and the effects on learning, reasoning and decision making in both humans and software agents, the interaction and socially inspired process models for such collaborating systems and techniques to support them, the design, build and testing of such systems, what design patterns are adequate, what methods and tools to use.

This third HCP conference was organized with a three-day conference covering most of themes of interest for the community, and three workshops on more specific HCP themes, namely, the modeling and reasoning in context, the ethical aspects of HCP in decision making, and the problem of supervisory control and the question of relationship between automation and human operators. The papers carefully selected for the main conference cover most of these themes. Three keynotes speakers, namely J.J. Meyer (Universiteit Utrecht, The Netherlands), R.R. Hoffman (IHMC, USA) and D.L. Olson (University of Nebraska-Lincoln, USA) have accepted to share their experience with participants in the context of the HCP-2008 conference.

Patrick Brézillon, Conference Chair
Gilles Coppin, Steering Committee Chair
Philippe Lenca, Workshop General Chair

ISBN: 978-2-908849-22-6