Apprentissage automatique et acquisition des connaissances : deux approches complémentaires pour les systèmes à base de connaissances. Application au système "ACASYA" d'aide à la certification des systèmes de transport automatisés

Habib Hadj-Mabrouk

Mémoire

Présenté à l'Université Polytechnique Hauts-de-France - Université de Valenciennes

Pour l'obtention du :

Doctorat en automatique et informatique des systèmes industriels et humains.

Par :

Habib HADJ-MABROUK

APPRENTISSAGE AUTOMATIQUE ET ACQUISITION DES CONNAISSANCES : DEUX APPROCHES COMPLEMENTAIRES POUR LES SYSTEMES A BASE DE CONNAISSANCES.

Application au système « ACASYA » d'aide à la certification des systèmes de transport automatisés.

Soutenu le 15 décembre 1992 devant la commission d'examen :

- Jean-Gabriel GANASCIA, Professeur, Université Pierre et Marie Curie - Paris VI (Rapporteur)
- Marcel STAROSWIECKI, Professeur à l'université de Lille I (Rapporteur)
- Jean DEFRENNE, Professeur à l'université de Valenciennes (Examinateur)
- Bao LE TRUNG, Directeur de recherche à l'INRETS (Examinateur)
- Bernard HOURIEZ, Professeur à l'université de Valenciennes (Co-directeur thèse)
- Patrick MILLOT, Professeur à l'université de Valenciennes (Directeur de thèse)

AVANT PROPOS

Je remercie vivement Monsieur Marcel STAROSWIECKI, Professeur à l'université de Lille I, de me faire l'honneur d'être rapporteur de ma thèse.

Je tiens à exprimer ma profonde reconnaissance à Monsieur Jean-Gabriel GANASCIA, Professeur à l'université de Paris VI, également rapporteur de cette thèse, qui m'a entr'ouvert le cercle fermé des initiés à l'apprentissage automatique.

Mes remerciements vont également à Messieurs Jean DEFRENNE, Professeur à l'université de Valenciennes et Bao LE TRUNG, Directeur de recherche à l'INRETS d'Arcueil pour l'intérêt qu'ils portent à ce travail en acceptant de participer à mon jury de thèse.

Je remercie également Monsieur le Professeur Patrick MILLOT, Directeur de l'équipe Informatique Industrielle et Communication Homme-Machine ainsi que Monsieur Bernard HOURIEZ, Maitre de conférences à l'université de Valenciennes, pour les conseils qu'ils m'ont prodigués, la confiance qu'ils m'ont témoignée tout au long de la réalisation de ce travail et le soin extrême qu'ils ont porté à la correction de ce mémoire.

Mes remerciements s'adressent aussi aux membres de l'INRETS (Institut National de Recherche sur les Transports et leur Sécurité) et du CRESTA (Centre de Recherche et d'Evaluation des Systèmes de Transport Automatisés) et notamment à :

- Monsieur Yves DAVID, Directeur du CRESTA de Lille, pour sa sympathie, sa présence et ses encouragements constants.
- Monsieur El-Miloudi EL-KOURSI, Chargé de recherche et expert de certification au CRESTA de Lille, pour sa précieuse collaboration.

J'associe à ces remerciements :

- Les membres du laboratoire de mathématiques de l'université de Valenciennes et tout particulièrement Monsieur Jalel TABKA, pour son soutien et ses conseils éclairés.
- Toute l'équipe Informatique Industrielle et Communication Homme-Machine pour sa présence amicale.
- Madame Isabelle FLAMME pour sa participation à la réalisation pratique de ce mémoire.

Enfin, j'adresse mes remerciements les plus vifs à Claudine PAMPIN pour ses conseils judicieux, le soin qu'elle a apporté à la lecture et la correction de ce mémoire ainsi que pour son indispensable soutien moral et la patience dont elle a su faire preuve.

RESUME :

Ce mémoire présente une contribution à l'amélioration des méthodes usuelles d'analyse de sécurité employées dans le cadre de la certification des systèmes de transport automatisés (STA). La mission des experts de certification consiste à apprécier le caractère sécuritaire d'un nouveau STA en évaluant la complétude des scénarios d'accidents envisagés dans l'étude de sécurité du constructeur. La méthodologie d'aide à la certification développée repose sur l'utilisation conjointe et complémentaire de l'acquisition des connaissances et de l'apprentissage automatique. La méthode d'acquisition de connaissances choisie s'est révélée efficace pour extraire et formaliser les connaissances historiques d'analyse de sécurité mais insuffisante pour acquérir en détail la démarche experte de certification qui est fortement intuitive et évolutive. Pour pallier cette lacune, notre étude s'est orientée vers l'utilisation des techniques d'apprentissage automatique. La difficulté de définir et choisir un système d'apprentissage adapté aux exigences d'une application industrielle nous a conduits à proposer une caractérisation du processus d'apprentissage.

ACASYA est l'environnement logiciel développé pour supporter la méthodologie d'aide à la certification. Il est composé de deux modules principaux : CLASCA et EVALSCA, respectivement dédiés à la classification et à l'évaluation des scénarios d'accidents. CLASCA que nous avons entièrement conçu est un système d'apprentissage symbolique-numérique, inductif, incrémental, non monotone et interactif. EVALSCA, développé autour du système d'apprentissage de règles CHARADE, a pour objectif de suggérer aux experts de certification d'éventuelles pannes non considérées par le constructeur et susceptibles de mettre en défaut la sécurité d'un nouveau STA. A ce jour, ACASYA a prouvé l'intérêt de la méthodologie pour formaliser, exploiter et pérenniser le savoir faire de l'expert de certification, en vue de tendre vers l'exhaustivité de l'analyse de sécurité. Par opposition aux systèmes d'aide au diagnostic, ACASYA peut être perçu comme un outil d'aide à la prévention des pannes, situé en aval des méthodes prévisionnelles classiques d'analyse de sécurité.

Mots clés :

Apprentissage automatique, acquisition de connaissances, système à base de connaissances, système de transport automatisé, certification, sécurité, classification, évaluation, scénarios d'accidents, prévention des pannes.

ABSTRACT

This dissertation presents a contribution to the improvement of the usual methods of safety analysis used within the framework of the certification of automated transport systems (ATS). The mission of the certification experts consists in assessing the safety character of a new ATSby evaluating the completeness of the accident scenarios envisaged in the manufacturer's safety study. The certification assistance methodology developed is based on the joint and complementary use of knowledge acquisition and machine learning. The method of knowledge acquisition chosen has proved effective in extracting and formalizing historical knowledge from safety analysis but insufficient to acquire in detail the expert certification approach which is highly intuitive and scalable. To fill this gap, our study focused on the use of machine learning techniques. The difficulty of defining and choosing a learning system adapted to the requirements of an industrial application led us to propose a characterization of the learning process.

ACASYA is the software environment developed to support the certification assistance methodology. It is made up of two main modules: CLASCA and EVALSCA, respectively dedicated to the classification and evaluation of accident scenarios. CLASCA which we have entirely designed is a symbolic-digital, inductive, incremental, non-monotonous and interactive learning system. EVALSCA, developed around the CHARADE rules learning system, aims to suggest to certification experts possible failures not considered by the manufacturer and likely to jeopardize the safety of a new ATS. To date, ACASYA has proven the interest of the methodology to formalize, exploit and perpetuate the know-how of the certification expert, in order to tend towards the exhaustiveness of the security analysis. In contrast to diagnostic assistance systems, ACASYA can be seen as a tool to help prevent breakdowns, located downstream of conventional forecasting methods of safety analysis.

Keywords:

Machine learning, knowledge acquisition, knowledge-based system, automated transportation system, certification, safety, classification, assessment, accident scenarios, failure prevention.

TABLE DES MATIERES

<u>CHAPITRE 1</u> : CERTIFICATION ET SECURITE DES SYSTEMES DE TRANSPORT TERRESTRES AUTOMATISES (STA)

INTRODUCTION GENERALE

Le processus intellectuel par lequel un opérateur humain évalue une situation, prédit un événement ou prend une décision est souvent difficilement modélisable sous forme d'algorithmes sûrs et définitifs. Cette difficulté peut être partiellement contournée en faisant appel aux techniques d'Intelligence Artificielle (IA). En effet, leur développement considérable a permis de pallier l'insuffisance de l'informatique traditionnelle. L'IA s'est donnée pour but l'étude et la simulation des activités intellectuelles humaines. Elle s'efforce de créer des machines capables d'un comportement "intelligent" et a pour vocation ambitieuse de doter l'ordinateur de quelques unes des facultés de l'esprit humain : apprendre, reconnaître, raisonner ... L'aptitude à comprendre le langage naturel et à raisonner est la clé de voûte de l'intelligence.

Les premiers résultats des recherches menées dans cette voie concernent les systèmes experts (SE) ou Systèmes à Base de Connaissances (SBC). Ils sont apparus comme des outils intelligents capables de reproduire certaines tâches intellectuelles habituellement accomplies par des experts. La capacité d'exploiter et surtout de pérenniser l'expérience de ces derniers confère aux SBC un pouvoir d'information et de décision en vue de guider les opérateurs humains non spécialistes. Depuis quelques années, les SBC font une entrée remarquée dans l'industrie, ils ne sont plus considérés comme des objets rares issus des laboratoires de recherche. Cependant ils atteignent très rarement les performances de l'expert humain et sont souvent mal adaptés aux besoins réels des utilisateurs finaux. Ceci est dû aux difficultés d'extraire l'expertise nécessaire auprès d'un ou plusieurs experts du domaine et de représenter ce savoir sans distorsion pour aboutir à un modèle cognitif de l'expert.

A cela s'ajoute une phase cruciale de remplissage manuel de la Base de Connaissances (BC) du SBC. "Capturer" la connaissance pour la conserver dans la BC d'un système expert est une tâche complexe et coûteuse en temps et souvent en moyens matériels et humains. Plusieurs travaux de recherche ont évoqué le problème du recueil et de la formalisation des connaissances manipulées par l'expert en situation de résolution de problème. L'expert peut avoir de grandes difficultés à décrire explicitement les étapes du raisonnement qu'il utilise pour prendre des décisions. Ceci nécessite de sa part un long travail de réflexion qui lui permettra d'expliquer la part inconsciente de sa démarche. De cette tâche difficile et parfois même douloureuse dépend le succès d'un projet de SBC. Or, à l'instar de nombreux auteurs, nous considérons cette tâche comme le "goulot d'étranglement" rencontré dans le développement d'un SBC.

En effet, compte tenu de la complexité des connaissances de l'expert et de la difficulté de ce dernier à expliciter ses processus mentaux, la connaissance extraite risque souvent d'être inexacte, incomplète voire incohérente. Diverses recherches en IA sont donc menées pour appréhender ce problème de transfert d'expertise. On perçoit aujourd'hui deux grandes activités de recherche indépendantes :

- l'acquisition de connaissances dont le but est de définir des moyens et des méthodes inspirés notamment du génie logiciel et de la psychologie cognitive pour mieux cerner le transfert d'expertise,
- l'apprentissage automatique qui propose la mise en œuvre de techniques inductives, déductives, abductives ou par analogie permettant de doter le SBC de capacités d'apprentissage.

Le développement d'un SBC d'aide à la certification nous a incité à utiliser de façon conjointe et complémentaire ces deux approches pour appréhender le domaine de la certification. Ce travail, effectué au Laboratoire d'Automatique Industrielle et Humaine (LAIH) dans l'équipe Informatique Industrielle et Communication Homme-Machine, s'appuie sur une collaboration avec l'Institut National de Recherche sur les Transports et leur Sécurité (INRETS) et le Centre de Recherche et d'Evaluation des Systèmes de Transports Automatisés (CRESTA). Cette étude s'inscrit dans le cadre du Groupement Régional Nord - Pas de Calais pour la recherche dans les Transports (GRRT) et est soutenue par une bourse de Doctorat cofinancée par l'INRETS et la Région Nord - Pas de Calais.

L'une des activités de recherche menées actuellement par l'INRETS-CRESTA s'intéresse à la certification des systèmes de transport à conduite automatique et à la sécurité des systèmes à commande numérique. Notre étude s'inscrit dans ce contexte et vise à concevoir et réaliser un outil logiciel d'aide à la certification des systèmes de transport automatisés (STA) afin d'apprécier l'adéquation des équipements de protection proposés. L'objectif de cet outil est, d'une part d'évaluer la complétude et la cohérence des scénarios d'accidents proposés par les constructeurs et d'autre part de générer de nouveaux scénarios susceptibles d'aider les experts de certification de l'INRETS-CRESTA qui doivent conclure sur le caractère sécuritaire d'un nouveau STA.

Dans ce mémoire, on se propose d'identifier les problèmes inhérents au domaine de la certification des systèmes de transport terrestres automatisés et on présente une approche méthodologique d'analyse et d'évaluation de la sécurité pour aider à la certification. Cette méthodologie repose sur l'emploi interactif des techniques d'acquisition de connaissances et d'apprentissage symbolique-numérique. Ces travaux sont présentés en cinq chapitres.

- Le premier chapitre présente le domaine de la certification et justifie le recours aux techniques d'Intelligence Artificielle pour réaliser un outil d'aide à la certification. Ce chapitre s'efforce de cerner le processus de certification et fait une synthèse des différentes méthodes impliquées dans ce processus. A l'issue de ce bilan, il conclut sur la nécessité de recourir à des outils logiciels "intelligents" et notamment aux systèmes à base de connaissances pour aider les experts certifieurs à accomplir leur mission.

- Le second chapitre présente les techniques, méthodes et outils issus des recherches récentes sur l'acquisition de connaissances pour transférer la connaissance détenue par les experts vers la base de connaissances du système. Cet état de l'art sur l'acquisition de connaissances débouche sur le choix d'une méthode appliquée au domaine de la certification et fondée sur les travaux de BENKIRANE /91/. Compte tenu de l'originalité et de la complexité de ce domaine ainsi que du mode de raisonnement des experts certifieurs qui est de nature intuitive, évolutive et créative, la méthode d'acquisition choisie ne permet pas d'accéder à la totalité des connaissances de certification et notamment à celles relatives au détail de la démarche experte.

- Pour ces raisons, le troisième chapitre appréhende l'apprentissage automatique comme un moyen permettant de combler les insuffisances de l'acquisition de connaissances. Il présente un état de l'art

qui débouche sur la caractérisation du processus d'apprentissage. Cette caractérisation permet de définir les propriétés requises par l'outil d'aide à la certification et aboutit au choix du système d'apprentissage CHARADE ainsi qu'à la nécessité de developper un nouveau système dédié à la classification et baptisé CLASCA.

- L'expérience enrichissante de cette application industrielle guidée par l'état de l'art en matière d'acquisition de connaissances et d'apprentissage automatique nous permet de proposer une méthodologie d'aide à la certification. Cette méthodologie s'articule autour de trois modules principaux : CLASCA, EVALSCA ET GENESCA, dédiés respectivement à la classification, à l'évaluation et à la génération des scénarios d'accidents. Le quatrième chapitre présente le système ACASYA d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés regroupant ces trois modules. Il détaille la conception et la réalisation des deux premiers modules CLASCA et EVALSCA.

- Le dernier chapitre présente, à partir d'un exemple d'application, l'évaluation globale de la maquette de faisabilité d'ACASYA puis une évaluation interne spécifique au module CLASCA. Il dresse un bilan des limites, intérêts et extensions d'ACASYA et de CLASCA. Enfin, nous proposons en perspectives le fruit des réflexions induites par le développement de ce projet, concernant la complémentarité de l'acquisition des connaissances et de l'apprentissage automatique pour améliorer le processus de transfert d'expertise.

Chapitre 1 :

La certification et la sécurité des systèmes de transport terrestres automatisés (STA)

INTRODUCTION

La certification est la reconnaissance par un service officiel qu'une fonction, un équipement, un système est conforme à un ensemble de règlements nationaux ou internationaux (AFNOR Z61-102 10). Dans les transports terrestres automatisés, on se réfère à l'article 31 bis du décret de mars 1942, qui stipule que le dispositif de protection ainsi que les règlements d'exploitation doivent être agrées par le ministre chargé des transports /EL KOURSI 92/. La certification des Systèmes de Transport Automatisés (STA) mis en service en France est assurée par L'INRETS-CRESTA. Ce centre a acquis une expertise certaine dans ce domaine après avoir participé à la certification d'une demi-douzaine de STA. L'une des étapes essentielles d'une opération de certification consiste à examiner dans le détail, sur le plan de la sécurité, le système conçu par le constructeur. Ce processus complexe d'analyse et d'évaluation consiste à vérifier que le système satisfait bien aux exigences du cahier des charges défini en liaison avec l'exploitant et en particulier qu'il ne peut en aucun cas être une source de risque pour les usagers.

La certification des STA repose en grande partie sur une analyse profonde des risques résultant des scénarios d'accident. Un scénario d'accident est une succession adéquate d'évènements pouvant être à l'origine d'un risque potentiel. Pour assurer la sûreté de fonctionnement du STA et par conséquent garantir la sécurité des voyageurs, il est indispensable d'impliquer les experts certifieurs dès la phase de conception du système. La mission de l'expert de certification est de juger et d'évaluer la complétude et la consistance du dossier de sécurité proposé par le constructeur. En final, le certifieur doit prouver le caractère sécuritaire d'un nouveau STA en démontrant que les critères et fonctions de sécurité sont bien respectés. Dans ce processus de certification, une difficulté majeure consiste à trouver les scénarios anormaux pouvant conduire à un risque particulier. C'est le point fondamental qui a motivé le présent travail.

Dans ce premier chapitre, on se propose d'identifier les problèmes inhérents au domaine de la certification des STA et on présente les approches adoptées pour aider à l'analyse de sécurité et par conséquent à la certification. Quatre parties sont successivement abordées.

La première partie présente les concepts de base de la sûreté de fonctionnement des systèmes et notamment ceux relatifs à la sécurité sur lesquels repose la certification. Le développement d'un nouveau STA requiert la définition d'objectifs quantifiés de sécurité en référence aux systèmes de transport existants et réputés sûrs. Afin de maintenir et respecter ce niveau de sécurité, deux approches de sécurité existent et sont employées par les constructeurs de STA : la sécurité "intrinsèque" et la sécurité "probabiliste".

Juger de la consistance de ces deux approches constitue une des tâches principales des experts certifieurs. La deuxième partie de ce chapitre s'efforce de cerner l'activité de ces experts tout au long du cycle de vie du STA.

Pour évaluer et juger le degré de sécurité d'un nouveau STA, les experts certifieurs doivent procéder à une double analyse, fonctionnelle et qualitative, de la sécurité du STA. La troisième partie établit un bilan des différentes méthodes impliquées dans le processus de certification. Ces méthodes prévisionnelles d'analyse de sécurité, bien que nécessaires pour pallier les défaillances critiques, ne sont pas suffisantes. En effet,

l'imagination de nouveaux scénarios d'accident par les experts certifieurs demeure indispensable pour tendre vers l'exhaustivité de l'analyse de sécurité.

A partir de cette identification du domaine de certification, nous proposons dans la dernière partie une spécification des propriétés requises par un système d'aide à la certification et à l'analyse de sécurité des STA ainsi que l'approche retenue pour résoudre ce problème. Celle-ci repose sur l'emploi interactif des techniques d'acquisition de connaissances et d'apprentissage automatique.

1. SURETE DE FONCTIONNEMENT DANS LES TRANSPORTS TERRESTRES AUTOMATISES

La sûreté de fonctionnement est une préoccupation permanente dans la plupart des secteurs industriels, elle est d'autant plus essentielle dans le secteur des transports terrestres automatisés en raison de la gravité des accidents et de leur impact dans le public. L'automatisation intégrale des systèmes de transport tels que le VAL (Véhicule Automatique Léger) de Lille et plus récemment le VAL d'Orly rend encore plus critique la notion de sûreté de fonctionnement. Dans ce contexte, la connaissance des défaillances, leur évaluation, leur mesure et leur maîtrise, basées à l'origine sur l'expérience acquise doivent désormais intégrer l'aspect prévention. Ainsi pour comprendre et quantifier le comportement du système de transport par rapport aux défaillances de ses composants et à leurs remèdes, les constructeurs et les responsables de la certification doivent recourir aux concepts et aux techniques de la sûreté de fonctionnement. De façon à se conformer aux exigences des pouvoirs publics et de l'exploitant, cette démarche procède depuis le stade de l'avant projet jusqu'à la réalisation et l'exploitation du système de transport.

1.1. Le concept de base de la sûreté de fonctionnement

La Sûreté de fonctionnement d'un système est la propriété qui permet à ses utilisateurs de placer une confiance justifiée dans le service qu'il leur délivre /LAPRIE 85, 87/. Au sens large, la sûreté de fonctionnement est considérée comme la science des défaillances. Elle est caractérisée dans /VILLEMEUR 88/ par les concepts suivants : la fiabilité (Reliability), la maintenabilité (Maintenability), la disponibilité (Availability) et la sécurité (Safety).

1.1.1. La fiabilité

La fiabilité est l'aptitude d'un dispositif à accomplir une fonction requise, dans des conditions données, pendant une durée donnée.

1.1.2. La maintenabilité

La maintenabilité caractérise l'aptitude d'un système à reprendre l'accomplissement de sa fonction après une défaillance, lorsque la maintenance est accomplie dans des conditions données avec des procédures et des moyens prescrits. De la fiabilité et de la maintenabilité on déduit la disponibilité.

1.1.3. la disponibilité

La disponibilité est la probabilité qu'un matériel soit apte à fonctionner à un instant donné. En matière de systèmes de transport urbain, tels qu'une ligne de métro, on parle souvent de disponibilité en exploitation /FAYOLLE 90/.

$$\text{Disponibilité pour l'exploitation} = \frac{\text{temps d'exploitation} - \Sigma \text{ temps de perturbations}}{\text{temps d'exploitation}}$$

En effet, pour les usagers, la disponibilité apparente est celle calculée pendant l'exploitation tandis que le temps de réparation (MTTR) est le temps passé pour restaurer le service après une interruption de trafic (c'est le temps pendant lequel l'exploitation est perturbée). La redondance d'équipement est communément employée pour améliorer la disponibilité du système de transport, notamment pour des équipements dont la moindre défaillance provoque une perturbation significative pour l'exploitation. Elle nécessite le doublement des équipements, l'un étant actif et l'autre passif, c'est à dire en attente et prêt à prendre le relais dès que le premier n'assure plus tout ou partie des fonctions requises.

1.1.4. La sécurité

La sécurité est l'aptitude d'une entité à éviter de faire apparaître, dans des conditions données, des événements critiques ou catastrophiques. L'aptitude contraire sera dénommée "insécurité". La sécurité est mesurée par le temps moyen entre défaillances catastrophiques (MTBUF). Les états d'insécurité sont généralement les pannes non déclarées et les commandes intempestives.

Les attributs de la sûreté de fonctionnement peuvent être groupés en trois classes /LAPRIE 87/ :

- les entraves à la sûreté de fonctionnement qui sont des circonstances indésirables, causes ou résultats de la non-sûreté de fonctionnement,
- les moyens pour la sûreté de fonctionnement qui sont les méthodes, outils et solutions permettant de procurer au système l'aptitude à délivrer un service de confiance,
- les mesures de la sûreté de fonctionnement qui permettent d'apprécier la qualité du service délivré.

Une vision plus complète de la sûreté de fonctionnement est décrite dans /LAPRIE 88, 89/. L'auteur considère la sûreté de fonctionnement selon différents points de vue :

- par rapport à la continuité du service, la sûreté de fonctionnement est perçue comme la fiabilité,
- par rapport à la non-occurrence de défaillances catastrophiques, la sûreté de fonctionnement est perçue comme la sécurité,
- par rapport à la préservation de la confidentialité et de l'intégrité des informations, la sûreté de fonctionnement est perçue aussi comme la sécurité.

Cette analyse révèle que la **sécurité** est un facteur fondamental de la sûreté de fonctionnement des systèmes qui oblige le constructeur d'un nouveau système de transport à conduite automatique (STA) à respecter l'ensemble des concepts relatifs à la sûreté de fonctionnement (fiabilité, maintenabilité, disponibilité et sécurité). De plus, dans le domaine du transport, la certification doit s'effectuer compte tenu du caractère automatique d'un véhicule qui constitue un facteur supplémentaire influant sur la sécurité des voyageurs. Le constat suivant relatif au Métro de Lille, tiré de /MARCOVICI 82/, illustre bien l'importance accordée à la notion de sécurité :

> *"L'analyse critique de la conception sous l'angle de la sécurité et la démonstration de cette sécurité ont représenté, aussi bien chez le Maître d'Œuvre que chez ses coopérants un volume de travail dépassant 10% du coût total en recherche et développement ".*

Le paragraphe suivant est consacré à la présentation des objectifs de sécurité pour les transports terrestres automatisés.

1.2. Evaluation de la sécurité dans les transports terrestres automatisés

La sécurité est une notion souvent utilisée de manière subjective. Pour la traduire en termes opératoires d'aide à la décision, des approches quantitatives pour la mesure et l'évaluation du risque encouru ont été développées. De nos jours, le progrès technique va dans le sens de l'accroissement de la complexité et donc des risques potentiels de dangers et il n'est plus concevable d'attendre que les accidents surviennent pour en tirer des enseignements. L'évaluation prévisionnelle du risque encouru devient indispensable dès la conception d'un système et pose le problème du niveau ou seuil d'acceptabilité du risque.

Le risque est défini comme la mesure d'un danger associant une mesure de l'occurrence d'un événement indésirable et une mesure de ses effets ou conséquences. Un danger est une situation pouvant nuire à l'homme, à la société ou à l'environnement. L'occurrence de l'événement indésirable est généralement mesurée par sa probabilité d'occurrence sur une période donnée. Les conséquences ou effets de l'événement indésirable peuvent être de nature humaine, économique ou porter sur l'environnement. Le risque est exprimé, par exemple, en unité monétaire par unité de temps, en nombre de morts par unité de temps ou en probabilité de mort par unité de temps. De nombreux objectifs quantifiés de sécurité ont été proposés dans /VILLEMEUR 88/ et /BARANOWSKI 90/ pour aider à des prises de décision tant pour la conception que pour l'exploitation des systèmes industriels. Trois approches ont été distinguées pour définir ces objectifs quantifiés de sécurité :

- fixation d'un objectif à partir du constat statistique : on part d'un constat statistique d'accidents pour telle activité et on souhaite faire mieux à l'avenir,
- fixation d'un objectif à partir d'un raisonnement économique : on compare le risque encouru et le bénéfice qu'en retire l'individu ou la collectivité,
- fixation d'un objectif à partir de considérations sur les risques individuels et collectifs acceptables : l'étude des risques individuels et collectifs encourus dans la société permet de situer des niveaux d'acceptabilité du risque.

Ces objectifs de sécurité sont généralement employés dans la conception des systèmes qui n'est reconnue satisfaisante que lorsque l'objectif de sécurité est atteint. Dans les paragraphes suivants, nous abordons la présentation de l'étude de sécurité des systèmes par la définition d'objectifs quantifiés visant à assurer la sécurité des systèmes de transports terrestres automatisés ainsi que les approches et moyens utilisés à cette fin.

1.3. Les objectifs quantifiés de sécurité des systèmes de transport terrestres

Les pouvoirs publics et les exploitants jugent la sécurité de manière continue pour un système en cours d'exploitation et de manière prévisionnelle pour un nouveau système. Les aspects de la sécurité d'un nouveau système sont jugés en deux temps : tout d'abord de manière prévisionnelle lors de la conception et ensuite de

manière continue pendant son exploitation. Ces jugements a priori et a posteriori sont appréciés en référence à un seuil de risque à ne pas dépasser. Ce seuil qui constitue un premier niveau global de sécurité à obtenir et à maintenir, présente l'avantage de prendre en compte non seulement la sécurité technique mais aussi tous les aspects de la sécurité des voyageurs /AMPELAS 77/.

De nombreuses réflexions menées sur la façon de définir des objectifs de sécurité ont permis de dégager une certaine philosophie résumée comme suit /DAVID 88/, /BARANOWSKI 90/ : l'application de technologies nouvelles dans le secteur des transports terrestres ne doit pas conduire à une situation plus défavorable du point de vue de la sécurité que celle que l'on observe dans les systèmes existants, étant entendu que, en matière de transports collectifs urbains, la référence de base a été le Métro de Paris. GABILLARD /79/ constate en effet que même sur des systèmes de transport bénéficiant d'une réputation très grande de sécurité, comme par exemple le réseau ferré urbain de la RATP, il se produit quand même bien que très rarement, certains accidents. Il est donc plus réaliste d'abandonner l'exigence de sécurité absolue et d'exiger pour un système de transport nouveau, qu'il soit "au moins aussi sûr" qu'un système de référence réputé sûr tel que le Métro de Paris.

En résumé, le principe de cette réglementation par objectifs quantifiés de sécurité, repose sur la définition, à partir de statistiques d'accidents, d'objectifs chiffrés en terme de taux d'accidents alloués a priori au nouveau système de transport, en exigeant un niveau de sécurité au moins analogue à celui des systèmes de référence. En pratique, les constructeurs appliquent ce principe et répercutent l'allocation globale sur les différents équipements. Il s'en suit une conversion du taux d'accidents en taux horaire moyen de défaillance contraire à la sécurité /THIBAULT 87/. A titre d'exemple, deux limites de taux d'accidents ont été définies pour le Métro de Lille à partir des statistiques /MARCOVICI 82/ :

- un taux d'accidents suite à accidents collectifs de 16×10^{-9}
- un taux d'accidents par accident individuel de 600×10^{-9}

Ces objectifs, compte tenu du trafic annuel prévu, sont d'abord transformés en taux d'accidents annuels puis en taux de défaillance-sécurité alloués à chacune des parties du système, en distinguant :

- les équipements dont la défaillance-sécurité est susceptible d'entraîner un "accident individuel",
- les équipements dont la défaillance-sécurité est susceptible d'entraîner un "accident collectif".

Après avoir fixé un objectif prévisionnel quantifié de sécurité du système de transport, le rôle des constructeurs et plus particulièrement des responsables de sécurité est de garantir le respect de cet objectif. Pour cela, à l'aide des méthodes spécifiques d'analyse de sécurité, ils doivent s'efforcer de minimiser les conséquences des sources d'accidents. Deux méthodes visent à assurer la sécurité d'un système de transport /AMPELAS 77/ :

- Méthode "fiabiliste" : la possibilité de pannes contraires à la sécurité est admise et on cherche à en diminuer la probabilité d'occurrence,
- Méthode classique utilisant la technique de sécurité "intrinsèque" : on cherche à éliminer toutes les pannes contraires à la sécurité par une conception astucieuse.

Le paragraphe suivant détaille ces deux types d'approches pour l'analyse de la sécurité.

1.4. La sécurité dans les transports terrestres automatisés

En matière de système de transport terrestre, on oppose souvent sécurité "probabiliste" à sécurité dite "intrinsèque".

1.4.1. Sécurité "intrinsèque"

La "sécurité intrinsèque" peut être définie de la façon suivante : "un système est dit en sécurité intrinsèque si on a l'assurance que toute défaillance d'un ou plusieurs composants ne peut le faire passer que dans une situation moins permissive que la situation dans laquelle il se trouve au moment de la défaillance, la situation la moins permissive étant l'arrêt complet" /DAVID 88/. La sécurité intrinsèque ne cherche donc pas à calculer la fiabilité des organes ni à faire en sorte qu'ils soient secourus par un organe de rechange en cas de défaillance (redondance), mais elle s'intéresse essentiellement à la manière dont ils peuvent tomber en panne. De ce fait, la mise en œuvre d'une méthode de sécurité intrinsèque nécessite une analyse profonde, fonctionnelle et structurelle, des équipements de protection en vue d'évaluer de façon exhaustive l'effet des pannes et/ou de leurs combinaisons non détectées après leur occurrence. Cet objectif est facilement réalisable lorsqu'il s'agit d'analyser des composants mécaniques, des circuits logiques câblés à relais et des composants électroniques isolés car les modes de défaillance peuvent être facilement recensés de façon exhaustive.

Cependant l'apparition des microprocesseurs dans les dispositifs de commande automatique de sécurité soulève d'importantes difficultés, dues non seulement à la complexité de ces circuits sur le plan matériel mais aussi au manque de moyens de qualification d'un logiciel /DAVID 90/, /EL KOURSI 90/. La complexité du dispositif entraîne plusieurs conséquences telles que la difficulté d'observer complètement l'état et les conditions de fonctionnement d'un tel circuit et l'impossibilité d'établir un inventaire complet de toutes les défaillances possibles et de leurs conséquences.

1.4.2. Sécurité "probabiliste"

Les limites de l'approche de type sécurité "intrinsèque" ont conduit au développement de nouveaux concepts de sécurité "probabiliste" mieux adaptés à l'étude des circuits à haute intégration. Les méthodes mises en œuvre et décrites dans /EL KOURSI 90/ reposent sur l'utilisation de redondances d'une part et de stratégies de tests systématiques d'autre part. Ces deux approches sont étroitement liées et sont classées en deux catégories :

- tests effectués par le microprocesseur lui-même ou par un circuit de sécurité auxiliaire,
- tests faisant appel à des configurations à plusieurs microprocesseurs.

1.4.3. Conclusion sur les deux approches de sécurité

Nous venons de voir que la démarche d'analyse de sécurité d'un système de transport automatisé repose sur la détermination d'un objectif quantitatif de sécurité en référence à d'autres systèmes de transport réputés sûrs. Afin d'atteindre et maintenir cet objectif de sécurité, il faut recourir à deux approches de sécurité l'une dite "intrinsèque" employée dans le cadre des composants discrets et l'autre dite "probabiliste" utilisée lorsqu'il s'agit de composants électroniques fortement intégrés (système numérique).

La mise en place, dès la conception, de ces deux approches de sécurité permet de minimiser les risques d'accidents catastrophiques. Néanmoins elle ne permet pas de supprimer totalement les défaillances inopinées. En effet, comme le confirme GABILLARD /79/, le niveau de sécurité conventionnelle admissible n'est pas défini objectivement car il dépend d'une appréciation subjective des concepteurs du système. C'est ainsi que le scénario dangereux auquel personne n'a pensé existe toujours et il n'appartient qu'à "la main du diable" qu'il se produise (tel, par exemple, l'outil oublié qui met en contact deux points éloignés ou le rat qui ronge un isolant ...). Les concepteurs et les responsables de certification doivent donc continuer à imaginer les scénarios d'insécurité et stimuler la recherche du palliatif avant la mise en service d'un nouveau système de transport.

2. CYCLE DE VIE ET CERTIFICATION DES SYSTEMES DE TRANSPORT TERRESTRES AUTOMATISES

Les experts « certifieurs » ont pour charge de garantir un degré de sécurité maximal du système de transport conçu. Ils ont à juger de la consistance des deux approches de sécurité (intrinsèque et probabiliste) employées par le constructeur. Le paragraphe suivant est consacré à la présentation de l'activité de ces spécialistes qui s'exerce tout au long du cycle de vie du système de transport automatisé (STA).

2.1. Cycle de vie d'un système de transport terrestre automatisé

Ce paragraphe présente la méthodologie générale de développement d'un système de transport et sa procédure d'agrément et enfin il propose un exemple d'échéancier de mise en exploitation d'un projet de système de transport automatisé.

2.1.1. Méthodologie de développement d'un système de transport terrestre automatisé (STA)

La méthodologie de développement d'un système de transport à conduite automatique, employée actuellement par les concepteurs, peut être définie et représentée sous la forme d'un cycle en "V", comprenant deux étapes, l'une de conception et l'autre de validation (figure 1.1).

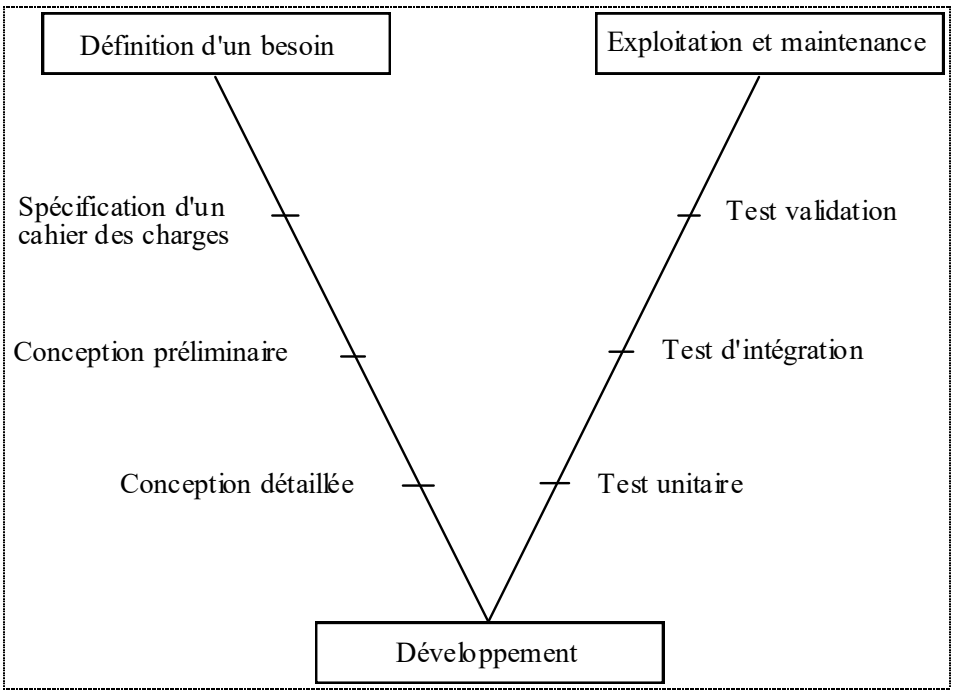


Figure 1.1 : méthodologie de développement d'un système de transport terrestre

Pour l'essentiel, le travail de certification des experts de l'INRETS-CRESTA concerne d'une part les phases de spécification, de conception préliminaire et détaillée et d'autre part la participation aux essais de validation de la sécurité.

2.1.2. Procédure d'agrément d'un système de transport automatisé

L'autorisation de conduite sur site d'un système de transport est accordée après agrément de tous les équipements automatiques. Ce paragraphe s'attache à exposer brièvement le déroulement de la procédure d'agrément en insistant particulièrement sur le rôle des experts de certification de l'INRETS-CRESTA. La procédure d'agrément se déroule de la manière suivante /LE TRUNG 84, 89/ :

- examen de sécurité avant la mise en service public (vérification, essais, ...) pour l'obtention d'un agrément provisoire,
- suivi de l'exploitation commerciale pendant un an, comportant l'analyse des incidents d'exploitation,
- attribution de l'agrément définitif.

En vue de l'attribution de cet agrément, il convient :

- d'élaborer des instructions techniques provisoires permettant d'assurer la sécurité globale du système de transport,
- de constituer une commission de sécurité regroupant l'exploitant, le constructeur et les représentants du ministère des transports, de l'INRETS-CRESTA, de la police, des pompiers, ...). Cette commission a pour charge de vérifier si le système de transport réalisé répond aux exigences de sécurité décrites par les instructions techniques provisoires,
- de désigner l'expert de certification qui sera impliqué dans les phases de conception et

d'exploitation et qui validera le dossier de sécurité du constructeur.

L'agrément provisoire ou définitif sera donné au vu des conclusions de la commission de sécurité et de l'expert de certification. Pour assurer ce rôle d'agrément, les experts de certification de l'INRETS-CRESTA doivent non seulement acquérir une bonne connaissance fonctionnelle et structurelle du système de transport, mais aussi procéder à une vérification du dossier de sécurité présenté par le constructeur, à l'aide des méthodes classiques d'analyse prévisionnelle de la sûreté de fonctionnement. La technique employée consiste à suivre une démarche inverse de celle suivie par le constructeur /LE TRUNG 84/. En effet, si par exemple le constructeur a analysé le système suivant une démarche déductive, le certifieur utilise une démarche inductive et la comparaison finale permet de juger de l'exhaustivité de l'analyse de sécurité.

2.1.3. Exemple d'échéancier d'un projet de STA

Afin de montrer l'implication importante des experts de certification dans le cadre du développement d'un système, un exemple d'échéancier d'un projet de système de transport est présenté en figure 1.2.

Plusieurs années sont nécessaires à la mise en service d'un système de transport, c'est au cours des dernières années que la contribution des experts de certification est la plus significative. Ils ont pour tâche principale de vérifier l'exhaustivité du dossier de sécurité du constructeur qui contient généralement :

- le profil de la mission,
- la description du système,
- une définition de l'objectif quantifié de sécurité,
- une liste de risques d'insécurité à éviter sous forme de scénarios d'accidents.

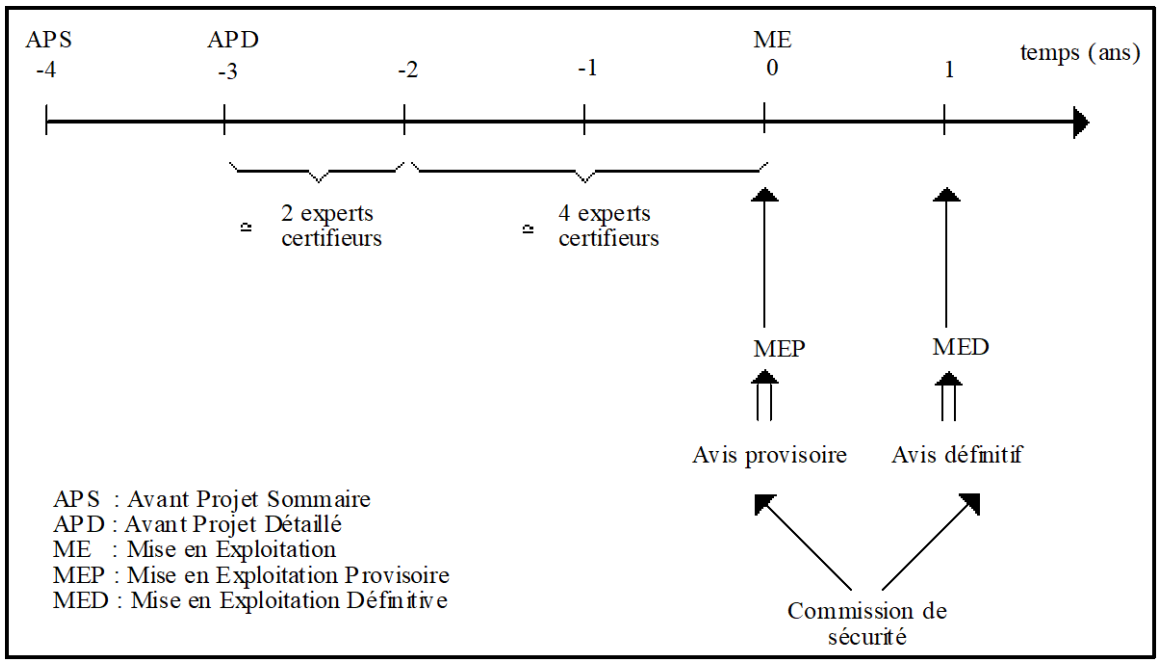


Figure 1.2 : Exemple d'échéancier de mise en exploitation d'un projet de système de transport

2.2. La certification des systèmes de transport terrestres automatisés

Afin de juger du respect des objectifs de sécurité définis précédemment et d'assurer au mieux l'exhaustivité de l'analyse des risques d'accidents dès la conception, la procédure de certification est intégrée dans le processus de développement des systèmes de transport automatisés (STA). L'activité de certification concerne à la fois les logiciels et les matériels de l'ensemble du système de transport .

2.2.1. Le but de la certification

L'agrément donné à un système de transport sur le plan de la sécurité entraîne soit une certification soit une homologation. On parle de certification lorsqu'il s'agit d'un nouveau système de transport (c'était le cas du VAL de Lille) et d'homologation pour un nouveau type d'un système déjà connu (VAL d'Orly et VAL de Toulouse conçus sur le modèle du VAL de Lille).

La certification repose essentiellement sur l'évaluation et l'analyse des risques d'accidents dûs à l'insuffisance des équipements de protection dans une configuration particulière d'évènements combinée avec une configuration géographique donnée. L'analyse du dossier de sécurité du constructeur des STA, basée notamment sur l'étude des scénarios d'accidents constitue une des phases fondamentales de la certification. Un scénario d'accident est une succession ordonnée d'évènements qui débouchent sur un risque potentiel. Chaque risque d'accident susceptible de mettre en cause la sécurité des voyageurs ou de provoquer l'altération de l'aptitude du STA à accomplir les fonctions de sécurité requises est traduit par les certifieurs en un scénario d'accident.

2.2.2. Démarche générale de certification

La certification dont l'objectif est de garantir le respect des fonctions et critères relatifs à la sécurité des STA, est un processus qui combine deux démarches :

- la première est la justification par le constructeur que l'ensemble des risques sont couverts et que le système est conforme aux objectifs de sécurité visés.
- la seconde est le contrôle de conformité, réalisé par plusieurs experts de certification qui évaluent la complétude et la consistance du dossier de sécurité du constructeur. A cette fin, ils analysent et critiquent les scénarios fournis pour en faire ressortir les éventuelles insuffisances quantitatives et qualitatives. De plus, ils sont souvent amenés à imaginer de nouveaux scénarios pour parfaire l'exhaustivité de l'étude. A l'issue de ce travail d'analyse de sécurité, une approbation provisoire ou définitive est établie.

Un modèle général de certification est présenté dans la figure 1.3.

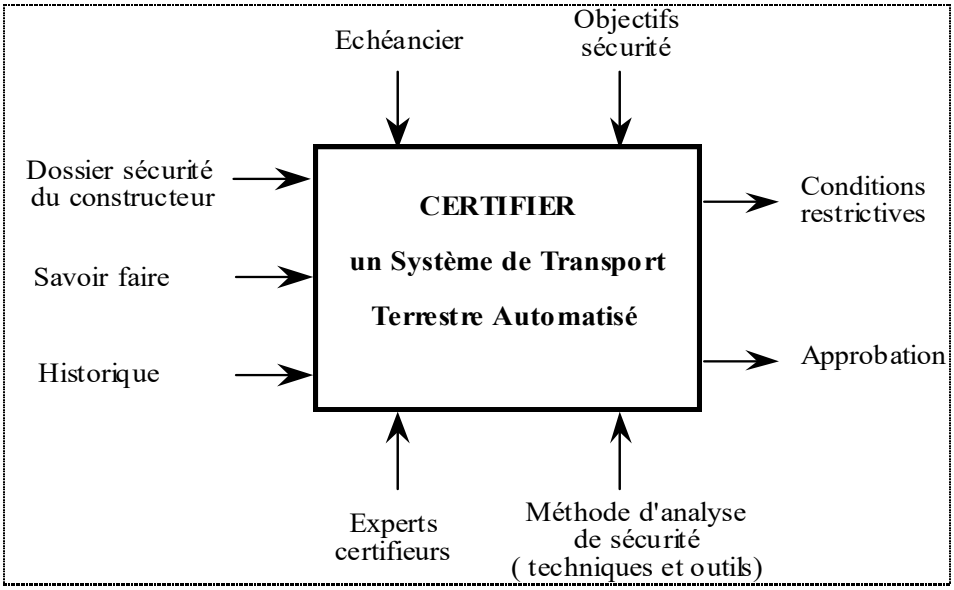


Figure 1.3 : Modèle général de certification d'un STA

A partir de l'examen du dossier de sécurité du constructeur, du savoir faire des experts, de l'historique des incidents survenus et compte tenu d'un certain nombre de contraintes (échéancier, objectifs de sécurité ...), l'expert certifieur s'appuyant sur des outils et méthodes d'analyse de sécurité est amené à élaborer un rapport de certification. Tant que l'expert certifieur démontre par le biais des scénarios d'accidents que les objectifs de sécurité ne sont pas atteints, le concepteur doit réviser et améliorer le système de transport sur le plan de la sécurité. En ce sens, le processus est itératif (fig. 1.4.).

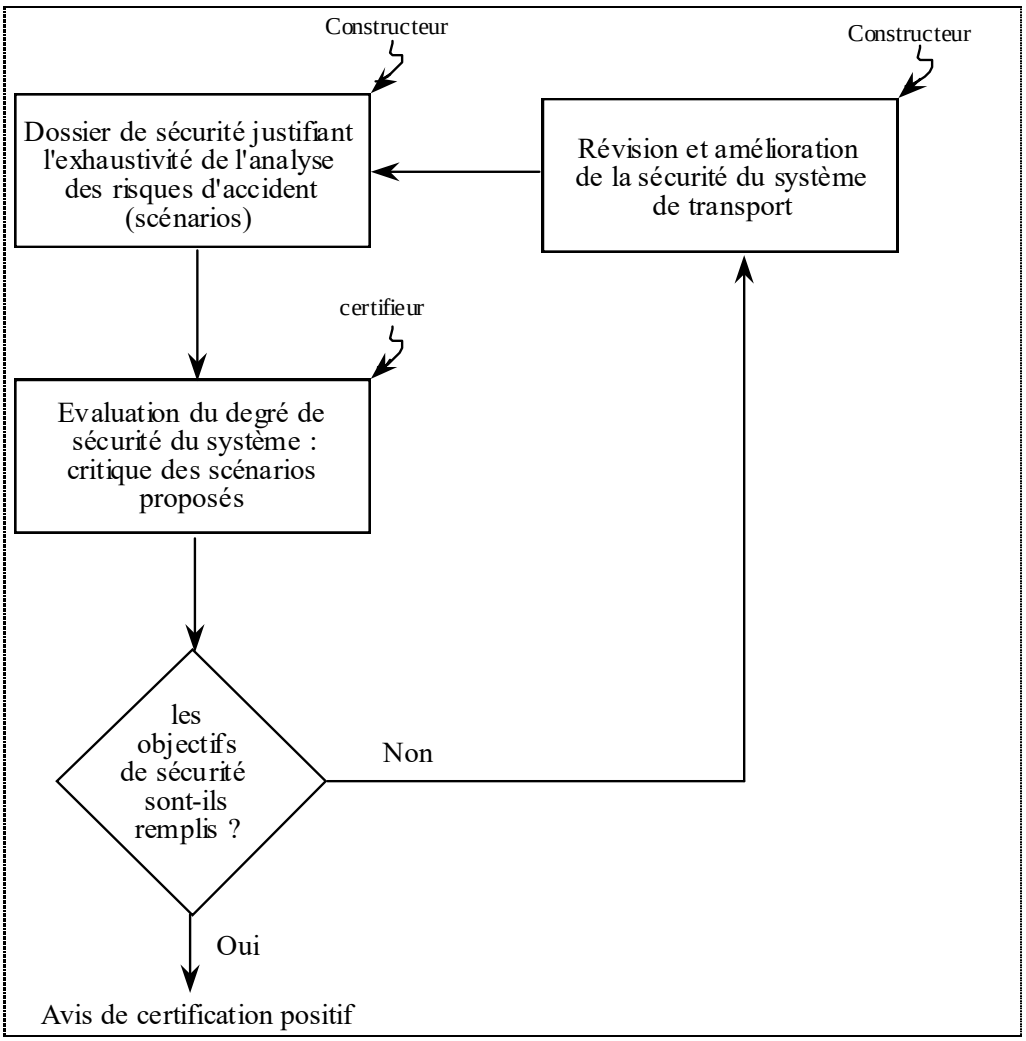


Figure 1.4 : Processus itératif de certification des systèmes de transport automatisés

3. METHODES, OUTILS ET TECHNIQUES UTILISES POUR LA CERTIFICATION DES "STA"

La démarche de certification est à la fois rationnelle et empirique. Elle fait appel à plusieurs techniques, méthodes et outils d'analyse prévisionnelle de la sécurité et est tributaire de l'évolution technologique des STA. Dans le paragraphe suivant, nous présentons deux types d'analyse de sécurité généralement combinés dans la certification des STA : l'analyse fonctionnelle et l'analyse qualitative.

3.1. Analyse technique et fonctionnelle du système de transport

L'analyse technique et fonctionnelle du système déroule en général les deux étapes suivantes /DAVID 90/ :

- Une décomposition hiérarchique du système en ses fonctions constituantes jusqu'à la désignation des fonctions les plus élémentaires. Parmi les méthodes utilisées pour aborder cette décomposition, nous pouvons citer SADT (Structured Analysis and Design Technics) /IGL Technology 89/ qui est couramment utilisée de nos jours.
- Les fonctions élémentaires sont ensuite modélisées par des graphes illustrant les différents états de l'automate et les conditions de changement d'état. Ces modèles sont couramment représentés par des automates à états finis, les réseaux de Pétri ou le GRAFCET.

Une revue des principales méthodes et des outils d'analyse fonctionnelle appliqués à la sûreté de fonctionnement des systèmes est proposée dans /BARBET 88/, /PIERREVAL 90/ et /CALVEZ 90/.

3.2. Analyse qualitative de la sécurité du système de transport

A partir de la définition du fonctionnement du système, l'analyse qualitative consiste à étudier globalement la sécurité du système en réalisant une analyse des risques. L'objectif est d'identifier les situations dangereuses, les accidents potentiels, les éléments ou équipements dangereux ainsi que la gravité des conséquences qui en résultent. L'utilisation des méthodes d'analyse prévisionnelle de la sûreté de fonctionnement contribue avantageusement à éliminer les défaillances ou à réduire leur probabilité d'occurrence. Plusieurs méthodes d'analyse existent dont les principales sont décrites dans /VILLEMEUR 88/ :

- l'Analyse Préliminaire des Dangers (APD)
- l'Analyse des Modes de Défaillance et de leurs Effets (AMDE)
- la Méthode du Diagramme de Succès (MDS)
- la Méthode de la Table de Vérité (MTV)
- la Méthode de l'Arbre des Causes (MAC)
- la Méthode des Combinaisons de Pannes Résumées (MCPR),
- la Méthode de l'Arbre des Conséquences (MACQ)
- la Méthode du Diagramme Causes-Conséquences (MDCC)
- la Méthode de l'Espace des Etats (MEE)

L'ensemble de ces méthodes repose sur deux démarches fondamentales, l'une de type inductive et l'autre de type déductive. Dans la démarche inductive, le raisonnement va du plus particulier au plus général, ce qui conduit à une étude détaillée des effets d'une défaillance sur le système et son environnement. L'AMDE, la MTV, la MCPR et la MACQ sont des méthodes inductives. Dans la démarche déductive, le raisonnement va du plus général au plus particulier de telle façon que, face au système défaillant, on recherche les causes de la défaillance. L'analyse effectuée à la suite de catastrophes pour retrouver les causes est de nature déductive. La principale méthode déductive est la Méthode de l'Arbre des Causes (MAC).

Dans le processus de certification, les experts utilisent dans la mesure du possible des méthodes d'analyse de sécurité complémentaires à celles employées par le constructeur. Le paragraphe suivant détaille les principales méthodes d'analyse de sûreté de fonctionnement utilisées en matière de certification des STA.

3.2.1. Analyse Préliminaire des Dangers (APD)

La méthode d'Analyse Préliminaire des Dangers (APD) a pour objectifs :

- d'identifier les dangers d'un système et leurs causes possibles,
- d'évaluer la gravité des conséquences induites par les situations dangereuses et les accidents potentiels.

Cette analyse débouche sur la définition de tous les moyens, toutes les actions correctrices visant à éliminer ou à maîtriser les situations dangereuses ou les risques potentiels d'accidents. Comme son nom l'indique, cette méthode est généralement considérée comme préliminaire à la réalisation d'études complémentaires de sûreté de fonctionnement. La mise en évidence d'un danger majeur nécessite généralement l'utilisation d'autres méthodes ou techniques prévisionnelles de la sûreté de fonctionnement.

3.2.2. Méthode de l'Arbre des Causes (MAC)

La méthode de l'Arbre des Causes (MAC) est aussi connue sous le nom d'Arbre de Défaillance, d'Arbre des Défauts ou d'Arbre des Fautes. C'est la méthode la plus largement utilisée pour l'analyse de la fiabilité, de la disponibilité et de la sécurité des systèmes. L'Arbre des causes représente graphiquement les combinaisons d'événements qui conduisent à la réalisation d'un événement indésirable unique ; ce dernier constitue l'événement "sommet" de l'arbre des causes et est déterminé par une analyse préliminaire des dangers (APD). L'Arbre des causes est formé de niveaux successifs d'événements tels que tout événement d'un niveau est généré à partir des événements du niveau inférieur combinés par des opérateurs logiques ET et OU. Ces événements sont généralement des défauts associés à des défaillances de matériels ou à des erreurs humaines pouvant conduire à l'événement indésirable. Ce processus déductif est poursuivi jusqu'à ce que l'on obtienne des événements élémentaires, indépendants entre eux et dont on connaît la probabilité d'occurrence.

3.2.3. Analyse des Modes de Défaillance et de leurs Effets (AMDE)

L'AMDE est une méthode inductive permettant d'effectuer une analyse complète des modes de défaillance des composants, de leurs causes, et de leurs effets sur le système. Elle permet d'évaluer l'importance des modes de défaillance vis à vis de la sûreté de fonctionnement (disponibilité, fiabilité, maintenabilité ou sécurité) du système. On distingue quatre étapes principales pour réaliser une AMDE :

1. définition du système, de ses fonctions et de ses composants,
2. établissement des modes de défaillance des composants et de leurs causes,
3. Etude et évaluation des modes de défaillance sur les fonctions du système,
4. Conclusions et recommandations.

Une extension naturelle de l'AMDE est "l'Analyse des Modes de Défaillance, de leurs Effets et de leur Criticité" (AMDEC). Pour chaque mode de défaillance, elle permet d'évaluer le couple "probabilité-gravité". Quatre classes de gravités permettent de caractériser les effets ou conséquences d'un mode de défaillance : mineurs, significatifs, critiques ou catastrophiques. La criticité du mode de défaillance d'un composant est appréciée à l'aide du couple "probabilité-gravité" : plus la probabilité est grande et plus les effets sont jugés

pénalisants, plus la criticité du mode de défaillance est importante et plus il devient nécessaire de prendre des mesures correctives.

3.2.4. Méthode des Combinaisons de Pannes Résumées (MCPR)

La Méthode des Combinaisons des Pannes Résumées (MCPR), issue du domaine de l'aéronautique, a été formalisée conjointement par la Société Nationale des Industries Aéronautiques et Spatiales (SNIAS) et les Autorités de certification du Ministère de l'Air français, pour l'analyse de la sécurité des avions Concorde et Airbus /LIEVENS 76/. L'analyse des Modes de Défaillance et de leurs Effets (AMDE), qui met généralement en évidence les défaillances simples, doit être complétée par l'étude des combinaisons de défaillances qui aboutissent à des événements indésirables. La MCPR, utilisée dans le prolongement de l'AMDE, détermine de manière inductive de telles combinaisons de défaillances. L'utilisation de la MCPR nécessite la mise en oeuvre de 4 étapes /VILLEMEUR 88/ :

1. décomposition du système et élaboration d'une AMDE : pour chaque composant du système, on étudie principalement les modes de défaillance et leurs effets sur le système considéré ainsi que sur les autres systèmes.

2. Elaboration des Pannes Résumées Internes (PRI) : on constate pour un mode de défaillance ou pour plusieurs combinés entre eux, que les effets ou les conséquences sur le ou les systèmes considérés sont identiques. Ces modes de défaillance sont alors regroupés en des ensembles de pannes appelés "Pannes Résumées internes".

3. Elaboration des Pannes Résumées Externes (PRE) : suite au déroulement des étapes précédentes, tous les systèmes (par exemples s1, s2, ... sn) sont analysés. Les Pannes Résumées Internes (ou leurs combinaisons) d'autres systèmes peuvent affecter le fonctionnement du système étudié (s1 par exemple) : on les appelle les "Pannes Résumées Externes" de s1.

4. Elaboration des pannes résumées globales (PRG) : une PRG englobe les pannes résumées internes, externes et/ou leurs combinaisons ayant les mêmes effets sur le système étudié que sur les autres systèmes.

Des relations très étroites existent entre le concept de panne résumée (interne, externe ou globale) et celui de Mode de Défaillance :

- les PRI d'un système correspondent aux modes de défaillance dus à des causes uniquement internes,
- les PRE d'un système correspondent aux modes de défaillance dus à des causes uniquement externes,
- les PRG d'un système correspondent aux modes de défaillance dus à des causes internes et/ou externes.

La MCPR, méthode purement inductive, analyse en premier lieu les effets des modes de défaillance des composants puis étudie les combinaisons de ces modes de défaillance afin de définir des ensembles de pannes correspondant à des fonctionnements anormaux ou à des événements indésirables pour les systèmes à

analyser. La MCPR s'attache donc à extraire seulement les combinaisons significatives sur le plan de la sécurité et se présente alors comme une extension de l'AMDE.

La MCPR est employée par les experts certifieurs de l'INRETS-CRESTA sous le nom de méthode de "Recherche de Combinaison de Pannes Significatives" (RCPS). On distingue trois types de pannes en fonction de leurs conséquences sur la sécurité du système de transport : PE : Panne Elémentaire, PR : Panne Résumée, PG : Panne Globale. Sont regroupées sous le terme de Pannes Elémentaires (PE) les pannes dont les conséquences sont identiques sur le comportement du système de transport. Leur nombre habituellement important rend pénible leur exploitation. Il est d'usage de regrouper en une panne unique dite Panne Résumée (PR) les PE dont la ou les conséquences sur le système sont généralement bénignes. Combinées entre elles, les PR peuvent aggraver les conséquences sur le système; leur association constitue une Panne Globale (PG). Les PG sont souvent dangereuses et résultent d'une combinaison de PR prises dans un ordre d'occurrence bien déterminé.

Dans la suite, pour la formalisation et la structuration des connaissances de sécurité, nous nous intéressons essentiellement aux PR qui ont un caractère universel de par leur indépendance vis à vis de la technologie.

3.3. Limites des méthodes prévisionnelles d'analyse de la sécurité pour la certification des systèmes de transport automatisés

Les méthodes prévisionnelles d'analyse des défaillances présentent un intérêt indéniable pour l'analyse de la sécurité des systèmes de transport automatisés (STA). Elles permettent d'évaluer et d'analyser les dangers liés à l'utilisation du système de transport et d'identifier les effets, les causes ainsi que les combinaisons des modes de défaillance des composants. Cependant, leur mise en œuvre présente au moins trois inconvénients :

- la nécessité d'utiliser plusieurs méthodes à la fois.
- la difficulté d'analyse due à la taille considérable des arbres de causes élaborés.
- la difficulté d'être exhaustif pour l'analyse des défaillances significatives dans le cas d'un équipement ou système de transport complexe.

En conclusion, ces méthodes d'analyse prévisionnelle de la sécurité des STA, bien que nécessaires pour mettre en évidence des situations "indésirables", ne sont pas suffisantes /HADJ-MABROUK 92d/. Aucune méthode, à elle seule, ne permet d'assurer l'exhaustivité de l'analyse de sécurité. Il est souvent nécessaire de recouper les résultats obtenus par une méthode avec ceux obtenus par une autre complémentaire. Pour de telles méthodes, l'exhaustivité de l'analyse de sécurité demeure essentiellement fondée sur l'intelligence et l'intuition humaine et il faut continuer d'affiner l'étude de sécurité tout au long de la vie utile des systèmes de transport.

De ce fait, nos travaux se sont notamment efforcés de spécifier une méthodologie d'analyse et d'évaluation de la sécurité basée sur les techniques d'Intelligence Artificielle (IA) dont la finalité est de compléter et renforcer les méthodes actuelles.

4. SPECIFICATION DES BESOINS EN MATIERE DE CERTIFICATION

Ce paragraphe précise les propriétés requises par un système d'aide répondant aux objectifs assignés pour l'évaluation de la sécurité des STA. Les solutions retenues pour concevoir et mettre en oeuvre un tel outil d'aide sont présentées et commentées.

4.1. Trois objectifs pour un système d'aide à l'analyse de sécurité

La conception d'un outil d'aide à l'analyse de la sécurité des STA a été motivée par diverses constatations révélées par la phase d'identification du problème /HADJ-MABROUK 91b/. Les principales sont les suivantes :

- le besoin d'améliorer l'expertise et la qualité de décision dans le domaine de certification par l'archivage, la formalisation et la diffusion du savoir-faire des certifieurs,
- la difficulté d'exploiter la masse considérable de connaissances impliquées par l'analyse de sécurité,
- le souci de juger l'exhaustivité du dossier de sécurité du constructeur,
- la nécessité de rechercher des similarités et des analogies entre plusieurs configurations ou situations d'insécurité,
- l'exigence d'un formalisme précis, rigoureux et explicite pour représenter les scénarios d'accidents.

Ces raisons ont orienté le cahier des charges vers la conception et le développement d'un outil d'aide à la certification des systèmes de transport automatisés fondé sur des techniques d'IA et répondant aux trois objectifs connexes suivants :

1. structurer et classifier les scénarios historiques existants de façon à pouvoir rechercher les scénarios similaires à celui proposé par le constructeur,
2. évaluer puis confirmer le contenu et la consistance des scénarios d'accidents fournis dans le dossier de sécurité du constructeur,
3. aider à générer d'autres scénarios susceptibles de parfaire l'exhaustivité de l'analyse des risques d'accidents.

4.2. Le recours aux techniques d'Intelligence Artificielle

Les modes de raisonnement utilisés en matière de certification (inductif, déductif, par analogie ...) ainsi que la nature même des connaissances de sécurité (incomplètes, évolutives, empiriques, qualitatives ...) confirment qu'une solution informatique conventionnelle n'est pas adaptée et que le recours aux techniques de l'Intelligence Artificielle (IA) semble approprié. L'IA s'est donnée pour but l'étude et la simulation des activités intellectuelles humaines /LESCORT 85/. Elle s'efforce de créer des machines capables d'un comportement intelligent et a pour vocation ambitieuse de doter l'ordinateur de quelques unes des facultés de l'esprit humain : apprendre, reconnaître, raisonner ou encore s'exprimer à l'aide d'un langage /GRUNDSTEIN 88/. L'aptitude à comprendre le langage naturel et l'aptitude à raisonner sont deux clés de voûte de l'intelligence /HOLSAPPLE 87/.

Pour nos travaux, nous avons eu recours à trois aspects particuliers du domaine de l'IA : l'acquisition de connaissances, l'apprentissage automatique et les Systèmes à Base de Connaissances. L'élaboration de la base de connaissances d'un SBC nécessite le recours aux techniques et méthodes d'acquisition de connaissances pour recueillir, structurer puis formaliser les connaissances. L'acquisition de connaissances n'a pas permis, à elle seule, d'extraire efficacement certaines connaissances expertes de certification. Aussi, l'utilisation conjointe de l'acquisition de connaissances et de l'apprentissage automatique apparait-elle comme une solution très prometteuse.

4.3. Approche retenue pour l'aide à l'analyse de la sécurité et la certification des Systèmes de Transport Automatisés (STA)

Les solutions retenues pour concevoir et mettre en œuvre un outil d'aide à l'analyse de la sécurité des STA implique les deux grandes activités suivantes /HADJ-MABROUK 91a/ :

- extraire, formaliser et archiver les situations d'insécurité, de façon à constituer une bibliothèque de cas types couvrant l'ensemble du problème. Celle ci est appelée Base de Connaissances Historiques des Scénarios (BCHS). Cette activité a nécessité le recours aux techniques et méthodes d'**acquisition de connaissances**.

- exploiter les connaissances historiques archivées afin d'en dégager un savoir-faire en certification susceptible d'aider les experts certifieurs à juger l'exhaustivité de l'analyse de sécurité proposée par le constructeur. Les approches envisagées pour cerner cette deuxième activité sont basées sur l'emploi des méthodes d'**apprentissage automatique**.

Les deux chapitres suivants présentent ces deux activités impliquées dans la méthodologie d'analyse de la sécurité des systèmes de transport terrestres automatisés.

CONCLUSION

Le présent chapitre a dégagé les différents aspects techniques et organisationnels du domaine de la certification des systèmes de transport terrestres automatisés (STA). Il a permis de caractériser le domaine et de définir les concepts impliqués avant de formuler les objectifs de notre travail.

Garantir les fonctions et les critères relatifs à la sécurité des STA est l'objectif principal de l'activité de certification qui repose sur une analyse profonde des risques mis en évidence par des scénarios d'accidents. La mission des experts certifieurs consiste à évaluer la complétude du dossier de sécurité proposé par le constructeur afin de prouver que le STA construit répond aux exigences de sécurité définies et qu'il ne peut en aucun cas être une source de risque pour les usagers. Le développement d'un nouveau STA requiert la définition d'objectifs quantifiés de sécurité en référence aux systèmes de transport existants et réputés sûrs. Afin de maintenir et respecter ce niveau de sécurité, deux approches de sécurité sont employées par les constructeurs de STA : la sécurité "intrinsèque" et la sécurité "probabiliste". Afin de juger de la complétude de ces deux approches de sécurité, les certifieurs disposent de plusieurs méthodes d'analyse de la sécurité. Ces méthodes, bien que nécessaires pour pallier les défaillances critiques, ne sont pas suffisantes. L'imagination de nouveaux scénarios d'accidents par les experts de certification demeure indispensable pour tendre vers l'exhaustivité de l'analyse de sécurité.

La prise en compte de cette réalité nécessite le recours à des outils logiciels "intelligents" visant à compléter les méthodes actuelles d'analyse de sécurité. L'approche envisagée est centrée sur les techniques d'Intelligence Artificielle (IA) et notamment sur l'emploi des Systèmes à Base de Connaissances (SBC). La démarche que nous avons développée et qui fait l'objet de cette thèse repose sur l'utilisation conjointe de l'acquisition de connaissances et de l'apprentissage automatique pour concevoir et mettre en œuvre un SBC pour l'aide à la certification des STA. La réalisation d'une base de connaissances de certification nécessite en premier lieu le recours aux techniques, méthodes et outils d'acquisition de connaissances que nous présentons dans le chapitre suivant.

Chapitre 2 :

L'acquisition de connaissances pour l'élaboration d'une base de connaissances de certification.

INTRODUCTION

L'acquisition de connaissances, reconnue comme un "goulot d'étranglement" dès l'avènement des systèmes experts ou plus généralement des Systèmes à Base de Connaissances est encore de nos jours considérée comme une tâche cruciale de leur réalisation. L'"extraction" ou l'"élicitation" désigne le recueil des connaissances auprès de l'expert du domaine alors que les notions de "transfert" ou "transmission" d'expertise désignent le recueil puis la formalisation des connaissances d'un expert humain. Le terme "Acquisition de connaissances" désigne quant à lui l'ensemble des démarches nécessaires à l'élaboration d'une base de connaissances d'un système expert.

Après avoir présenté l'enjeu de l'acquisition de connaissances et quelques caractéristiques des systèmes à base de connaissances, nous détaillons dans ce chapitre les techniques, méthodes et outils actuellement utilisés pour effectuer le transfert de la connaissance depuis le monde réel vers un système artificiel. Cette présentation débouche sur le choix d'une méthode d'acquisition de connaissances. Les apports et les limites de cette méthode pour le développement du système à base de connaissances d'aide à la certification sont présentés dans la deuxième partie du présent chapitre.

1. BUT DE L'ACQUISITION DE CONNAISSANCES

L'Acquisition des Connaissances (AC) constitue l'un des thèmes centraux des recherches sur les SBC et l'une des clés non seulement du succès du développement d'un tel système, mais aussi de son intégration et de son utilisation en milieu opérationnel. L'AC fait intervenir essentiellement deux groupes d'acteurs : l'expert, détenteur d'un savoir-faire par nature difficilement exprimable et le cogniticien ou ingénieur de la connaissance qui doit extraire et formaliser les connaissances relatives à ce savoir faire, généralement implicite chez l'expert. Ce processus long et délicat est pourtant fondamental pour réaliser une base de connaissances efficace. Initialement centrée sur le couple Expert/Cogniticien, l'Acquisition des connaissances a très vite soulevé des problèmes cruciaux tels que l'identification des besoins des utilisateurs ou le choix d'un mode de représentation des connaissances. Le décalage trop important entre le langage utilisé par les experts pour décrire leur problème et le niveau d'abstraction des formalismes de représentation des connaissances a motivé de nombreuses recherches visant à faciliter le transfert d'expertise.

Les nouvelles approches de l'AC visent la définition de méthodologies plus efficaces et la conception de logiciels permettant d'aider ou de remplacer partiellement le cogniticien. Certains travaux proposent de voir la conception d'un SBC comme un processus de construction d'un modèle conceptuel, à partir de toutes les sources de connaissances (humaines ou documentaires) disponibles sur la résolution du problème à traiter. Dans ce contexte, l'AC est perçue comme une activité de modélisation. D'autres travaux soulignent l'intérêt de méthodologies visant à guider le cogniticien dans ce processus de transfert/modélisation. Des outils et méthodes permettent notamment de faciliter la verbalisation, les interviews d'experts et les analyses de documents. Les méthodes d'AC disponibles actuellement proviennent pour l'essentiel de la psychologie cognitive (modèles de raisonnement humain, techniques de recueil des connaissances), de l'ergonomie (analyse de l'activité de l'expert et du futur utilisateur), de la linguistique (pour rendre plus efficace l'exploitation des documents ou guider l'interprétation de données verbales) et du génie logiciel (description du cycle de vie d'un SBC).

En résumé l'AC peut être définie comme l'ensemble des démarches nécessaires pour recueillir, structurer puis formaliser des connaissances dans le processus de conception d'un SBC. Pour aider l'expert ou le cogniticien dans cette phase de transfert d'expertise, il existe :

- Des méthodologies de développement d'un SBC. Dans la littérature, on distingue généralement deux approches pour décrire le cycle de vie d'un SBC :

  - Le "prototypage rapide" dont l'objectif est de réaliser dès que possible une maquette de faisabilité.
  - L'"acquisition structurée" des connaissances qui s'attache à modéliser la totalité de la connaissance avant son implémentation.

- Des méthodes d'acquisition de connaissances qui proposent un cadre méthodologique pour l'acquisition des connaissances lors de la conception d'un SBC.

- Des techniques d'extraction ou de recueil d'expertise souvent inspirées des travaux en psychologie. Parmi les techniques fréquemment utilisées, notons les interviews, les questionnaires, l'analyse de protocoles et le tri conceptuel.

- Des outils d'extraction des connaissances qui visent à automatiser certaines étapes du processus d'acquisition en élaborant directement la base de connaissances d'un SBC. Ces outils (exemple AQUINAS, ROGET, MOLE) sont basés sur l'emploi d'une ou plusieurs techniques d'extraction de connaissances.

Ces moyens d'acquisition de connaissances permettent actuellement d'identifier divers aspects du raisonnement humain. Toutefois, il n'existe pas à ce jour de solution type et le cogniticien est souvent démuni devant le problème du choix d'une méthode adaptée au contexte des connaissances expertes.

2. SYSTEMES A BASE DE CONNAISSANCES (SBC)

Avant d'aborder le problème de l'acquisition de connaissances, il convient de rappeler les principales caractéristiques d'un SBC. Cette présentation est faite sans en détailler les spécificités et le fonctionnement, fort bien décrits par ailleurs dans /LAURIERE 84/, /FARRENY 85, 87/, /BENCHIMOL 86/, /CORDIER 87/ et /HART 88/. La dernière partie de ce paragraphe met l'accent sur les étapes du développement d'un SBC.

2.1. Caractéristiques d'un SBC

Depuis la naissance de l'informatique, les ordinateurs sont considérés comme des machines à programmer. Il existe une grande variété de langages de programmation qui possèdent en commun les caractéristiques suivantes : ils ordonnent à la machine des opérations déterminées, ces opérations sont elles-mêmes entrées de façon ordonnée, figée et seuls les informaticiens peuvent y accéder. Cela limite donc leur intérêt car, si la machine sait exécuter les ordres, elle ne peut ni dialoguer avec l'utilisateur ni lui expliquer son raisonnement. Le recours aux SBC a permis de pallier cette lacune. Ceux-ci tentent de reproduire la démarche intellectuelle

d'un spécialiste, ce qui en fait à l'heure actuelle, un outil précieux déjà opérationnel dans plusieurs secteurs. Habituellement, un SBC se présente comme l'association d'une Base de Connaissances (BC), d'un moteur d'inférences et d'une interface Homme-Machine.

- La base de connaissances détient la connaissance spécifique d'un domaine d'application (savoir-faire et modes de raisonnement de l'expert). Elle est couramment organisée autour des deux entités suivantes :
  - une base de faits représentant les informations qui décrivent des situations établies par l'utilisateur ou déduites par le moteur d'inférences (ou mécanisme de raisonnement),
  - une base de règles qui constitue le savoir-faire sur le domaine et indique donc les actions à entreprendre lorsqu'on est en présence d'une situation précise.

- Le moteur d'inférence exploite les données contenues dans la base de connaissances en vue d'élaborer la solution des problèmes posés. Il met en œuvre des mécanismes déductifs (fonctionnement en chaînage avant) ou inductif (fonctionnement en chaînage arrière). L'objectif est de résoudre un problème décrit par les données contenues dans la base de faits, en sélectionnant et déclenchant les règles contenues dans la base de règles.

- L'interface Homme-Machine permet d'assurer le dialogue avec les utilisateurs et/ou l'expert du domaine. Cette interface peut être scindée en deux parties :
  - une interface utilisateur à travers laquelle l'utilisateur décrit les données du problème posé et reçoit la solution accompagnée d'une trace, reflet de la résolution du problème.
  - une interface d'acquisition permettant le remplissage, la révision et la mise à jour de la base de connaissances.

L'évocation de ces trois composantes révèle ce qui différencie fondamentalement un SBC d'un logiciel classique : la séparation entre les connaissances du domaine et le mécanisme chargé de leur traitement.

Depuis MYCIN et ses dérivés, les applications des SBC se sont multipliées mais les systèmes développés ne sont toujours pas à la hauteur des espoirs qu'ils avaient suscités /AUSSENAC 89/. En effet, il est difficile de constituer une "bonne" base de connaissances qui permette au système de traiter d'autres problèmes que des cas classiques ou triviaux et même de résoudre de nouveaux problèmes. Les SBC plus récents, qualifiés de SBC de "deuxième génération", visent à surmonter ces obstacles grâce à des moteurs d'inférence plus sophistiqués ainsi qu'à la prise en compte de nouveaux modes de représentation de la connaissance. Outre les règles de production, les SBC de seconde génération utilisent les langages de représentation à base d'objets, des langages qui reposent sur les notions de "frames" ou de réseaux sémantiques. Dans le cadre des SBC de seconde génération, on distingue deux types de connaissances considérées comme des propriétés caractéristiques de ces systèmes : les connaissances "profondes" par opposition aux connaissances "de surface". Les connaissances de surface sont des heuristiques traduisant purement le raisonnement. En revanche les connaissances profondes sont souvent qualifiées de connaissances théoriques.

2.2. Acteurs impliqués dans le développement d'un SBC

Traditionnellement, on distingue six groupes d'acteurs autour d'un projet SBC : un comité directeur, les experts, les cogniticiens, le groupe de développement, les utilisateurs et le groupe de maintenance. Les trois principaux groupes sont le groupe expert, le groupe cogniticien et l'utilisateur.

- Le groupe des experts humains détient une grande partie de la connaissance spécifique à un domaine et possède le savoir-faire lié à la résolution du problème. Par nature les connaissances sont vagues, ambiguës et évolutives. Leur organisation dans la mémoire de l'expert est complexe, ce qui rend difficile leur extraction.
- Le groupe des cogniticiens est chargé d'identifier, recueillir, expliciter, analyser et formaliser les connaissances et modes de raisonnement de l'expert pour élaborer la base de connaissance du SBC. Le cogniticien joue un rôle primordial dans le transfert d'expertise, malgré le biais ou la distorsion que peut provoquer son intervention.
- Le groupe des utilisateurs exploite finalement les connaissances expertes emmagasinées dans le SBC.

L'évocation du rôle de ces acteurs nous amène à mettre en évidence des problèmes majeurs liés à l'élaboration de la base de connaissances d'un SBC. La conception d'une base de connaissances nécessite l'extraction, l'analyse, la structuration et la formalisation du savoir-faire d'un domaine auquel on accède à travers un ou plusieurs individus, qualifiés d'experts. Dès lors, le transfert de cette expertise soulève les questions délicates suivantes : qui détient réellement l'expertise ?, comment peut-on y accéder ?, comment l'extraire ?, comment la formaliser sans la déformer ?, quelle représentation choisir ?, comment valider et maintenir les connaissances recueillies ?

Diverses recherches sont menées pour mieux cerner ces problèmes inhérents à l'acquisition de connaissances et à la conception d'un SBC. Des moyens (méthodes, techniques, outils) pour l'acquisition des connaissances sont aujourd'hui accessibles au cogniticien et à l'expert et offrent un cadre méthodologique pour le développement d'un SBC.

3. MOYENS DE L'ACQUISITION DE CONNAISSANCES

Sans prétendre à l'exhaustivité, l'objectif de ce paragraphe est d'effectuer une synthèse des méthodes, outils et techniques proposés récemment dans le domaine de l'acquisition de connaissances. Les travaux de /AUSSENAC 89, 90, 91/, /DIENG 90/ et /BENKIRANE 91/ retracent de façon assez complète l'évolution des recherches les plus récentes et présentent un éventail des méthodes et outils existants.

3.1. Méthodologie de développement d'un SBC

Généralement, les méthodologies de développement existantes sont associées à une certaine vision du cycle de vie du SBC. En effet, les SBC sont avant tout des logiciels et les étapes de développement proposées par de nombreux auteurs sont inspirées des résultats de recherche en génie logiciel. Dans la littérature, on distingue généralement deux approches pour décrire le cycle de vie d'un SBC /DIENG 90/ : le "prototypage rapide" et l'"acquisition structurée" des connaissances.

3.1.1. Le prototypage rapide

L'objectif essentiel du prototypage rapide est de montrer la faisabilité du problème identifié en réalisant dès que possible une maquette. Cette démarche est incrémentale et cyclique et en cela entraîne des retours-arrière. Plusieurs méthodologies de développement de SBC s'inscrivent dans le cadre de cette approche /BENCHIMOL 86/, /NASSIET 87/, /GALLOUIN 88/, /HART 88/. La thèse de D. Nassiet /NASSIET 87/ présente une méthodologie de développement de ce type d'approche, particulièrement complète et structurée. L'intérêt de ce document est de rapporter parallèlement à l'énoncé théorique d'une méthode, l'expérience de son application au domaine du diagnostic technique. Cette méthode, très clairement exposée, rappelle une idée simple et indispensable au succès du développement d'un SBC : il est important de procéder progressivement, en validant les résultats obtenus à chaque étape. Cinq phases chronologiques sont impliquées dans la méthodologie : identification, réalisation d'une maquette, développement d'un prototype, évaluation des performances du prototype et exploitation du système sur site.

Identification : au cours de cette étape, le cogniticien définit et caractérise l'expertise. Il se familiarise avec les principaux concepts attachés au domaine : le type de problème à traiter, les aspects du problème risquant de soulever des difficultés, les principaux éléments de connaissances manipulés par l'expert, les sources et la nature des connaissances, les modes de raisonnement de l'expert, la stabilité, la complétude et la consistance des informations manipulées par l'expert. Cette phase d'identification aboutit à la rédaction d'une première version du cahier des charges, point de départ de la deuxième étape dédiée à la réalisation de la maquette de faisabilité.

Réalisation d'une maquette : l'objectif escompté de cette phase est d'étudier la faisabilité du système final à travers le traitement d'un sous-problème représentatif du domaine, de familiariser le cogniticien au domaine d'expertise et de sensibiliser les experts aux problèmes du recueil de connaissances. La maquette présente une forme réduite du système et s'attache à cerner toutes les caractéristiques du produit. Elle nécessite une étape d'extraction de connaissances permettant au cogniticien d'identifier les connaissances en terme de source, nature, forme ainsi que les stratégies et modes de raisonnement développés par l'expert pour la résolution du sous-problème étudié. La phase de maquettage se termine par une procédure d'évaluation et de validation des méthodes et outils utilisés et permet finalement de conclure sur la faisabilité du système. Le but essentiel de la maquette est la comparaison entre les raisonnements du système et ceux de l'expert. Dans l'hypothèse de la faisabilité du système, la maquette débouche sur la définition des objectifs de la phase de prototypage.

Réalisation d'un prototype : cette phase de développement commence par l'identification de tous les sous-problèmes que le système devra être en mesure de résoudre. Le prototypage est une extension ou modification totale de la base de connaissances de la maquette. Contrairement à la maquette, le prototype commence à être doté de modules d'interfaces, d'explication, etc...

Evaluation du prototype : la complétude du système ainsi que son aptitude à répondre aux besoins de l'utilisateur sont évalués selon plusieurs critères tels que la rigueur des raisonnements, la performance, la rentabilité et l'évolution du système, la communication Homme-Machine (évaluation des interfaces).

Mise en service et exploitation industrielle du système : la dernière étape du cycle de vie d'un SBC concerne l'intégration et la mise en exploitation industrielle du système. Il est alors du ressort du groupe de maintenance choisi, d'assurer la vie future du système en collaboration avec les utilisateurs et les experts afin d'acquérir les nouvelles connaissances expertes et les nouveaux besoins informationnels issus de l'utilisation du système.

3.1.2. L'acquisition structurée des connaissances

En alternative à l'approche précédente, les recherches récentes se sont fixé comme objectif de rendre plus efficace la démarche en visant une modélisation totale (représentation intermédiaire) de la connaissance avant son implantation sur machine. Cette approche repose notamment sur l'activité de modélisation du domaine d'expertise, indépendamment du problème de l'implantation. Le cycle de vie d'un SBC construit de la sorte fait généralement intervenir quatre principales étapes : définition du problème, modélisation de la connaissance, conception du système et enfin implantation et tests.

L'analyse de ces deux approches de développement d'un SBC met en évidence, tout au long du cycle de vie du système, l'omniprésence de l'acquisition de connaissances : extraction, structuration, formalisation, conceptualisation et analyse des connaissances recueillies, introduction effective dans la base de connaissances, essai, enrichissement et validation du système. Nous présentons dans les paragraphes suivants les techniques, méthodes et outils d'acquisition de connaissances.

3.2. Techniques de recueil de connaissances

Nous avons présenté dans le paragraphe précédent les différentes étapes de développement d'un SBC. Cependant, un point crucial demeure imprécis : comment capter les connaissances détenues par l'expert en vue d'élaborer la base de connaissances d'un SBC ?

Il existe actuellement diverses techniques, souvent inspirées des travaux en psychologie cognitive, pour accéder et rendre le plus explicite possible les connaissances manipulées par l'expert du domaine. Plusieurs termes sont proposés dans la littérature pour traduire le mot "recueil" de connaissances. On parle des techniques d'élicitation, d'énonciation, d'extraction, d'extériorisation, de capture ou de sollicitation des connaissances. Ces techniques sont généralement qualifiées de méthodes "cogniticiennes", de méthodes "manuelles" ou de méthodes "empiriques". Les techniques possibles pour extraire les connaissances ont été étudiées et présentées par de nombreux auteurs, /MAHE 87/, /NASSIET 87/, /BOY 88/, /GALLOUIN 88/, /HART 88/, /VISSER 88, 90/, /AUSSENAC 89, 91/, /DIENG 90/, /RIALLE 90/, /BENKIRANE 91/ et /BEAUNE 92/.

L'ensemble de ces techniques est souvent classé en deux groupes selon le mode d'accès aux connaissances qu'elles autorisent :

- Les techniques directes reposent généralement sur des méthodes de recueil informelles et s'attachent à extraire directement auprès de l'expert ce qu'il sait exprimer de manière verbale.

- Les techniques indirectes sont des méthodes formelles plus sophistiquées. Elles tentent de révéler les compétences de l'expert à l'aide d'indicateurs interprétables par le cogniticien. Les données acquises par ces techniques sont essentiellement des connaissances utilisées de façon implicite que l'expert ne sait pas formuler explicitement.

Avant d'appliquer une technique, il est fondamental de savoir quel type d'information elle permet de recueillir, d'en connaître les limites de validité et les conditions de mise en œuvre. Il est souhaitable que les cogniticiens disposent de moyens leur indiquant les techniques les mieux adaptées au caractère figé ou évolutif du domaine d'expertise, au mode de raisonnement des experts ainsi qu'aux caractéristiques des connaissances disponibles. En pratique, une seule technique est insuffisante pour extraire la totalité des connaissances et il est nécessaire d'en combiner plusieurs en fonction de la complexité du domaine. Le rôle du cogniticien est non seulement de choisir les techniques adéquates mais aussi de définir leur ordre d'application. Dans le cadre de notre travail concernant le développement d'un SBC pour l'aide à la certification, plusieurs techniques de recueil dont l'interview, les questionnaires, l'analyse de protocole et le tri conceptuel sont utilisées pour extraire des connaissances impliquées dans l'analyse de sécurité .

3.3. Outils d'extraction des connaissances

Afin de pallier l'insuffisance et la complexité d'une démarche empirique de recueil des connaissances, des outils d'extraction ont été développés. Ils visent à automatiser certaines étapes du processus d'acquisition en élaborant eux-mêmes la base de connaissances du SBC. Ces outils, tels par exemple AQUINAS, ROGET, MOLE, sont basés sur l'emploi d'une ou plusieurs techniques de recueil de connaissances évoquées précédemment. Ils sont généralement dédiés à des formes de connaissances ou à des catégories de problèmes particuliers comme le diagnostic ou la conception. Un ensemble d'outils est présenté en détail dans /DIENG 90/ et /BENKIRANE 91/.

3.4. Méthodes d'acquisition des connaissances

Il n'y a pas, à ce jour, de solution complète et générale pour traiter du problème global de l'acquisition des connaissances. Plusieurs constatations issues de la littérature sur l'acquisition de connaissances convergent sur le manque de maturité et de repères théoriques dans la communauté scientifique : ambigüité du vocabulaire, absence de définitions de référence, manque de critères précis de comparaison et d'évaluation des différents travaux. L'acquisition de connaissances est un thème pluridisciplinaire qui bénéficie des apports fructueux de la psychologie cognitive, du génie logiciel et de la linguistique. Elle peut être perçue comme une activité de modélisation qui vise à élaborer un modèle conceptuel (KADS) ou un modèle cognitif de l'expert (KOD, MACAO).

Cependant, devant l'arsenal de ces méthodes, il est difficile de distinguer celle qui répondra le mieux à un besoin particulier. De plus, la plupart des systèmes basés sur ces méthodes sont encore des produits de recherche non commercialisés /DIENG 90/. Pour BOY /88/, ils ne sont encore pour la plupart que des prototypes et réalisent essentiellement de l'édition plutôt que de l'acquisition de connaissances. L'auteur considère que l'acquisition de connaissances demeure une tâche vaste et délicate, liée à une présence humaine. Il s'agit toujours d'un travail d'équipe lourd, difficile à expérimenter, car nécessitant une grande disponibilité des experts. Selon VISSER /88/, les méthodes d'acquisition de connaissances se tarissent rapidement dans certaines situations concrètes, pour deux raisons principales :

- D'une part, les méthodes existantes ont été développées essentiellement pour des problèmes bien structurés.

- D'autre part, ces méthodes partent souvent du principe que l'activité du sujet au moment du traitement est accessible et qu'il est possible d'utiliser des méthodes d'observation directe ou de verbalisation simultanée de l'activité. Or, ce n'est pas toujours le cas, car l'activité peut se dérouler sur un laps de temps trop long pour pouvoir être observée ou verbalisée, ou alors les contraintes temporelles de l'étude sont trop fortes.

Selon AUSSENAC /90/, très peu de systèmes d'acquisition sont directement exploitables par les utilisateurs, en dehors d'outils très spécifiques ou de produits de laboratoire et il n'existe pas de solution individuellement satisfaisante pour la capitalisation de l'ensemble de connaissances d'un domaine. Selon BENKIRANE /91/, la plupart des méthodes d'acquisition de connaissances ne prennent pas en considération les phénomènes imprévisibles rencontrés généralement dans des applications industrielles. Bien que ces méthodes soient très riches du point de vue des modèles de connaissances, elles n'en demeurent pas moins limitées selon BEAUNE /92/ au niveau de leur applicabilité dans des domaines où de nombreux experts sont à consulter et où la connaissance évolue.

Ce paragraphe expose les principales méthodes d'acquisition des connaissances développées au cours des dernières années.

KOD /Vogel 88, 89/

La méthode KOD (Knowledge Oriented Design), marque déposée de CISI Ingénierie, a été conçue à partir des travaux de C. VOGEL /88/. L'outil support de KOD est la K-station, développée et distribuée par Ilog. Dans le développement d'un SBC, l'objectif de KOD est de proposer des cadres de collecte et de modélisation de la connaissance et de permettre le passage de cette connaissance à une information manipulable par la machine. La K-station quant à elle se présente comme une plate-forme logicielle qui offre un ensemble d'outils facilitant la mise en œuvre de la méthode KOD tels que l'hypertexte, les éditeurs graphiques, les algorithmes de validation de la base de connaissances. Très schématiquement, la méthode KOD propose d'enchaîner deux phases complémentaires : une phase d'analyse et de traitement du discours de l'expert et des sources documentaires qui débouche sur l'élaboration du *modèle pratique* et une phase de structuration et de formalisation des connaissances recueillies en vue de concevoir le *modèle cognitif.* La méthode KOD a été utilisée dans plusieurs projets industriels parmi lesquels on peut citer le projet ORIENT /ANGELETTI 91/ destiné à fournir aux consultants du groupe France Télécom les moyens de réaliser des études de réseaux de télécommunications d'entreprises.

MACAO /AUSSENAC 88, 89, 90/

La méthode d'acquisition de connaissances MACAO utilise des techniques de psychologie cognitive pour accéder aux différents aspects de l'expertise. Elle est définie comme une trame pour guider la démarche du cogniticien dans le transfert d'expertise humaine. Cette méthode de recueil et de formalisation des connaissances repose sur un modèle cognitif qui constitue un cadre formel pour la modélisation de la nature et du fonctionnement des connaissances expertes. Ce modèle cognitif s'appuie sur des schémas empiriques (ou procédures) décrivant la connaissance utilisée pour résoudre certains aspects du problème ainsi que sur des schémas conceptuels, plus abstraits, constituant une sorte de "méta-connaissance". A ce modèle sont associées des techniques et méthodes d'élicitation (extraction) permettant de recenser les éléments clés des connaissances expertes. La méthode MACAO se déroule en quatre étapes fondamentales : Identification de l'expertise, recueil puis analyse des protocoles de résolution de problèmes et enfin validation par l'expert. A ce jour, MACAO a fait l'objet d'une maquette de laboratoire qui est en cours de perfectionnement.

KADS /BREUKER 85, 87/

La Méthode KADS (Knowledge Acquisition and Design Support) résulte d'un projet ESPRIT1 et a débouché par la suite sur KADS 2. L'objectif de KADS est de proposer un guide méthodologique pour le processus de développement des SBC. Ce processus est bâti en une construction successive de plusieurs modèles : modèles conceptuels du domaine et du raisonnement de l'expertise, modèle de la coopération entre le futur système et ses utilisateurs, modèle intégrant l'analyse des besoins qui débouche sur le modèle de conception fonctionnelle, modèle logique puis modèle physique du système. L'accent est mis sur le "modèle conceptuel" des connaissances, organisé selon quatre niveaux : le niveau "domaine" qui décrit les connaissances statiques

de l'application, les niveaux "inférence", "tâche" et "stratégie" qui précisent leur exploitation lors des raisonnements. Une bibliothèque de tâches génériques, appelées modèles d'interprétation, a été développée pour permettre l'acquisition structurée de connaissances. La méthode KADS repose sur l'hypothèse que les phases d'analyse et de conception sont séparées : il n'y a pas de retour sur la phase d'analyse, à partir de celle de conception ou d'implémentation. KADS est supporté par un environnement d'outils dénommé SHELLEY que l'on peut qualifier d'atelier de génie cognitif.

CARMEN /TONG 91/

CARMEN est une méthodologie guidant la structuration des connaissances et l'organisation de la résolution de problèmes. Elle répond aux besoins techniques générés par la construction des SBC de seconde génération. CARMEN contribue au domaine de l'acquisition de connaissance par son approche pour la modélisation des connaissances. Elle s'articule autour de cinq niveaux distincts permettant la localisation des connaissances selon leur rôle dans la résolution de problèmes : deux niveaux pour la conception du domaine et trois niveaux pour spécifier la structure de raisonnement. MAC (Modélisation Automatique des Connaissances) est l'outil support de la méthode CARMEN.

OPEN KADS, KADS 2, VITAL, VALID et ACKnowledge sont des projets en cours de réalisation qui s'inscrivent dans le cadre de projets ESPRIT2. Ils visent à approfondir les idées développées dans KADS tout en produisant des avancées sur ses points faibles : formalisation, standardisation de la représentation des connaissances, gestion de grandes bases de connaissances et essentiellement construction et validation d'outils en vue d'une exploitation industrielle des résultats. Parmi ces projets, ACKnowledge /CALABRETTA 90/ et VITAL visent l'amélioration du processus d'acquisition de connaissances en intégrant notamment des techniques d'**apprentissage automatique**. Le système d'apprentissage de règles CHARADE /GANASCIA 87/ est impliqué dans le projet VITAL.

4. MODELE CONCEPTUEL DES SYSTEMES D'INGENIERIE DE CONNAISSANCES DE BENKIRANE /91/ POUR L'ACQUISITION DES CONNAISSANCES DE CERTIFICATION

En présence du domaine de la certification qui est nouveau, complexe, de grandeur réelle, mal structuré, évolutif, basé sur l'intuition et où l'expertise est détenue par 7 experts de certification, l'interactivité directe expert/cogniticien nous semble plus judicieuse. C'est pourquoi notre choix s'est porté sur le modèle de BENKIRANE /91/ qui est particulièrement complet et structuré car il prend en compte quatre dimensions : les phases de développement du SBC, les étapes d'extraction de connaissances, l'environnement du problème et les facteurs humains. L'auteur a proposé une approche méthodologique qui s'attache à organiser le processus d'extraction de connaissances dans le domaine du diagnostic technique. Ces travaux ce sont appuyés sur le développement d'une application industrielle concernant l'aide au diagnostic des causes des défauts dans un procédé sidérurgique.

Le paragraphe suivant présente, dans un premier temps, le modèle conceptuel des systèmes d'ingénierie de connaissances de BENKIRANE /91/ et détaille, dans un deuxième temps, son application au domaine de la certification des systèmes de transport terrestres automatisés.

4.1. Modèle conceptuel des systèmes d'ingénierie de connaissances de BENKIRANE

Le modèle proposé dans /BENKIRANE 91/ fait intervenir quatre dimensions (figure 2.1.).

- La première dimension est composée des étapes d'acquisition de connaissances décrites dans /ROOK 89/.
- La deuxième dimension est constituée de l'environnement du problème proposé dans /GAINES 87/.
- La troisième dimension regroupe les phases de développement d'un SBC proposées dans /NASSIET 87/.
- La dernière dimension s'attache à intégrer les facteurs humains inhérents au processus d'extraction de connaissances.

Ces quatre dimensions sont détaillées dans les paragraphes suivants.

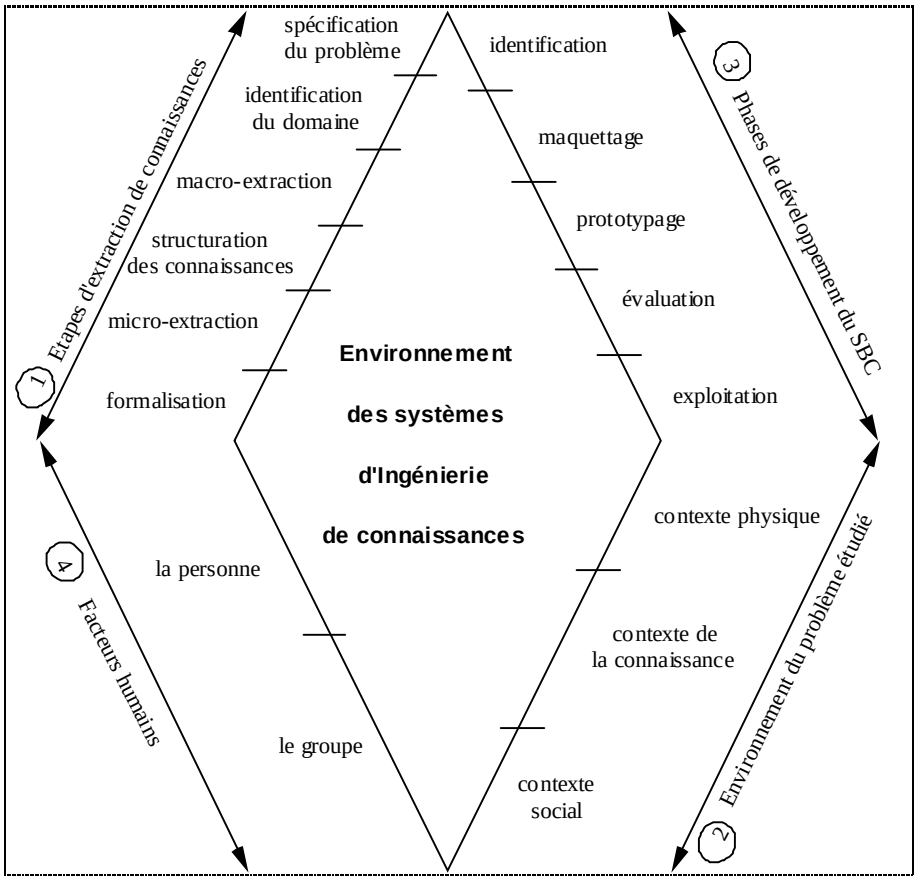


Figure 2.1. : Modèle conceptuel des systèmes d'ingénierie de connaissances

4.1.1. Etapes d'extraction de connaissances

Spécification du problème : cette étape permet de définir le problème et de le situer dans le domaine d'expertise. Elle permet au cogniticien de se faire une première idée du problème et d'acquérir un premier niveau de langage opératoire. Une visite de l'environnement physique du problème à traiter peut être effectuée pour sensibiliser et familiariser le cogniticien au domaine de l'étude. De même, une présentation d'ensemble du problème par l'un des experts permet au cogniticien d'établir les premières spécifications de

l'étude. Dans plusieurs méthodes d'acquisition de connaissances, cette étape de spécification du problème est liée à l'étape d'identification du domaine.

Identification du domaine : cette étape de la méthodologie d'acquisition de connaissances permet de vérifier que le contexte du problème est favorable pour recueillir les connaissances. Elle permet notamment de vérifier l'existence et la disponibilité d'un ou plusieurs experts, la présence d'une expertise en adéquation avec les objectifs de l'étude et la possibilité d'assister à des traitements de cas réels sur le site. L'étape d'identification du domaine permet aussi d'identifier la nature et l'origine des informations utilisées par l'expert ainsi que la spécification de la forme de solution à élaborer par le SBC. Plusieurs techniques de recueil de connaissances telles que l'interview ou les questionnaires sont employées pour identifier le domaine d'expertise. A l'issue de cette étape, le cogniticien est familiarisé avec le vocabulaire du domaine.

Macro-extraction des connaissances : après avoir bien défini le champ de l'étude, l'étape de macro-extraction consiste à identifier les grandes lignes de l'activité de l'expert face à des problèmes réels. Elle s'attache à identifier de manière générale les mécanismes et modes de raisonnement de l'expert, les stratégies et heuristiques de résolution du problème, les relations entre les concepts clés du domaine etc ... Il s'agit d'une description de haut niveau des classes d'objets. La technique d'analyse de protocoles peut être mise en œuvre au cours de cette étape afin d'observer l'expert à l'œuvre.

Structuration des connaissances : les connaissances acquises précédemment sont ensuite analysées et structurées dans des modules. Chacun de ces modules est associé à une tâche de résolution du problème.

Micro-extraction des connaissances : c'est la spécification détaillée des différents modules de connaissances ainsi que l'affinement des concepts utilisés et de leurs relations. L'étape de micro-extraction permet, en effet, d'identifier précisément les mécanismes de raisonnement de l'expert. Pour réaliser cette étape, l'utilisation de plusieurs techniques de recueil de connaissances est nécessaire. La technique d'analyse de protocole peut être appliquée pour acquérir la partie observable de l'activité de l'expert. En revanche, la technique du tri conceptuel peut être utilisée pour identifier l'organisation mentale des concepts manipulés par l'expert. L'étape de micro-extraction débouche finalement sur l'identification des principales facettes du raisonnement de l'expert. Les résultats de cette étape sont notamment les descriptions détaillées de toutes les connaissances (objet, entité) ainsi que les heuristiques de résolution de problèmes.

Formalisation des connaissances : il s'agit ici de transcrire les connaissances issues de la phase de micro-extraction dans un module de représentation adéquat tel que les règles de production, les réseaux sémantiques ou les frames, en vue de leur implémentation.

Les étapes de la méthodologie d'acquisition de connaissances, présentées ci-dessus, s'enchaînent de manière cyclique afin de prendre en compte la variabilité humaine /BENKIRANE 90/.

4.1.2. Environnement du problème

Selon GAINES, l'environnement du problème étudié fait intervenir trois contextes :

- le contexte physique, défini par les objets physiques du domaine qui peuvent être, par exemple, des machines de fabrication ou des outils informatiques.
- le contexte de la connaissance qui permet de décrire l'état des connaissances et son évolution au cours du raisonnement de l'expert.
- le contexte social environnant qui influence l'expert lors de la résolution d'un problème. La présence notamment d'autres experts ou d'autres cogniticiens peut modifier la démarche de l'expert.

4.1.3. Phases de développement du SBC

La méthodologie de développement d'un SBC, impliquée dans le modèle conceptuel des systèmes d'ingénierie de connaissances de BENKIRANE, s'inscrit dans l'approche de prototypage rapide (ou vertical). Elle positionne les cinq phases du cycle de vie d'un SBC proposées par /NASSIET 87/ : identification, maquettage, prototypage, évaluation et exploitation évoquées précédemment.

4.1.4. Facteurs humains

En plus de la méthodologie de développement du SBC, des étapes d'extraction de connaissances impliquées dans le cycle de vie du SBC et de l'environnement du problème étudié, le processus d'extraction de connaissances est tributaire des facteurs humains inhérents à la nature interactive de ce processus /BENKIRANE 91/. En effet, plusieurs individus experts et cogniticiens participent généralement au développement d'un SBC. L'auteur distingue deux acteurs : la personne (expert ou cogniticien) avec sa compétence, sa formation, sa motivation, sa disponibilité et sa personnalité et le groupe avec sa stabilité, la coopération et la confiance entre ses membres.

4.2. Application du modèle conceptuel de BENKIRANE pour l'acquisition de connaissances de certification

Dans le cadre de l'étude de faisabilité d'un SBC d'aide à la certification, nous avons inscrit notre démarche dans l'esprit du modèle proposé par BENKIRANE. Ce travail a permis de soulever des problèmes inhérents à un projet de grande envergure, dans un environnement industriel complexe. Dans ce paragraphe, nous présentons notre approche "de terrain" en insistant plus particulièrement sur les étapes essentielles d'extraction de connaissances : spécification, identification, macro-extraction, structuration, micro-extraction et formalisation des connaissances. Malgré notre souci de respecter la distinction faite par l'auteur au niveau des étapes de spécification et d'identification, il est difficile en pratique de les dissocier.

4.2.1. Spécification du problème de certification

L'étape de spécification du problème, développée dans le premier chapitre de ce mémoire, a finalement débouché sur la définition d'un outil d'aide à la certification des systèmes de transport terrestres automatisés,

permettant aux experts certifieurs de juger l'adéquation des équipements de protection aux normes de sécurité établies dans le cahier des charges imposé aux constructeurs.

4.2.2. Identification du domaine de certification

Cette phase permet d'identifier l'ensemble des données du domaine. Elle recouvre plusieurs objectifs d'ordre technique, économique et organisationnel et tente de cerner les caractéristiques générales du domaine d'expertise en spécifiant les moyens, les objectifs et l'environnement humain. La phase d'identification est une étape critique. Un problème mal défini peut rendre le futur SBC inexploitable par l'utilisateur final. Ce constat motive la collaboration étroite entre experts, cogniticiens et utilisateurs pour spécifier plus précisément le cahier des charges. Dans le cadre du projet de certification, cette phase a conduit à :

- la définition d'une structure de conduite du projet,
- et à l'identification :
  - des sources de connaissances de certification,
  - des types de connaissances employées pour la certification,
  - des caractéristiques des connaissances manipulées et des modes de raisonnement des experts certifieurs,
  - des propriétés essentielles d'un SBC d'aide à la certification,
  - du domaine de compétence du SBC.

4.2.2.1. Structure de conduite du projet SBC

L'élaboration d'un projet SBC fait intervenir un certain nombre d'acteurs. Il est donc souhaitable de discerner les différents groupes ainsi que leur rôle dans le projet. La désignation des participants au projet a abouti à la constitution des différents groupes de travail :

- le comité directeur qui prend les décisions et joue un rôle d'arbitre entre les autres groupes. Il est constitué de deux responsables de l'INRETS-CRESTA et de deux responsables du LAIH de Valenciennes.
- le groupe des experts qui fournit l'essentiel des connaissances est composé d'un expert du CRESTA de Lille et d'un expert de l'INRETS d'Arcueil.
- le groupe des cogniticiens chargé du recueil et de la formalisation de connaissances est formé de deux chercheurs du LAIH de Valenciennes.
- le groupe de développement est constitué d'un chercheur du laboratoire, occasionnellement assisté d'étudiants en projets de fin d'études.
- le groupe des utilisateurs à qui est destiné le système est constitué de jeunes experts de certification de l'INRETS-CRESTA.
- le groupe de maintenance qui assure le bon fonctionnement du futur SBC est attribué à un expert de certification du CRESTA.

4.2.2.2. Les sources de connaissances de certification

Les connaissances de certification et notamment les connaissances relatives à l'analyse de sécurité des systèmes de transport, proviennent des trois sources suivantes :

- les ingénieurs concepteurs qui fournissent dans leurs dossiers de sécurité les connaissances structurelles, fonctionnelles et d'analyse des risques d'accidents à travers des scénarios d'accidents,
- les experts de certification de l'INRETS-CRESTA qui détiennent les connaissances expérimentales et le savoir-faire (méthodologies, stratégies, heuristiques ...) : 2 experts au CRESTA de Lille et 5 experts à l'INRETS d'Arcueil (Paris),
- les historiques tels que les archives, manuels, comptes-rendus de réunions, base de données ... .

4.2.2.3. Les types de connaissances employées pour la certification

Deux types de connaissances sont manipulés par les experts pour certifier un nouveau système de transport :

- Le premier regroupe des connaissances historiques de sécurité qui reposent notamment sur l'analyse des risques d'accidents. Ces connaissances sont en grande partie statiques car issues des historiques provenant des systèmes de transport existants.
- Le second englobe les connaissances de certification traduisant la démarche et les stratégies employées par les certifieurs afin d'évaluer le contenu du dossier de sécurité proposé par le constructeur du système de transport. Ces connaissances sont fondées essentiellement sur l'imagination et l'élaboration des scénarios d'accidents. En ce sens elles sont dynamiques et nuancées d'une part de subjectivité.

4.2.2.4. Caractéristiques des connaissances de certification et mode de raisonnement des experts certifieurs

Les connaissances détenues par les experts de certification concernent essentiellement l'analyse des risques d'accidents ou des situations dangereuses sur le plan de la sécurité, décrites sous forme de scénarios d'accidents. La manipulation de ces connaissances fait appel à différents modes de raisonnements que l'on peut qualifier ainsi :

Raisonnement incertain *:* Le raisonnement est un processus de traitement qui produit des connaissances nouvelles à partir de connaissances acquises. Le critère important qui permet de classifier les raisonnements est le degré de certitude des connaissances. C'est ainsi que GRUNDSTEIN /88/ distingue trois types de raisonnement :

- les raisonnements rigoureux qui procèdent du certain au certain,
- les raisonnements non rigoureux qui procèdent du certain au plausible,
- les raisonnements incertains qui procèdent du plausible au plausible.

Le raisonnement des experts de certification peut être qualifié d'"incertain". Dans le cadre de l'analyse de sécurité d'un nouveau STA, le constructeur propose un ensemble de situations d'insécurité plausibles. A partir de ces scénarios d'accidents potentiels, l'expert certifieur imagine de nouvelles situations d'insécurité qui ne peuvent être que plausibles.

Raisonnement empirique : le raisonnement impliqué dans la démarche de certification peut être qualifié d'empirique. Il s'appuie, en effet, sur l'expérience vécue de situations concrètes d'insécurité rencontrées par les systèmes de transport déjà certifiés ou homologués.

Raisonnement intuitif : un des problèmes cruciaux de la certification réside dans la simulation de raisonnements intuitifs issus du bon sens des experts certifieurs lors de l'imagination de nouveaux scénarios d'accident. Ce raisonnement intuitif qui, selon Bonnet /86/ résulte d'une expérience accumulée, est difficilement justifiable. En effet, l'expert certifieur peut avoir le sentiment qu'une panne ou défaillance risque de mettre en cause la sécurité du système ou des voyageurs, sans pour cela être capable de justifier son raisonnement et de décrire clairement les séquencements d'événements conduisant à l'état d'insécurité.

Raisonnement qualitatif : les connaissances manipulées dans le raisonnement de vérification et de certification des systèmes de transport à conduite automatique sont essentiellement de nature qualitative plutôt que quantitative. En effet, l'analyse des risques d'accidents nécessite le traitement d'une quantité importante de connaissances symboliques en majorité qualitatives.

Raisonnement évolutif : le domaine de certification est non borné et instable, du fait de l'évolution de la technologie des systèmes de transport à conduite automatique. En ce sens, les connaissances relatives à l'analyse de sécurité de systèmes évoluent constamment ce qui rend nécessaire leur mise à jour périodique.

Raisonnement non unique : le problème de certification ne connaît pas de solution théorique, unique et explicite. Plusieurs cheminements sont possibles car l'activité de certification se compose d'un enchaînement de plusieurs tâches élémentaires de natures différentes. Le certifieur fait appel en permanence au diagnostic, à l'imagination, à l'intuition, à la prédiction ... Ces raisonnements font appel à des connaissances et concepts difficilement identifiables ainsi qu'à des modes de raisonnement difficilement reproductibles et formalisables.

4.2.2.5. Récapitulatif des propriétés essentielles d'un système d'aide à la certification

Le cahier des charges s'est focalisé sur l'étude de faisabilité d'un SBC d'aide à la certification des systèmes de transport terrestres automatisés, répondant aux objectifs suivants /LE TRUNG 92/ :

- évaluer puis confirmer la complétude des scénarios d'accidents fournis dans le dossier de sécurité du constructeur.
- générer de nouveaux scénarios susceptibles de parfaire l'exhaustivité de l'analyse des risques d'insécurité.

Ce système d'aide à l'analyse de sécurité et par conséquent à la certification doit satisfaire les principales aspirations suivantes, révélées lors de la phase d'identification du problème /HADJ-MABROUK 91b/ :

Avoir une approche plus fiable pour l'examen de la complétude du dossier de sécurité du constructeur et donc améliorer la qualité de décision

Compte tenu du caractère automatique du système de transport mettant en jeu la sécurité des voyageurs, la tâche de certification pèse largement sur les experts certifieurs et doit être réalisée avec une attention accrue. En ce sens, l'analyse des risques d'accident doit être perçue comme exhaustive avant qu'un avis de certification favorable soit accordé. Juger de l'exhaustivité de l'analyse des risques d'accident présentés sous forme de scénarios d'accidents dans le dossier du constructeur, constitue une tâche fondamentale de la certification des STA. Dans ce contexte, l'essentiel de la tâche de l'expert certifieur consiste à imaginer de nouveaux scénarios d'accidents susceptibles soit de démontrer le caractère exhaustif de l'analyse de sécurité effectuée par le constructeur soit de la contredire. En ce sens, améliorer la qualité de décision de certification par l'évaluation et la génération des situations d'insécurité représente une aide substantielle pour les experts certifieurs.

Définir un formalisme précis, rigoureux et explicite pour la représentation des scénarios d'accidents

L'intérêt d'élaborer un modèle de représentation des connaissances qui sont en grande partie des scénarios d'accidents, est majeur. Cela permet d'une part, d'aider les experts de certification à mieux structurer leurs connaissances et d'autre part de proposer éventuellement aux constructeurs des systèmes de transport, un cadre pour une définition exhaustive des scénarios d'accidents.

Aider à rechercher des similarités et des analogies entre plusieurs configurations ou situations d'insécurité

En présence d'une situation de sécurité dangereuse décrite sous forme de scénario par le constructeur, l'expert de certification raisonne par analogie. Il tente de rapprocher cette nouvelle situation d'insécurité de certaines situations vécues sur des équipements ou systèmes analogues déjà certifiés ou homologués. Pour faire cette analyse, le certifieur recherche des parallèles entre le cas qui lui est soumis et l'ensemble des cas typiques simulés ou vécus auxquels il a déjà été confronté.

Pérenniser l'expertise du domaine

L'expertise existante en matière de certification et d'analyse de sécurité des systèmes est rare, vulnérable et répartie entre plusieurs certifieurs (7 experts). Comme elle n'est ni consignée dans des manuels, ni enseignée, les experts de certification souhaitent l'approfondir en vue de conserver et de diffuser au mieux ce capital intellectuel auprès des jeunes experts. Bonnet /86/ constate : "combien de connaissances sont-elles perdues parce qu'elles ne sont pas intégrées systématiquement dans une base, sous forme pratique". Cette base de connaissances doit être évolutive et enrichie au fur et à mesure de la certification des nouveaux STA car on ne peut raisonnablement pas attendre des certifieurs qu'ils délivrent d'emblée et exhaustivement leurs connaissances.

4.2.2.6. Domaine de compétence du système d'aide à la certification

La plupart des connaissances relatives au domaine de certification sont liées à la sécurité des systèmes de transport et découlent pour l'essentiel de l'analyse des risques d'insécurité représentés sous forme de scénarios d'accidents potentiels. Ces derniers sont conçus par les Ingénieurs concepteurs et les experts de certification de l'INRETS-CRESTA. L'élaboration d'un scénario d'accident s'inspire notamment des historiques relatifs aux systèmes de transport déjà certifiés et/ou homologués. L'extraction et la formalisation de l'ensemble des scénarios éprouvés à ce jour est un travail conséquent. En effet, la sûreté de fonctionnement des systèmes de transport automatisés requiert la prise en compte de tous les risques (collision, déraillement, électrocution ...) auxquels peuvent être associés de nombreux scénarios. Dans le cadre d'une étude de faisabilité, nous avons volontairement limité les développements à un seul risque : la collision. Néanmoins, l'architecture du système réalisé est ouverte et pourra donc accueillir d'autres risques.

4.2.2.7. Bilan de la phase d'identification

La phase d'identification est apparue comme une étape primordiale pour cerner les problèmes liés au domaine de la certification. Elle a permis d'en dégager les principaux aspects ainsi que l'environnement humain et technique et de cadrer le projet sur les deux aspects suivants :

- étude de faisabilité d'un système d'aide à la certification,
- focaliser en première approche sur le risque de "collision".

Plusieurs techniques de recueil de connaissances ont été utilisées pour identifier le domaine de certification. Par exemple, la technique d'interview a été employée pour identifier la nature et l'origine des informations manipulées par les experts certifieurs ainsi que leur mode de raisonnement pour certifier un nouveau système de transport automatisé. Finalement, la phase d'identification a permis au cogniticien de se familiariser avec le vocabulaire du domaine de certification et d'établir un lexique des concepts impliqués dans la certification. Ce lexique a été mis à jour tout au long du projet. Le problème étant bien identifié, nous allons aborder l'étape de macro-extraction dont la finalité est de faire émerger les caractéristiques essentielles de la démarche experte.

4.2.3. Macro-extraction des connaissances

La Macro-extraction des connaissances permet d'identifier les grandes lignes de l'activité de l'expert face à des problèmes réels. Afin d'acquérir la "trame" générale de résolution d'un problème par l'expert certifieur, trois techniques de recueil de connaissances ont été utilisées : l'interview, le questionnaire et l'analyse de protocole. La Macro-extraction a permis de mettre en évidence la démarche générale experte d'évaluation du degré de sécurité d'un nouveau système de transport terrestre automatisé. Le principe général de la stratégie utilisée par les experts certifieurs est illustré en figure 2.2.

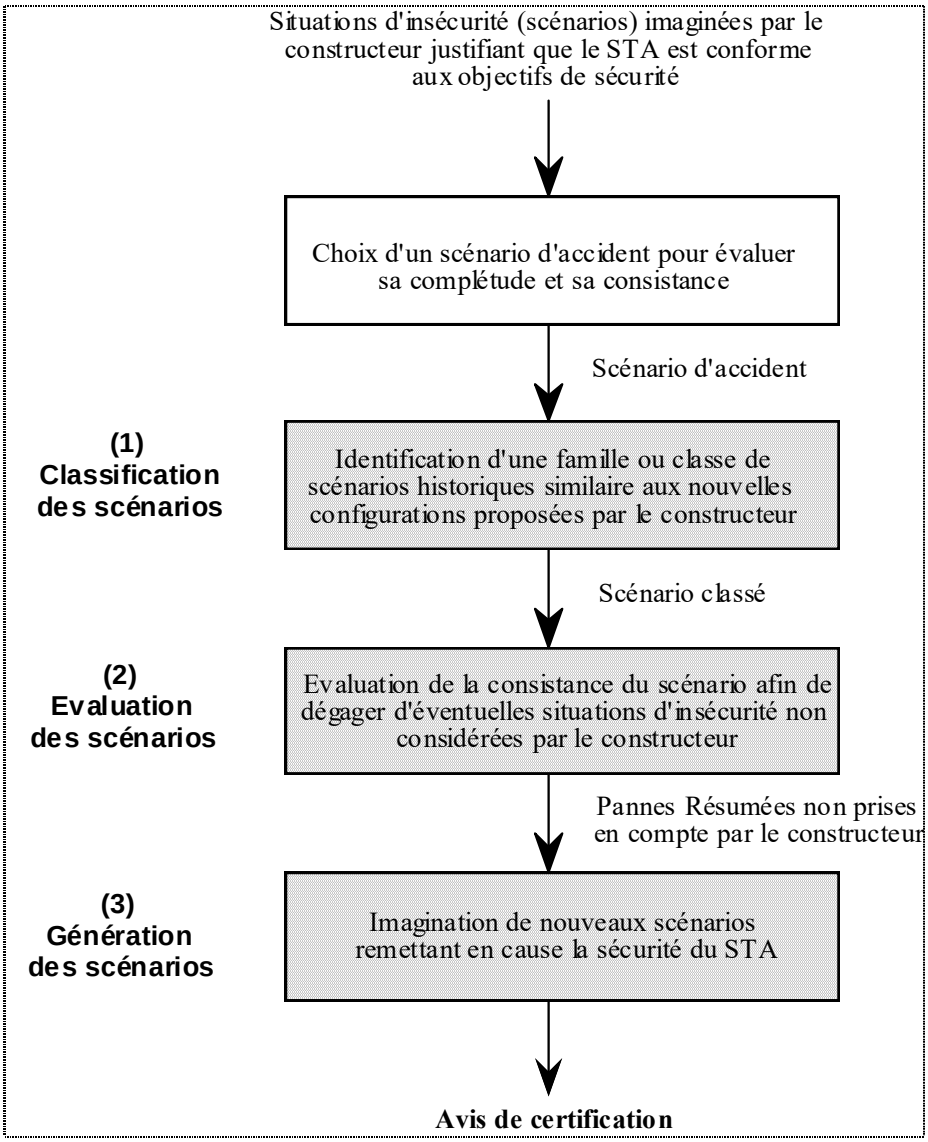


Figure 2.2. Démarche experte d'évaluation du degré de sécurité d'un STA

La démarche de certification fait appel à une première phase de reconnaissance qui apparente le scénario étudié à une famille de scénarios connue de l'expert, d'où la nécessité de disposer de classes de scénarios. Dans une seconde phase, le certifieur s'attache à évaluer la consistance de ce scénario afin de dégager d'éventuelles situations d'insécurité non considérées par le constructeur. De telles situations stimulent l'expert pour la formulation de nouveaux scénarios d'accident. L'analyse de la démarche experte de certification fait ressortir la succession de trois grandes tâches pour l'évaluation de la complétude et de la consistance du dossier de sécurité proposé par le constructeur des systèmes de transport automatisés /HADJ-MABROUK 91b/ :

1. une tâche de classification des scénarios d'accidents,
2. une tâche d'évaluation des scénarios d'accidents,
3. une tâche de génération des scénarios d'accidents.

Dès lors, l'extraction des connaissances relatives à l'ensemble des concepts fondamentaux impliqués dans un scénario d'accident, est indispensable. Les connaissances identifiées lors de cette étape de macro-extraction ont été organisées dans l'étape de structuration présentée ci-après.

4.2.4. Structuration des connaissances

Les connaissances relatives aux concepts fondamentaux impliqués dans les risques d'accidents des systèmes de transport terrestres automatisés ont été structurées en une hiérarchie (figure 2.3), organisée selon six niveaux : types d'accidents, risques d'accidents, fonctions de sécurité, sous-fonctions de sécurité (classes d'accidents), scénarios d'accidents et paramètres descriptifs d'un scénario. Cette hiérarchie a fait l'objet d'une validation par les experts de certification.

Afin de maîtriser la complexité de la démarche de certification et d'en appréhender les problèmes de sécurité, la stratégie adoptée par les experts a été formalisée dans un diagramme SADT (Structured Analysis and Design Techniques) figure 2.4. Le cœur de la démarche "traiter une fonction de sécurité" est ensuite développé en figure 2.5.

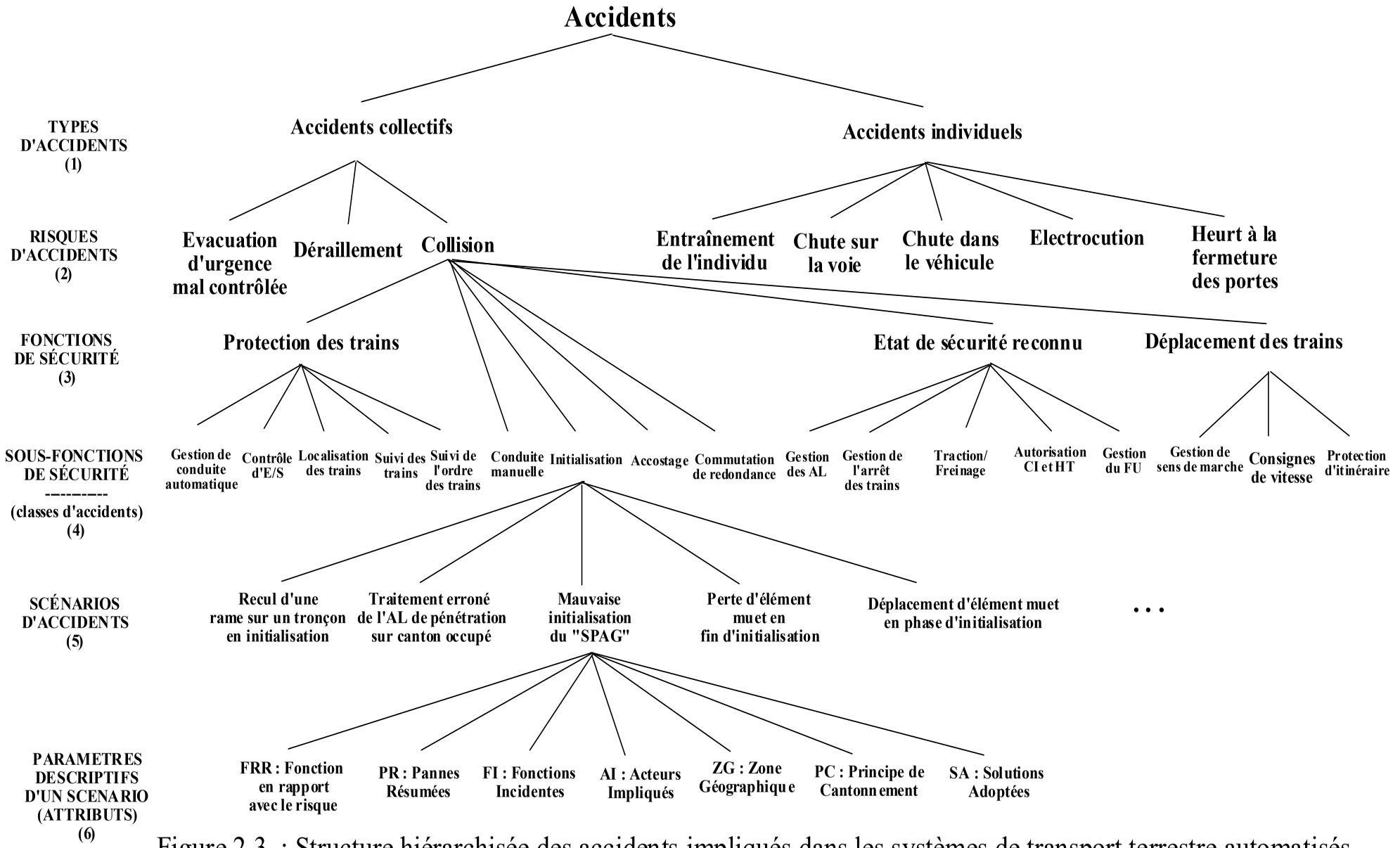


Figure 2.3. : Structure hiérarchisée des accidents impliqués dans les systèmes de transport terrestre automatisés

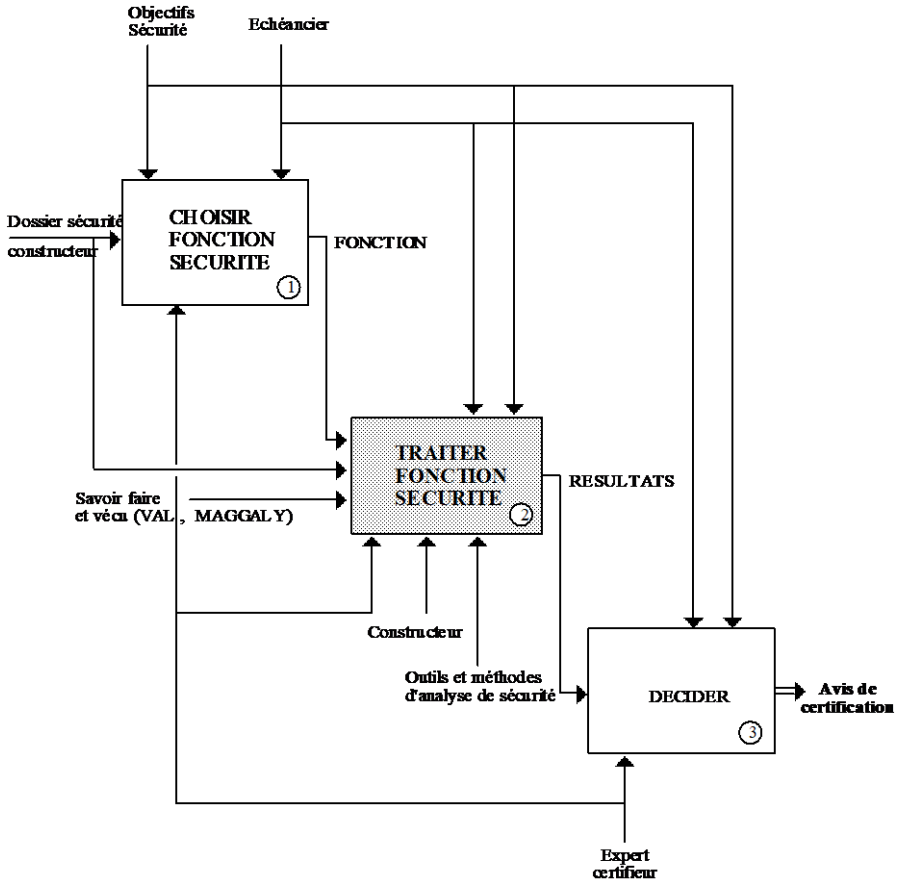


Figure 2.4 : Démarche générale de certification (niveau AO)

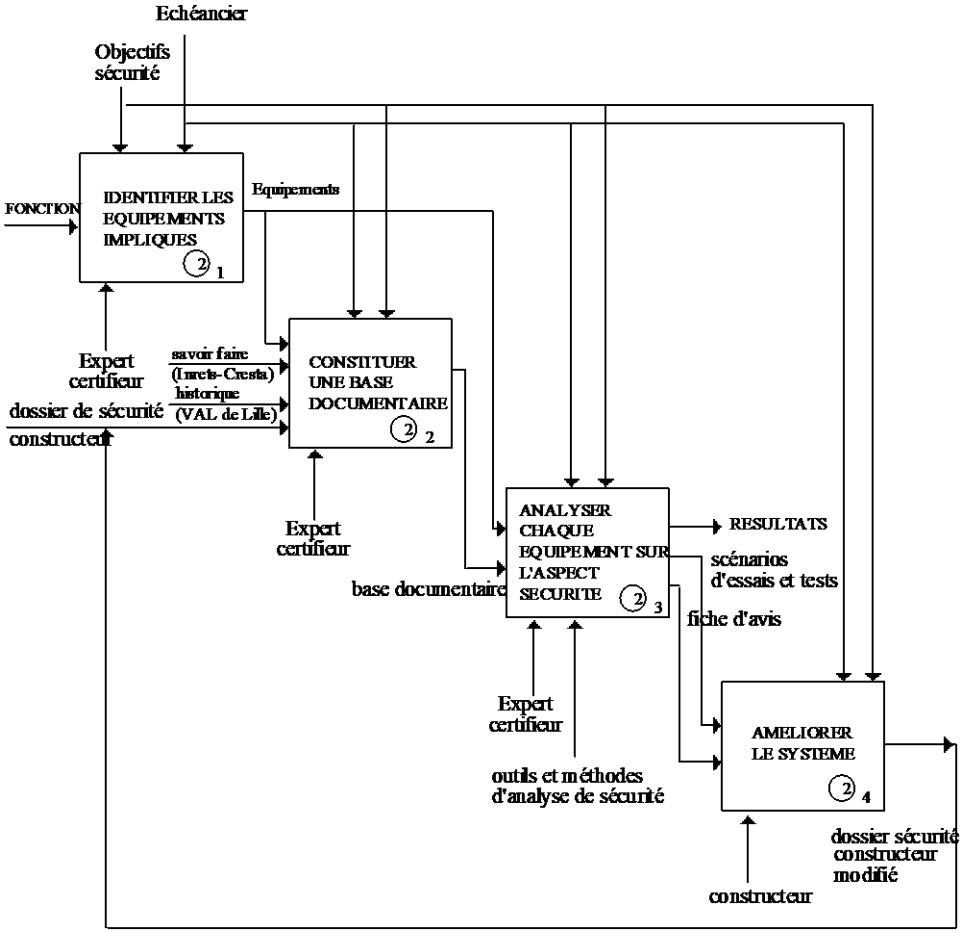


Figure 2.5 : Traitement d'une fonction de sécurité (niveau A1)

4.2.5. Micro-extraction des connaissances

La Micro-extraction des connaissances vise à extraire d'une manière détaillée les mécanismes de raisonnement de l'expert, les stratégies de résolution du problème et les relations entre les concepts clé du domaine d'expertise /BENKIRANE 91/. Cette étape débouche en final sur l'identification des principales facettes du raisonnement des experts. Malgré tous les soins apportés à cette étape et le nombre important de sessions d'extraction de connaissances (une trentaine de sessions) ainsi que l'emploi de plusieurs techniques de recueil de connaissances (interviews, questionnaires, analyse de protocoles, tri conceptuel ...), la Micro-extraction a débouché uniquement sur une caractérisation précise des scénarios d'accidents (ensemble des paramètres descriptifs) avant de se heurter à la difficulté de faire émerger les stratégies, heuristiques et mécanismes de raisonnement des experts de certification. Cet obstacle à l'acquisition de la démarche experte de certification est dû notamment au caractère intuitif et évolutif de la connaissance. Nous présentons ci-dessous les résultats de cette étape concernant l'analyse et la caractérisation d'un scénario d'accident.

4.2.5.1. Caractérisation d'un scénario d'accident

Un scénario d'accident décrit un concours de circonstances qui peut conduire à une situation non désirable voire dangereuse. Il est caractérisé par un contexte et un ensemble d'événements et de paramètres. La phase de Micro-extraction des connaissances a débouché sur la mise en forme d'un modèle basé notamment sur l'identification des huit paramètres décrivant un scénario d'accident (figure 2.6.).

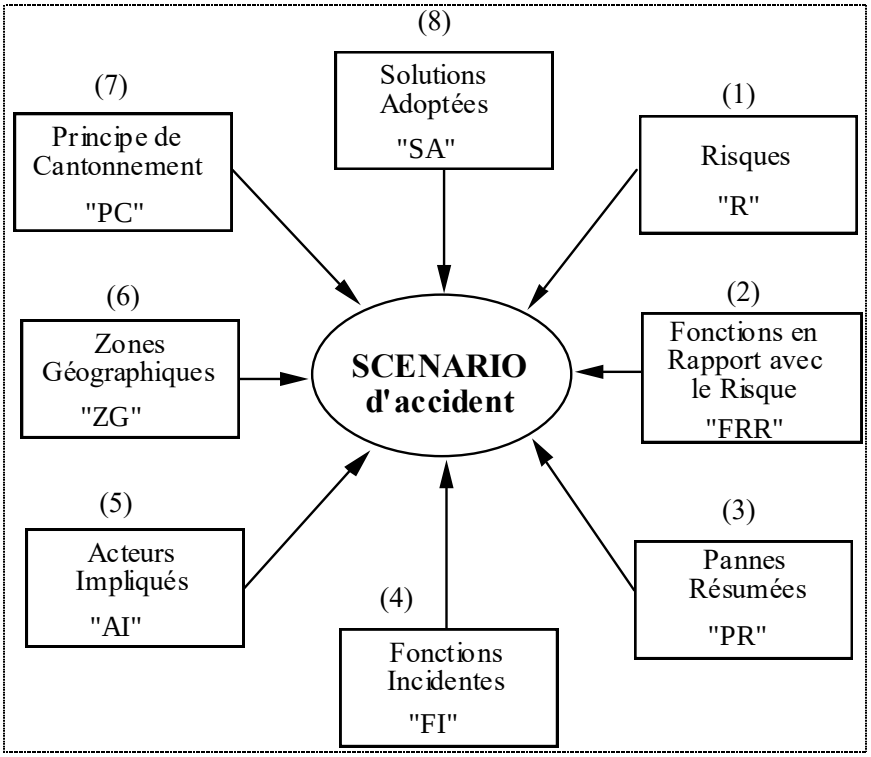


Figure 2.6 : paramètres descriptifs d'un scénario d'accident

4.2.5.2. Définition des paramètres descriptifs d'un scénario

Risque (R) *:* le risque est la mesure du danger. Il est caractérisé par le couple "gravité des conséquences" et "probabilité d'occurrence". Dans notre contexte, on utilise le terme de risque pour désigner une circonstance qui peut conduire à un dommage corporel. Dans le domaine de la sécurité des systèmes de transport automatisés, plusieurs risques potentiels ont été recensés tels que la collision, le déraillement, l'électrocution.

Fonctions en Rapport avec le Risque (FRR) *:* ce sont des fonctions de protection destinées à rendre le risque nul ou acceptable pour l'usager. Un risque particulier peut faire intervenir plusieurs sous-fonctions du système qu'elles soient de protection ou purement fonctionnelles. Par exemple, les fonctions commutation de redondance, accostage, localisation des trains, sont relatives au risque de collision.

Pannes Résumées (PR) *:* une Panne Résumée est une panne **générique** résultant du groupement d'un ensemble de pannes élémentaires ayant la même conséquence sur le comportement du système (cf chapitre I, paragraphe 3.2.4.). Chaque scénario fait intervenir une ou plusieurs PR. Une liste des PR de l'ensemble des scénarios acquis à ce jour relatives au risque de collision a été établie. Un échantillon de quelques PR est donné ci-dessous :

PR1 : Pénétration par recul d'une rame sur un canton occupé
PR2 : Rame en panne d'émetteur anticollision
PR3 : Masquage d'alarme par Initialisation
PR7 : Maintien d'un Itinéraire à tort par le Pilote Automatique en attente
PR24 : Panne de traction permanente
PR32 : Commutation inopinée de mode E/S

Fonctions Incidentes (FI) *:* ce sont des fonctions liées à l'exploitation du système qui peuvent favoriser l'incidence d'un scénario agissant sur la sécurité du système. Ces fonctions peuvent jouer un rôle de catalyseur. On a pu distinguer quatre fonctions incidentes fondamentales : gestion des itinéraires, régulation de trafic, transmission, consignes d'exploitation (consistance des consignes et vigilance des opérateurs).

Acteurs Impliqués (AI) *:* on a pu déterminer une liste des acteurs impliqués dans l'ensemble des scénarios d'accidents tels que les rames, l'opérateur au Poste de Commande Centralisé (PCC), le Pilote Automatique (PA) avec redondance, etc.

Zones Géographiques (ZG) *:* les cinq principales zones géographiques recensées sont : Terminus, station, ligne, zone d'injection de rame, limite de tronçon.

Principe de Cantonnement (PC) *:* on distingue actuellement deux principes de cantonnement dans les systèmes de transport automatisés : canton fixe ou canton mobile.

Solutions Adoptées (SA) *:* le constructeur propose pour chaque scénario qui met en défaut la sécurité, plusieurs solutions dont une est réalisée. Une liste des SA relatives à l'ensemble des scénarios acquis a été établie. On peut en citer l'extrait suivant :

SA1 : Faire intervenir la position d'aiguillage

SA4 : Faire le bilan des rames présentées en ligne

SA7 : Augmenter la longueur du Canton Mobile Déformable

SA10 : Le canton amont ne doit être libéré que si l'aval est occupé

SA14 : Interdire l'effacement d'un élément en mode "Train Secouru"

4.2.6. Formalisation des connaissances

Dans cette dernière étape de la méthode d'acquisition de connaissances, le cogniticien formalise les connaissances issues de la Micro-extraction dans un mode de représentation adéquat. Pour la certification, cette étape a débouché sur l'élaboration d'un modèle "générique" pour la représentation des scénarios d'accidents.

L'analyse du concept de "Scénario" a révélé deux aspects fondamentaux. Le premier, **statique**, permet de caractériser le contexte, et le second, **dynamique**, met en évidence les possibilités d'évolution dans ce contexte tout en soulignant le cheminement qui débouche sur une situation d'insécurité /HADJ-MABROUK 91a/. Le formalisme retenu pour la description statique est celui d'une fiche dans laquelle plusieurs paramètres descriptifs essentiels sont décrits en termes de paires attribut/valeur. Pour la description dynamique nous avons repris le formalisme déjà utilisé des Réseaux de Pétri tout en le structurant de façon à mettre en évidence trois composantes essentielles : l'environnement externe du système (partie opérative), l'environnement interne (partie commande) et l'interface qui assure la communication entre les deux environnements. Le Réseau de Pétri permet de focaliser l'étude sur un environnement particulier du STA pour lequel diverses séquences d'animation peuvent être envisagées. Chacune de ces séquences correspond à un scénario et est décrite sous forme d'un tableau dit "*Tableau de Séquencement du Marquage*".

4.2.6.1. Description statique d'un scénario d'accident

La description statique recense les principaux paramètres d'un scénario d'accident, dans une "fiche descriptive" élaborée sous forme d'un tableau de données (Attributs/Valeurs). Les attributs correspondent aux huit paramètres et à chaque attribut est associée une liste des valeurs possibles. Cette fiche a par la suite été utilisée comme formulaire de base pour l'acquisition des scénarios ; elle se veut exhaustive et permet d'aborder l'étude de tout type de système de transport automatisé (VAL, MAGGALY, POMA). En résumé, la description statique d'un scénario a débouché sur la définition d'un premier langage de description des exemples de scénarios. Il s'agit d'une représentation classique par des couples Attribut-Valeur. Ce langage d'expression est proche du langage de l'expert certifieur et a l'avantage d'être compatible avec la structuration des données historiques de sécurité.

4.2.6.2. Description dynamique d'un scénario d'accident

La description dynamique repose sur l'emploi de deux modes de représentation d'un scénario : le Réseau de Pétri et le Tableau de Séquencement du Marquage.

Le **réseau de Pétri**, outil formel de spécification et de description des systèmes a été utilisé pour modéliser le comportement du ou des équipements du système de transport mis en cause par le scénario. Pour tenir compte des différents éléments constituant le scénario modélisé, on dissocie trois aspects :

- Le premier aspect décrit l'environnement externe du système (par exemple, le passage d'une rame d'une station à une autre).
- Le second aspect tient compte de l'environnement interne, à savoir les automatismes tels que les pilotes automatiques, les alarmes, ...
- le troisième décrit l'interface entre ces deux types d'environnement qui assure l'échange d'informations.

Modéliser un scénario d'accident par un Réseau de Pétri permet ensuite par simulation d'examiner toutes les situations remarquables jugées dangereuses ou indésirables sur le plan de la sécurité. En pratique, le Réseau de Pétri élaboré pour un scénario, permet une étude plus exhaustive des diverses situations d'insécurité possibles. Chacune de ces situations est décrite sous la forme d'un tableau dit "de séquencement du marquage".

Le **Tableau de Séquencement du Marquage**, présente la chronologie des événements précédant l'altération partielle ou totale des conditions de sécurité. Il est caractérisé par trois concepts fondamentaux :

- un état initial (marquage initial),
- une chronologie de situations,
- une situation critique (c'est en quelque sorte un pré-accident car elle annonce l'accident inévitable).

4.2.6.3. Exemple de scénario d'accident formalisé

Afin d'illustrer l'ensemble des concepts évoqués dans les paragraphes précédents, un exemple de scénario relatif à la commutation de redondance est représenté par :

- une description textuelle,
- un synoptique (figure 2.7),
- une fiche descriptive des paramètres (figure 2.8),
- un Réseau de Pétri (figure 2.9),
- un tableau de séquencement du marquage (figure 2.10).

Description textuelle du scénario :

Afin d'améliorer la disponibilité des systèmes de transport automatiques, il est d'usage de doubler certains équipements, l'un étant actif, l'autre passif et capable de remplacer le premier en cas de défaillance. Le scénario présenté ici concerne la mémorisation à tort d'itinéraires incompatibles sur deux pilotes automatiques (PA) redondants dans un terminus (figure 2.7.). Il s'agit d'une rame 1 ayant rebroussé chemin et se trouvant à la station de départ sous le contrôle du PA actif (PA A). Une rame 2 arrive suivant l'itinéraire I1. Suite à une défaillance, le PA redondant (PA B) n'a pas détruit l'itinéraire I2 alors que ce dernier l'a été sur le PA A, ce qui provoque une alarme de discordance PA. A la suite de cette alarme, le poste central de contrôle (PCC) pourrait commuter le PA A sur le PA B. Ce dernier autoriserait le rebroussement de la rame 2, ce qui pourrait engendrer une collision.

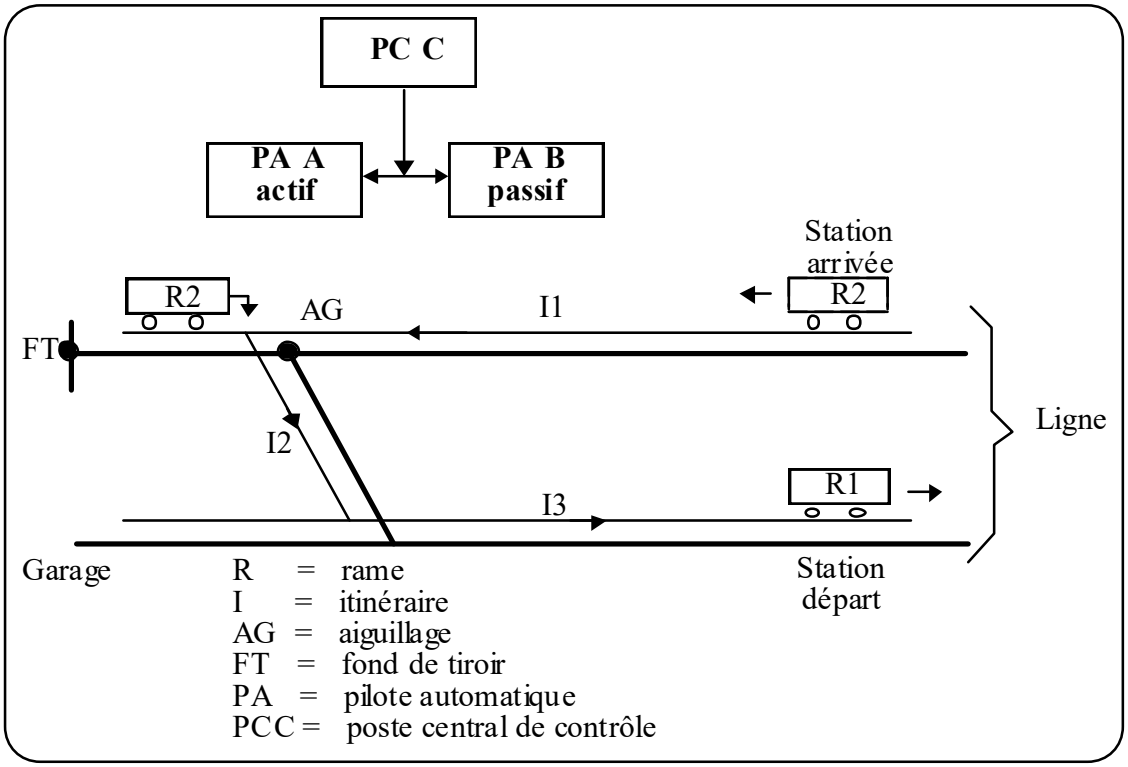


Figure 2.7 : Synoptique du scénario

	LISTE DES ATTRIBUTS	LISTE DES VALEURS POSSIBLES		Valeurs(X) choisies	concepts(*) clés
1	PRINCIPE DE CANTONNEMENT	Canton fixe		X	
		Canton mobile			
2	RISQUES	Collision		X	*
		Déraillement			
		Evacuation d'urgence mal contrôlée			
		Chute dans le véhicule			
		Chute sur la voie			
		Entraînement de l'individu			
		Electrocution			
		Heurt à la fermeture des portes			
3	FONCTIONS EN RAPPORT AVEC LE RISQUE	Gestion de conduite automatique			
		Localisation des trains		X	
		Contrôle d'entrée/sortie			
		Suivi des trains			
		Gestion de sens de marche		X	
		Consigne de vitesse			
		Gestion de l'arrêt des trains		X	
		Sécurité Quai-Voie			
		Autorisation CI/HT			
		Commutation de redondance		X	*
		Initialisation			
		Conduite manuelle			
		Gestion des alarmes		X	
		Evacuation			
		Accostage			
		Protection des itinéraires			
		Traction/Freinage			
4	ZONES GEOGRA-PHIQUES SPECIFIQUES AU SCENARIO	Terminus		X	*
		Station			
		Ligne			
		Zone d'injection de rame			
		Limite de tronçon			
5	ACTEURS IMPLIQUES DANS LE SCENARIO	Nombre de rames		2	*
		Opérateur au PCC		X	*
		Opérateur itinérant			
		PA avec redondance		X	*
		PA sans redondance			
6	FONCTIONS INCIDENTES	Gestion des itinéraires		X	*
		Régulation de trafic			
		Consignes (consistance, vigilance)		X	*
		Communication (transmission)			
7	PANNES RESUMEES ENGENDREES PAR LE SCENARIO	PR7	Maintien d'un Itinéraire à tort par le PA en attente	X	*
		PR8	Commutation à tort du PA fixe par le PCC	X	*
8	SOLUTIONS PROVISOIRES	SP	Isoler un PA	X	
	SOLUTIONS OPTIMALES	SO	Faire intervenir la position d'aiguillage	X	

Figure 2.8 : Fiche descriptive des paramètres

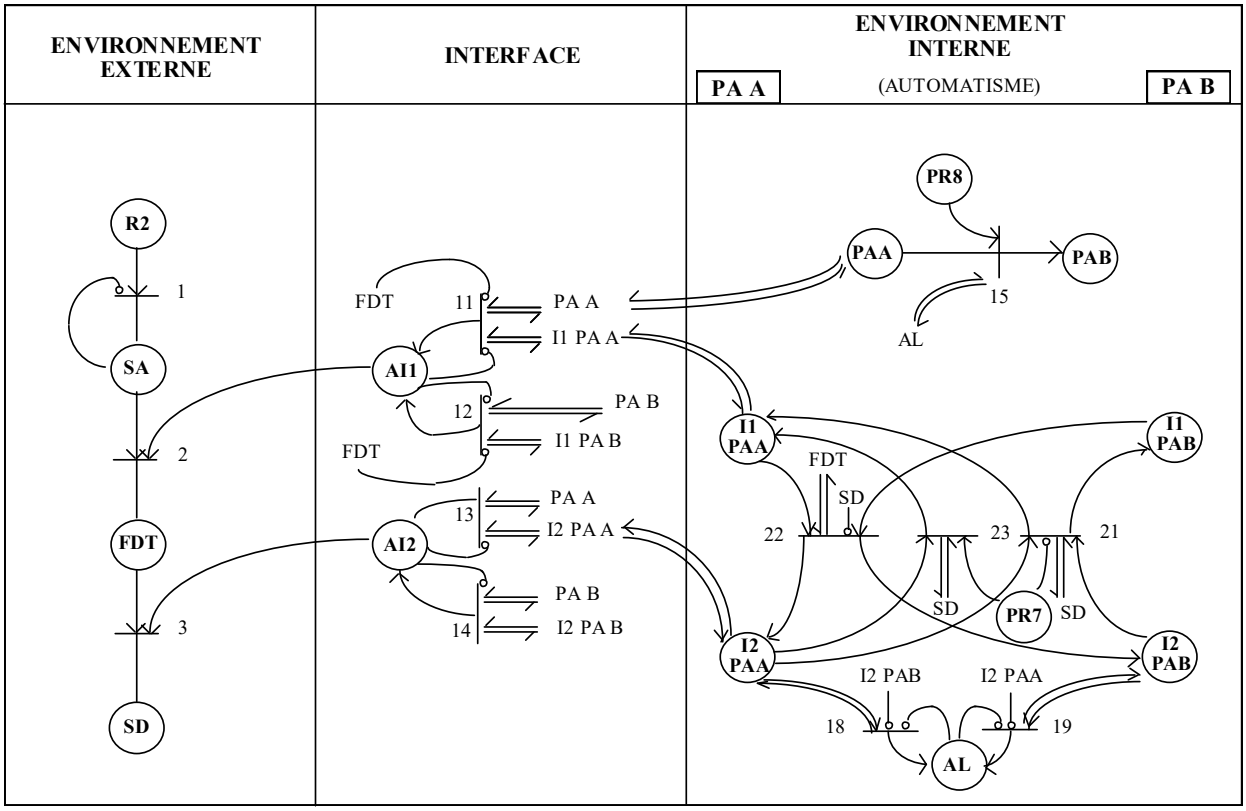


Figure 2.9 : Modélisation du scénario par R. Pétri

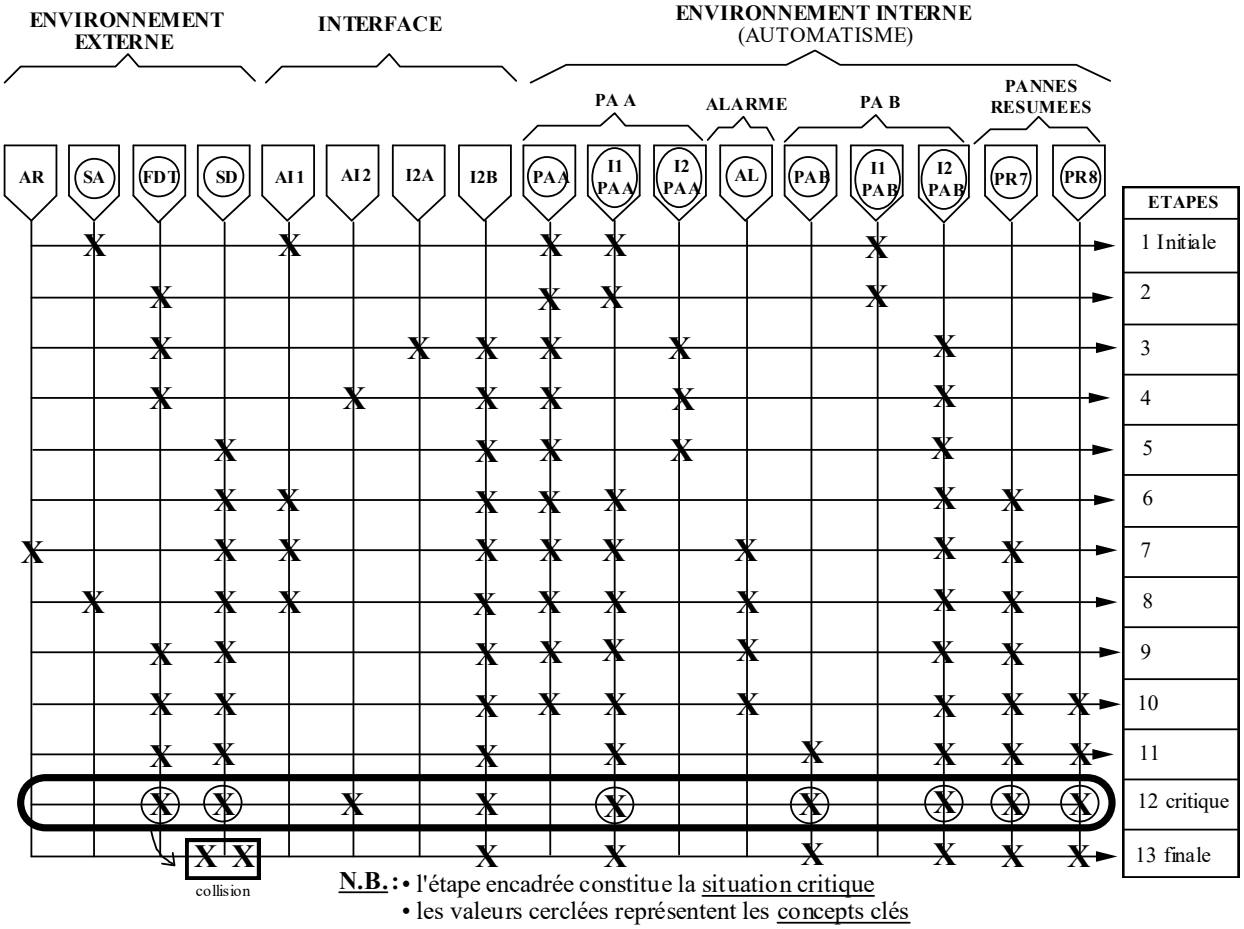


Figure 2.10. : Tableau de Séquencement du marquage

4.3. Limites des moyens d'acquisition de connaissances appliqués au domaine de la certification des STA

L'acquisition des connaissances a permis de caractériser précisément le domaine d'expertise, d'en recenser l'ensemble des concepts fondamentaux et de recadrer les objectifs de l'étude pour tenir compte des réalités. Elle a débouché sur l'élaboration d'un modèle "générique" pour la représentation des scénarios d'accidents. Les scénarios collectés à ce jour dans la base de connaissances historiques (BCHS), concernent le problème de la *collision* /HADJ-MABROUK 91b, 92f/.

Toutes les étapes d'extraction des connaissances proposées dans la méthode de BENKIRANE (identification, macro-extraction, structuration, micro-extraction et formalisation) ont été réalisées pour extraire les connaissances historiques d'analyse de sécurité (scénarios). En ce qui concerne la démarche experte de certification, l'étape de macro-extraction a permis de dégager les trois grandes activités chronologiques impliquées dans l'évaluation des scénarios : classification, évaluation et génération des scénarios. En revanche, l'étape de micro-extraction qui, selon l'auteur, permet d'identifier de manière détaillée les mécanismes de raisonnement de l'expert, ses stratégies et heuristiques de résolution de problème n'a pu être franchie, compte tenu du caractère intuitif, évolutif et créatif du mode de raisonnement des experts certifieurs. De ce fait, la réalisation de la maquette de faisabilité du SBC de certification n'a pas été possible et le recours à d'autres moyens d'acquisition de connaissances automatique (apprentissage automatique) s'est avéré nécessaire.

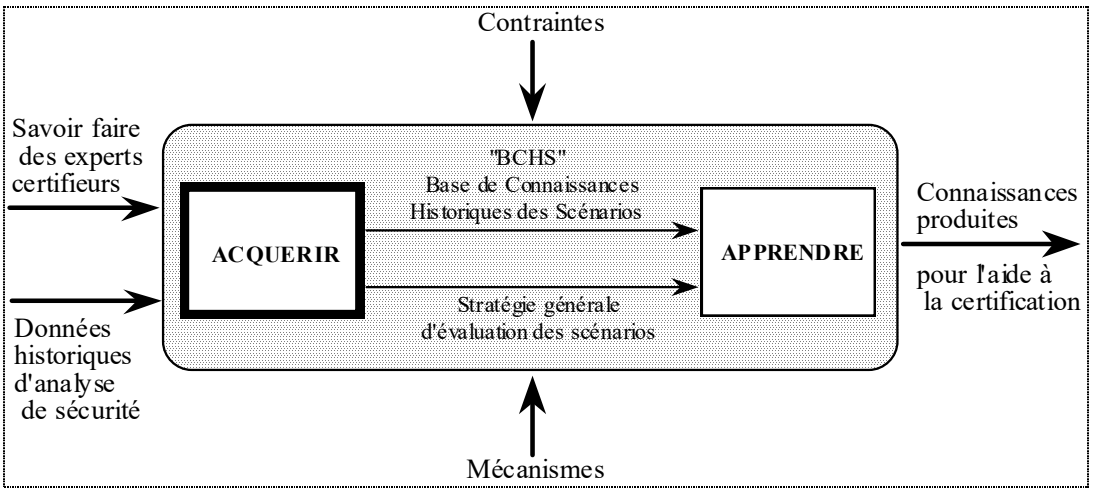


Figure 2.11. : Démarche générale d'acquisition de connaissances de certification

Notre approche (figure 2.11.) consiste à exploiter par apprentissage la Base de Connaissances Historiques des Scénarios, en vue de produire des connaissances susceptibles d'aider les experts de certification dans leur mission d'évaluation du degré de sécurité d'un nouveau STA. Plus précisément, notre objectif est de reproduire en détail le contenu des trois activités impliquées dans la démarche de certification et identifiées au cours de l'étape de macro-extraction, en substituant à l'étape infructueuse de micro-extraction une phase d'apprentissage automatique. L'étude correspondante fait l'objet des chapitres suivants.

CONCLUSION

L'aboutissement d'un projet de développement d'un SBC dépend fortement du problème complexe et crucial de l'acquisition des connaissances manipulées par les experts. L'état de l'art sur les travaux menés dans le domaine de l'acquisition de connaissances a permis de choisir une méthode pour le développement d'un SBC d'aide à la certification des systèmes de transport terrestres automatisés. Cette méthode, appliquée au domaine de la certification, a montré son intérêt pour extraire et formaliser les connaissances historiques d'analyse de sécurité (essentiellement des scénarios d'accident) et ses limites au niveau de l'extraction de la démarche experte de certification, basée notamment sur l'intuition et l'imagination.

Généralement, les méthodes actuelles d'acquisition des connaissances ont été conçues pour des problèmes bien structurés. Elles n'abordent pas les spécificités liées à la multi expertise et à la cohabitation de connaissances diverses et n'autorisent pas l'accès aux connaissances subjectives et intuitives liées à un domaine fortement évolutif et non borné comme l'est celui de la certification. Si la psychologie cognitive et le génie logiciel ont généré des méthodes et outils d'aide à l'acquisition des connaissances, l'exploitation de ces outils demeure encore limitée, dans un contexte industriel complexe.

Nous estimons que, situé en aval, l'apprentissage automatique peut avantageusement contribuer à compléter et renforcer les moyens conventionnels d'acquisition de connaissances. En fait, ces deux approches sont complémentaires et interactives pour optimiser le processus d'acquisition de connaissances : l'apprentissage peut être perçu comme un moyen pour fournir des connaissances au système à partir d'exemples et les résultats de l'acquisition peuvent amorcer le mécanisme d'apprentissage. Le chapitre suivant montre l'intérêt de l'apprentissage pour l'élaboration d'une base de connaissances de certification.

Chapitre 3 :

Apport de l'apprentissage automatique pour le développement d'un SBC d'aide à la certification des systèmes de transport terrestres automatisés.

INTRODUCTION

Le chapitre précédent a présenté l'étude, la mise en œuvre ainsi que les apports et limites des moyens d'acquisition de connaissances pour le développement d'un SBC d'aide à la certification. L'application de ces moyens a débouché essentiellement sur l'élaboration d'un modèle "générique" de représentation des scénarios d'accidents et sur la constitution d'une Base de Connaissances Historiques des scénarios (BCHS) qui regroupe une quarantaine de scénarios relatifs au risque de collision. L'acquisition de connaissances s'est toutefois heurtée à la difficulté d'extraire l'expertise évoquée à chaque étape de la démarche de certification. Cette difficulté émane de la complexité de l'expertise qui incite naturellement les experts à décliner leur savoir faire au travers d'exemples significatifs ou scénarios d'accidents vécus sur des systèmes de transport automatisés déjà certifiés ou homologués.

Dès lors, la mise à jour de l'expertise doit se faire à partir d'exemples. Pour renforcer les moyens usuels d'acquisition de connaissances, nous avons orienté ce travail vers l'utilisation des méthodes d'apprentissage automatique. L'apprentissage automatique (Machine Learning) permet de faciliter le transfert de connaissances, notamment à partir d'exemples expérimentaux. Il contribue à l'élaboration des bases de connaissances des SBC tout en réduisant l'intervention du cogniticien. Dans notre approche, l'apprentissage exploite la Base de Connaissances Historiques des scénarios (BCHS) pour engendrer de nouvelles connaissances susceptibles d'aider les experts de certification à évaluer le degré de sécurité d'un nouveau système de transport. Mener à bien cette étude nécessite dans un premier temps de faire un état de l'art sur les méthodes et systèmes d'apprentissage. Cet état fait l'objet du présent chapitre, composé de trois parties.

La première partie présente une définition générale du domaine de l'apprentissage automatique. L'étude bibliographique approfondie des diverses méthodes, techniques et systèmes d'apprentissage développés met en évidence la difficulté de définir et de choisir un système d'apprentissage adapté aux exigences d'une application industrielle. Nous proposons dans la deuxième partie une caractérisation du processus d'apprentissage par ses données d'entrée, ses contraintes, ses mécanismes et ses données de sortie. Cette caractérisation est exploitée en troisième partie pour effectuer le choix des systèmes d'apprentissage nécessaires à l'élaboration de la base de connaissances de certification.

1. GENERALITES SUR L'APPRENTISSAGE

Dans ce paragraphe, nous abordons l'apprentissage en évoquant dans un premier temps l'analogie fondamentale qui existe entre l'apprentissage humain et l'apprentissage automatique. Dans un deuxième temps, nous présentons les principaux objectifs de l'apprentissage automatique, un bref historique de son évolution ainsi que les problèmes qu'il soulève.

1.1. Apprentissage humain et apprentissage automatique

La faculté d'apprendre apparait comme une composante essentielle de la définition de l'intelligence dont tous les êtres humains ont l'apanage. En effet, selon LAURIERE /87/, l'homme est préprogrammé pour apprendre et il ne sait rien faire sans mémoriser ce qu'il fait et ceci parfois contre sa volonté. Le nouveau né possède déjà des attitudes réflexes et est par nature fondamentalement curieux. L'homme, toute sa vie ne peut s'empêcher d'aller à la quête de l'information, de tourner le regard vers quelque chose qui bouge, de tendre l'oreille vers un son inhabituel. L'ordinateur est a priori dénué d'un tel programme de recherche et d'apprentissage général. Aussi, les méthodes et approches de recherches développées actuellement en apprentissage automatique ne peuvent-elles, en aucun cas, se substituer aux mécanismes d'apprentissage et au comportement intelligent dont les hommes sont dotés. THAYSE /91/ relève que la majeure partie des résultats obtenus à ce jour pour l'apprentissage automatique est basée sur des hypothèses de travail fort simplificatrices par rapport aux mécanismes réels de l'apprentissage humain.

Les concepts appris par un système sont souvent décrits dans un formalisme logique plus ou moins expressif et appartiennent à un même niveau d'abstraction alors que l'être humain a tendance à organiser ses concepts en de multiples niveaux d'abstraction. A l'opposé de l'apprentissage humain, de nombreux systèmes d'apprentissage sont non-incrémentaux au sens où les données nécessaires à l'apprentissage doivent être entièrement disponibles au démarrage du processus. Quelque soit le degré d'intelligence de la machine et sa capacité d'acquisition, l'apprentissage automatique constitue un outil précieux qui nécessite l'intervention de l'homme pour fournir et contrôler les connaissances apprises. La non prise en compte de la coopération entre l'homme et les mécanismes d'apprentissage constitue l'une des principales faiblesses de la majorité des systèmes développés actuellement.

En résumé, l'apprentissage est un terme très général qui décrit le processus selon lequel l'être humain ou la machine peut accroître sa connaissance. Apprendre c'est donc raisonner : découvrir des analogies et des similarités, généraliser ou particulariser une expérience, tirer parti de ses échecs et erreurs passés pour des raisonnements ultérieurs. Les nouveaux acquis sont utilisés pour résoudre de nouveaux problèmes, accomplir une nouvelle tâche ou accroître les performances dans l'accomplissement d'une tâche existante, expliquer une situation ou prédire un comportement. Les domaines d'activité humaine sont de plus en plus complexes et font intervenir des quantités d'informations que l'esprit humain synthétise avec difficulté. Extraire de cette masse de données des connaissances pertinentes et utiles dans un but explicatif ou décisionnel constitue l'objectif principal de l'apprentissage automatique.

1.2. Objectifs de l'apprentissage automatique

L'émergence et le développement industriel des systèmes à base de connaissances exigent la conception d'outils d'aide à l'acquisition des connaissances, incluant des mécanismes d'apprentissage. L'apprentissage suscite en effet un intérêt croissant depuis quelques années, comme le prouve le nombre impressionnant de publications et de congrès dont il fait l'objet. Cette discipline, considérée comme une solution prometteuse pour l'aide à l'acquisition de connaissances, tente notamment de répondre à certaines questions /KODRATOFF 85/ : comment représenter explicitement une masse de connaissances, comment la gérer, l'accroître, la modifier ? Selon GANASCIA /87/, l'apprentissage automatique se définit par un double objectif : un objectif scientifique, comprendre et mécaniser les phénomènes d'évolution dans le temps et d'adaptativité des raisonnements et un objectif pratique, acquérir automatiquement des bases de connaissances à partir d'exemples. En effet, il y a tant de connaissances diverses à posséder que l'idéal est qu'un système informatique les apprenne de lui même à partir des exemples plutôt que de les recevoir de l'homme une à une /LAURIERE 87/. BISSON et LAUBLET /89/ définissent un système d'apprentissage comme étant un système qui, en acquérant les événements significatifs de son environnement, est capable de construire et d'organiser une représentation interne de celui-ci et d'utiliser cette représentation. En ce sens, l'apprentissage peut être défini par l'amélioration des performances avec l'expérience. En effet, d'après GANASCIA /90/, l'apprentissage est intimement lié à la généralisation : apprendre c'est passer d'une succession de situations vécues à un savoir réutilisable dans des situations similaires.

1.3. Historique de l'apprentissage

L'apprentissage automatique est une branche importante de la recherche dans le domaine de l'IA. La naissance de cette discipline remonte aux années 1960 et le résultat le plus spectaculaire a été obtenu à cette époque par l'américain A. SAMUEL. Ce dernier a conçu un programme de jeu de dames "CHECKERS" qui, en mémorisant un grand nombre de coups, améliorait constamment sa stratégie et finissait par atteindre le niveau d'un champion dans cette discipline. Le principe de l'apprentissage était né : apprendre, c'est parfaire ses connaissances et améliorer ses performances en tirant parti des échecs passés. De nos jours, les recherches menées dans ce domaine connaissent un essor considérable issu des progrès mais surtout des besoins des SBC. L'une des principales lacunes des SBC concerne l'acquisition de connaissances. L'apprentissage automatique tente d'apporter un élément de réponse à ce problème. Dans les années 70, une nouvelle approche de l'apprentissage a vu le jour : l'approche IA qui vise l'explicabilité dans la base de connaissances formée. Ainsi, l'apprentissage qui n'était au début qu'une idée intéressante est devenu aujourd'hui une discipline indispensable au progrès des SBC.

1.4. Problématique de l'apprentissage automatique

Pour chacune des principales méthodes d'apprentissage existantes, trois types de problèmes se posent /KODRATOFF 86/. Le premier est celui du regroupement (qu'on appelle "classification en analyse de données") : étant donnée une masse de connaissances, comment découvrir des traits communs entre elles de

sorte qu'on puisse les regrouper en sous-groupes plus simples et ayant une signification? L'immense majorité des procédures de regroupement est de nature numérique. Le second problème (de discrimination) est celui de l'apprentissage de procédures de classification : étant donné un ensemble d'exemples de concepts, comment trouver une méthode qui permette efficacement de reconnaître chaque concept ? Les méthodes existantes reposent en grande majorité sur des évaluations numériques. Le troisième problème est celui de la généralisation : comment, à partir d'exemples concrets d'une situation, trouver une formule assez générale pour décrire cette situation et comment expliquer la capacité de description de cette formule ?

Les efforts menés dans le domaine de l'apprentissage automatique pour appréhender les problèmes de regroupement, de discrimination et de généralisation d'objets ont débouché sur une grande variété de méthodes, techniques, algorithmes et systèmes. Ce foisonnement rend difficile la perception du domaine, compte tenu de l'ambiguïté du vocabulaire qui lui est propre et de l'absence de définitions de référence. Ceci a conduit certains chercheurs à développer des approches de caractérisation et de modélisation du processus d'apprentissage. Parmi ces approches, on peut citer :

- La décomposition des différents niveaux d'apprentissage proposée dans /LAURIERE 87/.
- La classification typologique des systèmes d'apprentissage présentée dans /MICHALSKI 89/.
- Les modèles phénoménologique et comportemental de l'apprentissage décrits dans /GANASCIA 88, 89,90/.
- Le modèle fonctionnel des systèmes d'apprentissage détaillé dans /BISSON 89/.
- L'évaluation expérimentale des systèmes d'apprentissage évoquée dans /LOUNIS 91/.

Si ces approches présentent un intérêt sur le plan scientifique pour dégager des voies de recherche et des besoins futurs en matière d'apprentissage, elles ne permettent pas d'orienter l'utilisateur dans le choix d'un système d'apprentissage adapté à une application industrielle donnée. En effet, aucune de ces approches ne propose de critères pratiques de comparaison des différents systèmes existants tels que les caractéristiques, les principes de fonctionnement, le domaine d'application. Sur la base des travaux existants, notre contribution vise à synthétiser les concepts fondamentaux impliqués dans le processus d'apprentissage pour guider l'utilisateur dans la recherche du système d'apprentissage adapté à son problème. Ce travail est ensuite repris pour sélectionner les systèmes appropriés au développement de l'outil d'aide à la certification.

2. CARACTERISATION D'UN PROCESSUS D'APPRENTISSAGE

Inspirée notamment des travaux de /KODRATOFF 85, 86, 87, 89/, /GANASCIA 85, 87, 88, 89, 90, 91/, /VRAIN 87/, /MANAGO 88, 89/, /MICHALSKI 89/, /BISSON 89, 91/ et /THAYSE 91/, notre caractérisation du processus d'apprentissage est synthétisée en figure 3.1. et détaillée ci-après.

2.1. Présentation générale

Le processus d'apprentissage se caractérise par les données d'entrée et de sortie, les contraintes à respecter ainsi que les mécanismes (ou moyens) à mettre en œuvre pour effectuer cet apprentissage.

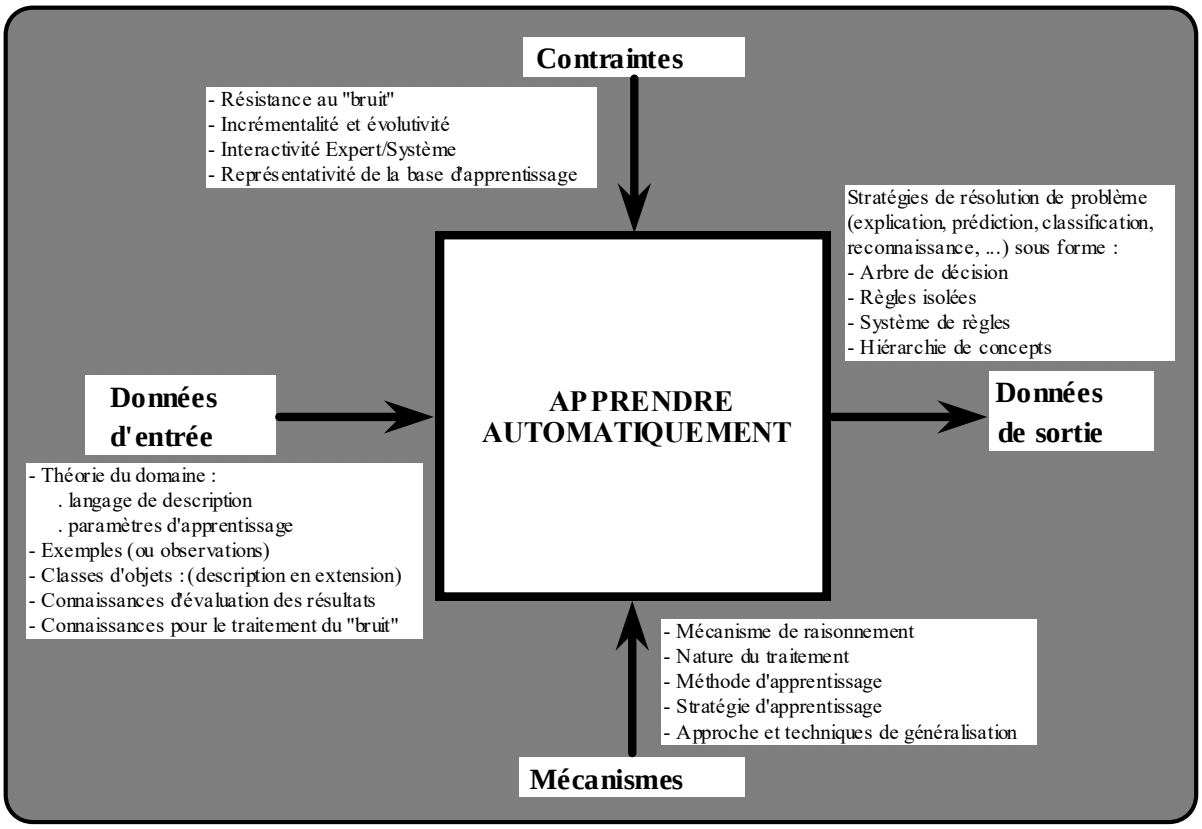


Figure 3.1. : Caractérisation d'un système d'apprentissage

Cette approche dérive des quatre interrogations traditionnelles suivantes :

- à partir de quoi un système d'apprentissage peut il apprendre ? il s'agit des données d'entrée du système.
- quelles sont les contraintes qui devront être prises en compte par le système pour "bien" apprendre ?
- comment apprendre ? Il s'agit des mécanismes nécessaires pour apprendre.
- qu'est-ce que le système a appris et sous quelle forme? il s'agit des données de sortie du système.

2.2. Données d'entrée d'un processus d'apprentissage

Pour qu'un système d'apprentissage puisse produire des résultats, il doit disposer essentiellement d'une théorie du domaine, d'un ensemble d'exemples d'apprentissage, de classes d'objets, de connaissances d'évaluation des résultats et de connaissances destinées au traitement des données bruitées.

2.2.1. Théorie du domaine

La théorie du domaine désigne les connaissances sur le domaine encore appelées connaissances "préalables" ou "a priori" ou connaissances de "fond". Elle exprime toute connaissance explicite propre au domaine d'application telle que le type et le domaine de définition des descripteurs ainsi que leurs relations. Elle englobe généralement le langage de description des exemples et les paramètres d'apprentissage.

Langage de description

L'existence d'un ensemble d'apprentissage présuppose la définition d'un langage "source" pour représenter les exemples et d'un langage "cible" pour formaliser les connaissances apprises. Le choix de ces langages résulte d'une analyse préalable de la nature des données.

- le langage source décrit les propriétés des objets en indiquant le domaine de définition des descripteurs, leur type (énuméré, continu, multivalué ...) ainsi que leurs relations dont le mode de représentation est divers : règles (axiomes ou implications logiques), taxonomies, hiérarchies conceptuelles ou théorèmes. Ces relations entre descripteurs s'attachent principalement à compléter les descriptions des exemples (données bruitées) ou à réaliser une généralisation. Par exemple, si on dispose d'une hiérarchie conceptuelle explicitant des relations de généralité entre figures, au cours du processus de généralisation des exemples, on peut remonter dans cette hiérarchie et généraliser la proposition "forme = carré" en "forme = polygone".

- le langage cible permet de formaliser les connaissances apprises. Les généralisations finales produites peuvent s'exprimer sous forme de règles logiques ou approximatives, isolées à un seul pas d'inférence ou structurées en un système de règles. Les connaissances produites peuvent aussi être décrites sous forme d'une hiérarchie de concepts ou d'un arbre de décision.

Paramètres d'apprentissage

Ces paramètres d'apprentissage sont souvent qualifiés de "prédisposition à apprendre" ou de "préconception". Ils sont constitués d'heuristiques et de stratégies permettant de guider le processus d'apprentissage et agissent sur la qualité du résultat que doit fournir l'apprentissage. Ils expriment des paramètres de réglage, des critères de cohérence ou de préférence entre diverses généralisations ainsi que des contraintes sur la structuration des règles /GANASCIA 88/. Ces paramètres de réglage permettent de restreindre l'exploration de l'espace des généralisations potentielles aux seules caractérisations pertinentes.

2.2.2. Exemples d'apprentissage

Les exemples constituent l'ensemble expérimental sur lequel va opérer le processus pour engendrer de nouvelles connaissances, des concepts, des caractérisations ou des règles. Pour aboutir à des règles consistantes, l'ensemble d'exemples doit être non bruité et suffisamment représentatif du domaine d'application. Un exemple est qualifié de positif s'il appartient à une classe d'objets ou de négatif s'il n'appartient pas à cette classe et qu'il est en fait un contre-exemple. Les exemples peuvent faire l'objet d'un classement préalable de la part de l'expert. A défaut de classement, on les appelle des observations.

2.2.3. Classes d'objets

L'énumération exhaustive des exemples constituant une classe d'objets déterminée par l'expert du domaine est qualifiée en apprentissage automatique de description en "extension" ou "étendue" de la classe.

2.2.4. Connaissances d'évaluation des résultats produits

Ces connaissances sont destinées à l'évaluation des résultats produits par le système d'apprentissage. Cette évaluation peut être réalisée en ayant recours par exemple à un lot d'exemples test proposé par l'expert du domaine. L'expert du domaine peut aussi contribuer à évaluer et valider les connaissances apprises. Cette dernière démarche implique une interactivité et une coopération intenses entre l'expert et le système d'apprentissage qui ne sont pas sans conséquence sur la conception des interfaces.

2.2.5. Connaissances pour le traitement des données bruitées

Si un algorithme d'apprentissage permet d'engendrer des règles ou concepts à partir d'exemples expérimentaux, il n'en demeure pas moins que la qualité des connaissances apprises dépend en grande partie de la qualité de la base d'exemples (information correcte, complète, consistante, riche, nombre suffisant d'exemples et de descripteurs). L'apprentissage automatique est particulièrement sensible à la pertinence des données disponibles. Le contrôle de cette qualité repose notamment sur l'acquisition et la mise à profit de connaissances complémentaires permettant de réduire le bruit diffus dans les exemples.

Après avoir décrit les données d'entrée d'un processus d'apprentissage, nous détaillons ci-après les principales contraintes à respecter pour réaliser un "bon" apprentissage.

2.3. Les contraintes pour un apprentissage efficace

La production de connaissances exploitables conduit à respecter diverses contraintes :

- assurer l'interactivité entre l'expert du domaine et le système d'apprentissage : le système doit expliquer son raisonnement en produisant des connaissances compréhensibles par l'expert dont le rôle est de contrôler, compléter et valider ces connaissances. Cette transparence de la démarche requiert notamment un soin particulier dans la réalisation des interfaces Homme/Machine,

- garantir la représentativité et la complétude de la base d'exemples d'apprentissage,
- assurer la résistance au "bruit" pour s'affranchir des effets perturbants d'exemples mal caractérisés,
- prendre en compte l'incrémentalité et l'évolutivité des connaissances pour faciliter la mise à jour des connaissances.

L'aptitude à acquérir une "bonne" connaissance malgré l'introduction de données erronées (entachées de bruit) et le caractère incrémental de l'enrichissement de la base de connaissances sont deux contraintes importantes pour un système d'apprentissage. Elles sont examinées ci-après.

2.3.1. Bruit et résistance au bruit

L'acquisition automatique des connaissances à partir d'exemples pose le problème crucial du traitement des données bruitées. La qualité et la capacité d'un système d'apprentissage dépendent en grande partie de la cohérence et de la complétude des données à traiter, comme en témoignent les résultats des travaux de KODRATOFF /87/, MANAGO /88, 89/, BEAUNE /90, 92/, SEBAG /90/, STAROSWIECKI /91/ et LOUNIS /91/. Selon MANAGO, il existe plusieurs sources de bruit : le bruit peut résulter de la formalisation de la théorie initiale du domaine par l'expert (absence ou ambiguïté des descripteurs), être intrinsèque au domaine (incertitude d'une mesure, par exemple) ou résulter d'erreurs lors de la saisie de la description d'exemples. Cette dernière source de bruit, due aux erreurs de transmission des données est qualifiée par SEBAG de "bruit blanc".

La détection et le traitement du bruit peuvent intervenir à plusieurs niveaux du processus d'acquisition des connaissances. On peut agir, comme le préconise KODRATOFF, durant la collecte des exemples pour renforcer la théorie du domaine lorsqu'elle s'avère insuffisante. On peut également exploiter, comme le suggère MANAGO, les divergences détectées entre les descriptions des mêmes exemples fournis par des personnes différentes, afin d'obtenir des informations sur la nature du bruit. Enfin, on peut traiter le bruit par le système d'apprentissage. Comme le constate BEAUNE, la plupart des systèmes développés actuellement introduisent des coefficients numériques à un stade ou à un autre du mécanisme d'apprentissage.

Un problème particulier relatif au traitement du bruit concerne les informations manquantes dans la description des exemples d'apprentissage et plus exactement la "valeur manquante d'attribut". Le traitement des valeurs manquantes d'attributs diffère suivant les systèmes :

- pour un attribut non valué, la valeur manquante est considérée comme quelconque ou indifférente et peut prendre toutes les valeurs possibles. L'algorithme PSG de B. ROGER /91/ et les travaux relatés dans /SEBAG 90/ pour le traitement des données bruitées en sont des exemples.
- l'attribut non valué est considéré comme vide de sens ou non significatif. Par exemple, dans le système KATE /MANAGO 89/, la notion de valeur d'attribut spécifique "non-applicable" est introduite.
- les attributs non valués reçoivent la valeur de l'attribut la plus courante dans la classe d'exemples, comme c'est le cas dans l'algorithme ID3 /QUINLAN 86/.
- l'incomplétude des exemples est gérée par la prise en compte d'une connaissance explicite portant sur le domaine et décrite sous forme de règles (ou axiomes). Cette approche est utilisée par le système

CHARADE /GANASCIA 87/.

- enfin dans d'autres approches, l'incomplétude des données est palliée par une évaluation probabiliste des valeurs manquantes.

De façon générale, nous pensons que le mode de traitement qui consiste à attribuer à un exemple incomplet une valeur hypothétique peut altérer la cohérence des exemples d'apprentissage et induire dès le départ des informations non justifiées sémantiquement. Ceci est d'autant plus conséquent s'il s'agit d'un attribut pertinent (ou clé) pour la description de l'exemple considéré.

En effet, dans certains cas pratiques, l'expert du domaine peut volontairement ne pas assigner de valeur à un attribut s'il le considère non pertinent pour discriminer un exemple de ses contre-exemples ou s'il estime que son jugement est trop imprécis pour être émis. Enfin, un oubli de sa part lors de la description d'un exemple n'est pas à exclure.

2.3.2. Incrémentalité, circularité ou itérativité de l'apprentissage

Selon GANASCIA /90, 91/, le comportement idéal d'un système d'apprentissage serait celui qui, bouclant sur lui-même, améliorerait continuellement ses connaissances au contact des expériences auxquelles il est confronté. Cette caractéristique des systèmes qui apprennent, s'adaptent ou évoluent est qualifiée de "dimension itérative" ou encore de "circularité" du savoir : une connaissance induite peut être utilisée en vue d'apprendre de nouvelles connaissances qui serviront elles-mêmes à en construire d'autres. L'auteur considère que les processus d'acquisition et d'évolution ont quelque chose à voir avec la "spirale" : un perpétuel retour sur soi qui a comme effet un accroissement général de la connaissance.

Pour qu'un processus d'apprentissage soit continu, il faut que l'essentiel de l'information contenue dans les exemples soit retenu afin que l'apprentissage se poursuive avec les nouveaux exemples, sans qu'il soit nécessaire de revenir sur les exemples initiaux. En d'autres termes, lorsque A est le résultat d'un apprentissage à partir d'un ensemble d'exemples E1, E2, ... En, on doit pouvoir oublier tous ces exemples de sorte que, si un nouvel exemple E_{n+1} se présente, le résultat d'un apprentissage à partir de A et E_{n+1} soit cohérent avec les exemples initiaux. Lorsque les systèmes d'apprentissage possèdent cette propriété, on dit qu'ils sont incrémentaux.

Ainsi la notion d'incrémentalité ne signifie pas uniquement que le système accepte les exemples les uns après les autres. Il faut, de plus, que lors de l'ajout d'un nouvel exemple, les modifications à apporter n'entraînent pas la reconstruction complète des connaissances obtenues à partir des exemples déjà utilisés /GENNARI 89/. Cette caractéristique de l'apprentissage est souvent nécessaire pour permettre l'utilisation des informations apprises lorsque l'ensemble d'apprentissage n'est pas encore suffisamment représentatif du problème considéré /ROYER 90/.

En résumé, un système d'apprentissage est dit incrémental s'il est doté de la capacité de faire évoluer les connaissances apprises au cours d'un cycle antérieur, sans devoir à chaque fois retraiter la totalité des exemples collectés. Ce type d'apprentissage est fort appréciable lorsque les exemples couvrant le domaine

sont en nombre considérable et qu'ils ne peuvent être recensés de façon exhaustive auprès des experts ou lorsqu'ils sont longs à acquérir. L'évolution de la connaissance au cours d'un processus d'apprentissage engendre deux types d'incrémentalité /SEBAG 90/ : monotone et non monotone. Les paragraphes suivants présentent ces deux approches d'apprentissage.

2.3.2.1. Apprentissage monotone

Dans l'incrémentalité monotone (croissance ou décroissance continue des connaissances) l'apprentissage ne fait que produire de nouvelles connaissances qui complètent les connaissances initiales sans remettre en cause les connaissances déjà apprises. Il s'en suit, comme le souligne GANASCIA /90/, que la capacité à reconnaître ses propres erreurs semble absente. L'incrémentalité monotone n'est pas adaptée au traitement des données bruitées ou évolutives. La méthode de l'espace des versions /MITCHELL 82/ est un exemple type de processus monotone. L'espace des versions considère deux espaces S et G où S désigne l'ensemble des versions les plus spécifiques du concept à apprendre et G désigne l'ensemble des versions les plus générales du concept à apprendre. S est initialisé à "rien" et G est initialisé à "tout". Un nouvel exemple E du concept à apprendre conduit à généraliser S afin de couvrir E. Un nouveau contre-exemple CE du concept conduit à spécialiser G pour rejeter CE. Ainsi la généralité de S croît pendant que celle de G décroît. Dès qu'un nombre suffisant d'E et de CE a été considéré, S et G sont égaux et le système convergence vers la solution finale. Cet exemple permet d'introduire la notion de "convergence du système" qui est l'une des propriétés intéressantes garantie par l'incrémentalité monotone, lorsque la base d'exemples n'est pas bruitée.

2.3.2.2. Apprentissage non monotone

L'algorithme ADECLU /DECAESTECKER 89/ qui construit incrémentalement des hiérarchies de concepts permet d'illustrer ce type d'apprentissage. Il s'attache à intégrer un nouvel objet (exemple) dans une hiérarchie existante, tout en restructurant cette dernière. L'arbre est construit en utilisant quatre opérateurs : création d'un nœud, suppression d'un nœud, fusionnement de deux nœuds et éclatement en deux d'un nœud. En présence d'un nouvel exemple à classer, le processus d'apprentissage peut défaire ce qu'il a appris dans l'étape antérieure : détruire un nœud déjà créé, éclater un nœud précédemment fusionné. Cette réversibilité du processus permet une évolution non monotone des connaissances.

Contrairement à l'approche monotone, l'apprentissage non monotone est mieux adapté aux données bruitées, en revanche le processus n'offre plus de garanties de convergence et devient théoriquement capable d'osciller ou de boucler /SEBAG 90/. Dans l'algorithme ADECLU, la convergence est assurée en diminuant progressivement l'influence des exemples considérés. Ainsi, au fur et à mesure que le processus avance et que les seuils numériques évoluent, les exemples sont de moins en moins déterminants ; la connaissance acquise prend progressivement le pas sur les nouveaux exemples. Ceci induit inévitablement une sensibilité à l'ordre de prise en compte des exemples.

Un processus d'apprentissage doit non seulement respecter les deux contraintes fondamentales d'incrémentalité et de résistance au "bruit" mais aussi mettre en oeuvre des mécanismes décrits ci-après.

2.4. Les mécanismes nécessaires pour apprendre

Pour qu'un processus d'apprentissage puisse produire des connaissances, il doit faire appel aux mécanismes suivants :

- . Le raisonnement ou mode d'inférence,
- . La nature du traitement,
- . La méthode d'apprentissage,
- . La stratégie d'apprentissage,
- . Le principe de généralisation.

Les paragraphes suivants détaillent successivement ces cinq mécanismes d'apprentissage.

2.4.1. Raisonnement ou mode d'inférence

A la différence de la classification typologique proposée dans /MICHALSKI 89/ qui distingue d'un point de vue théorique les systèmes déductif, inductif et analogique, un système d'apprentissage peut faire appel par exemple à la fois à l'induction et à la déduction pour engendrer des connaissances. C'est le cas, par exemple, des systèmes AGAPE /GANASCIA 87/ et OGUST /VRAIN 87/. En fait, l'apprentissage automatique repose sur quatre mécanismes de raisonnement ou modes d'inférence : induction, déduction, abduction et analogie. Le but du présent paragraphe est de définir ces quatre termes en référence à des définitions issues des travaux en apprentissage et en intelligence artificielle.

GRUNDSTEIN /88/ définit ces termes à travers un exemple :

La **déduction** procède d'une règle et d'un fait pour obtenir un résultat :

Règle : Tous les haricots de ce sac sont blancs
Fait : Ces haricots sont tirés de ce sac
Résultat : Ces haricots sont blancs.

L'**induction** aboutit à la règle en partant d'un fait et d'un résultat :

Fait : Ces haricots sont tirés de ce sac
Résultat : Ces haricots sont blancs
Règles : Tous les haricots de ce sac sont blancs

L'**abduction** aboutit au fait en partant de la règle et d'un résultat :

Règle : Tous les haricots de ce sac sont blancs
Résultat : Ces haricots sont blancs
Fait : Ces haricots sont tirés de ce sac.

L'**analogie** procède, quant à elle, du fait au fait : elle consiste à inférer qu'une propriété vraie d'un objet peut également l'être pour un autre présentant des similitudes avec le premier. La plausibilité de la conclusion dépend de la ressemblance entre ces deux objets. L'induction, l'abduction, et l'analogie sont qualifiées par GRUNDSTEIN de raisonnements non rigoureux. En effet, ces mécanismes s'appliquent à des connaissances certaines mais aboutissent à des conclusions qui ne sont que plausibles.

Dans le cadre des Journées Françaises de l'Apprentissage de Sète (JAF 91), J. QUINQUETON (INRIA - LIRMM) à défini la déduction, l'abduction et l'induction comme suit :

La déduction : "Socrate est un homme, les hommes sont mortels, donc Socrate est mortel".
L'abduction : "Socrate est mortel, les chats sont mortels, donc Socrate est un chat".
L'induction : "Je vois une Française, elle est rousse, dont toutes les Françaises sont rousses".

Très schématiquement, on peut dire d'après DUVAL /91/ que l'**abduction** raisonne des effets vers les causes probables, alors que la **déduction** raisonne des causes vers les effets. Voici un exemple d'illustration des ces deux notions : sachant que ∀ x [Nuage (x) => pluie (x)]. C'est une abduction de prédire des nuages quand on constate la présence de la pluie. Au contraire, c'est une déduction de prédire qu'il y aura de la pluie quand on constate la présence des nuages. Si l'on considère le fameux Modus Ponens, la **déduction** est un mode d'inférence certain : "de A et de A=>B, déduire (ou inférer) B". En revanche, l'**abduction** se caractérise par une inversion du Modus Ponens. L'**induction** cherche à déterminer des relations générales de cause à effet à partir des faits spécifiques. On peut représenter de façon plus formelle ces trois mécanismes de raisonnement:

• déduction :	à partir de	A	et de	A=>B,	on "déduit"	B
• abduction :	à partir de	B	et de	A=>B,	on "abduit"	A
• induction :	à partir de	A (z) =>B	et de	A (t) =>B,	on "induit"	A (x) =>B

L'**analogie** est utilisée en pratique pour comprendre ou interpréter de nouvelles situations à partir de situations antérieures déjà mémorisées. L'analogie combine à la fois la notion de similarité (ou ressemblance) et la notion de causalité (figure 3.2). Plus formellement, une analogie comporte une situation source de la forme (A, B) et une situation cible de la forme (A', B'). Il existe des relations de similarité (et de dissimilarité) entre A et A', respectivement B et B', ainsi que des relations de dépendance, généralement de nature causale, entre A et B, respectivement A' et B'.

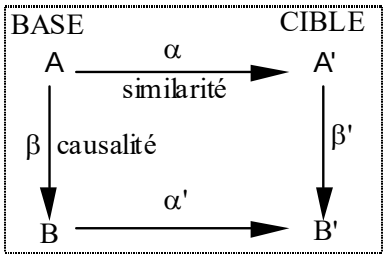


Figure 3.2.

Faire une analogie, c'est partir d'un schéma incomplet semblable à celui-ci de la figure 3.2. et le compléter en combinant similarité et causalité. En résumé, l'apprentissage par analogie consiste d'abord à reconnaître des similarités entre le concept cible à apprendre et un concept source connu et ensuite à déterminer qu'elles caractéristiques pertinentes peuvent être transférées de la "source" vers la "cible". De nombreux travaux ont traité de l'apprentissage par analogie /KODRATOFF 87/, /SEROUSSI 89/ et /BICHINDARITZ 91/.

Nous venons de présenter les quatre modes d'inférence utilisés dans le domaine de l'apprentissage automatique. Il convient à présent d'aborder la nature du traitement des données.

2.4.2. Nature du traitement

Dans un système d'apprentissage le traitement des données peut être de nature numérique, purement symbolique ou symbolique-numérique. Ces trois approches sont définies ci-après.

2.4.2.1. Traitement numérique

Le traitement de l'information de type numérique (statistiques et analyse de données) concerne généralement de gros volumes de données, statistiquement représentatives du problème étudié. Les techniques classiques d'analyse de données (AD) reposent, la plupart du temps sur une numérisation des données. En analyse de données, la connaissance du domaine (connaissance supplémentaire ou sémantique du domaine) est définie par le choix d'une mesure de dissimilarité entre les objets, généralement utilisée pour les classer. Si la représentation des connaissances, sous forme de tableaux de données, couramment utilisée par les techniques classiques de l'analyse de données s'est avérée très utile pour le traitement de simples observations, elle n'est pas adaptée au traitement de connaissances plus complexes /DE CARVALHO 90/.
L'approche numérique s'attache à optimiser un critère global comme l'entropie dans le cas de l'algorithme ID3 /QUINLAN 83/ ou bien une distance entre exemples en analyse de données. L'inconvénient majeur des méthodes numériques est non seulement l'appauvrissement des données initiales lors de leur traduction en nombre mais réside aussi dans le fait que la sémantique des opérations numériques diffère parfois de celle des données symboliques initiales. De plus, les connaissances engendrées sont souvent incompréhensibles par les humains. En revanche, l'approche numérique est efficace et dispose de méthodes pour résister au bruit induit par les données erronées.

2.4.2.2. Traitement symbolique

Le traitement symbolique de l'apprentissage vise à mimer le raisonnement humain en utilisant des descripteurs directement compréhensibles par des humains. La connaissance du domaine se présente sous des formes variées telles que des taxonomies, hiérarchies, des régles exprimant des contraintes entre propriétés, des théorèmes ou des axiomes. Un objet symbolique est une description en intention (ou compréhension) d'une classe d'objets élémentaires qui en constituent l'extension. Une description en intention est obtenue en sélectionnant les propriétés caractéristiques et communes des exemples. La dualité intention/extension et les techniques de généralisation/spécialisation sont amplement utilisées dans l'approche symbolique. Les objets symboliques s'expriment par des conjonctions logiques de descripteurs.
L'approche symbolique favorise l'aspect explicatif par le fait même que les connaissance produites par apprentissage sont essentiellement symboliques. La finalité des méthodes symboliques est d'utiliser les connaissances afin d'en produire de nouvelles qui ne sont pas trivialement présentées dans la description initiale du problème. Ces nouvelles connaissances constituent une explication d'un niveau plus haut que celui de l'observation en analyse de données classique /BOUCHON 90/. Dans l'approche symbolique on ne se demande plus ce qui est le plus efficace mais ce qui est le plus signifiant /KODRATOFF 85/. Toutefois, l'inconvénient majeur de l'approche symbolique est sa sensibilité au bruit qui rend son application difficile lorsque la population d'apprentissage est fortement incohérente /STAROSWIECKI 91/, /POMORSKI 91/.

2.4.2.3. Traitement symbolique-numérique

La distinction entre le traitement numérique et le traitement symbolique se situe au niveau des types d'opérations mises en œuvre dans l'algorithme d'apprentissage. L'apprentissage numérique repose sur des opérations de type numérique (l'analyse de données par exemple), alors que l'apprentissage symbolique repose essentiellement sur des opérations de type logique. L'intelligence artificielle s'est en premier lieu préoccupée du traitement symbolique de l'information alors que dans d'autres domaines est apparue conjointement la nécessité d'utiliser des méthodes à composantes numériques. L'approche symbolique est précieuse par sa capacité à fournir a priori des explications. Elle est néanmoins relativement inefficace pour traiter les données bruitées contrairement à l'approche numérique qui se caractérise par son efficacité et son manque d'explicabilité. Combiner ces deux approches qui poursuivent l'objectif commun de découvrir des connaissances utiles à partir des faits, pour inhiber leurs points faibles respectifs, constitue selon KODRATOFF /87, 89, 91/ et /DIDAY 89/ une avance technologique importante.

En matière de raisonnement, la présence simultanée des données quantitatives et symboliques et le caractère imparfait de certaines informations rendent insuffisantes les approches purement symboliques. L'opération symbolique de généralisation d'observations est un processus très complexe et hautement combinatoire (plusieurs généralisations possibles) qui exige un temps de calcul important. Ceci montre l'intérêt d'utiliser des données numériques relatives aux exemples à généraliser pour hâter la convergence du processus de généralisation /KODRATOFF 89/. L'approche symbolique est capable d'explications car elle opère sur des données sous forme de graphes conceptuels, réseaux sémantiques etc. Ceci induit selon PARODI /90/ une certaine "friabilité" du raisonnement et suppose la connaissance implicite d'un contexte. Ces constatations révèlent que l'usage d'une composante numérique est fondamentale voire indispensable pour optimiser le processus d'apprentissage et traiter des problèmes réels complexes où les connaissances du domaine sont souvent incomplètes, non exhaustives ou bruitées.

Le couplage "symbolique-numérique" est apparu à l'initiative de E. DIDAY et Y. KODRATOFF et a été consacré en 1987, par l'organisation en France des "journées symboliques-numériques pour l'apprentissage à partir de données". Les méthodes symboliques sont devenues symboliques-numériques par la force des choses, pour traiter des problèmes dans lesquels les données sont imparfaites. Ceci est confirmé par le développement croissant des systèmes d'apprentissage symboliques tels que INSTIL /KODRATOFF 86/ qui intègre des techniques numériques pour gérer le bruit ou le système ADECLU /DECAESTECKER 89/ qui utilise pour la formation incrémentale de concept, un critère d'adéquation numérique. Le système CHARADE /GANASCIA 87/ quant à lui effectue un traitement statistique des données pour sélectionner des règles incertaines. Dans le cadre de l'apprentissage par la découverte, FISHER /87/ effectue le regroupement des objets en classes par minimisation d'un critère numérique local, défini par le compromis entre la similarité intra-classe et la similarité extra-classe des objets.

L'utilisation conjointe des techniques numériques et symboliques est aussi employée dans le programme MAGGY /KODRATOFF 89/ qui combine l'algorithme numérique ID3 de QUINLAN et d'autres méthodes de généralisation symbolique. Le système de généralisation ALLY /LERMAN 90/ pose le problème de la classification conceptuelle comme étant celui de la détermination d'un ensemble de phases purement numériques et purement symboliques, interagissant dans la résolution du problème global.

L'ensemble de ces travaux atteste de l'intérêt de l'approche symbolique-numérique pour la réalisation de systèmes d'apprentissage efficaces intégrant la composante explication. L'utilisation de raisonnements symboliques permettant de favoriser l'explicabilité et l'interactivité avec les experts humains est un apport intéressant au domaine de l'analyse de données. L'approche numérique, quant à elle, est le complément indispensable des approches symboliques pour leur permettre de traiter des problèmes réels dont les données sont souvent bruitées.

Après avoir détaillé les trois natures de traitement employées en apprentissage, le paragraphe suivant présente les deux principales méthodes d'apprentissage.

2.4.3. Méthodes d'apprentissage

Au cours de l'International "Workshop on Machine Learning" de 1985, deux méthodes ou familles d'apprentissage ont été définies : l'apprentissage par recherche d'explications (ou EBL : Explanation Based Learning) et l'apprentissage par détection de similarités (ou SBL : Similarity Based Learning).

2.4.3.1. Apprentissage par recherche d'explications : EBL

En EBL (Explanation Based Learning), l'apprentissage se fait à partir des explications dérivées de l'analyse d'un exemple ou d'un contre-exemple du concept et grâce à une théorie du domaine. Souvent, ce type d'apprentissage résulte de la recherche d'une meilleure stratégie de résolution de problèmes ou de la résolution de problèmes similaires. Il s'applique particulièrement bien aux domaines où la connaissance experte repose, au moins en partie, sur une théorie bien formalisée. C'est le cas, par exemple, des mathématiques, de la physique ou de la médecine. A chaque fois que le système débouche sur une solution, un module évalue le résultat et analyse les raisons du succès ou de l'échec. Ces raisons, encore appelées "explications", servent à améliorer la capacité d'apprentissage du système. Ce type d'apprentissage est également nommé "apprentissage analytique" dans la mesure ou il améliore la formulation d'une connaissance existante, contrairement à l'apprentissage qualifié de "constructif" qui crée de nouvelles connaissances. Le mécanisme d'inférence utilisé dans l'EBL est en général purement déductif et nécessite une théorie simple, complète et correcte /DUVAL 91/. Certains systèmes mettent également en jeu des inférences abductives et analogiques.

2.4.3.2. Apprentissage par détection de similarités : SBL

L'objectif de ce type d'apprentissage, connu sous le sigle SBL (Similarity Based Learning), est d'apprendre en détectant d'une part des similarités dans un ensemble d'exemples et d'autre part des dissimilarités entre les exemples et des contre-exemples. L'apprentissage par détection des similarités peut être scindé en deux approches : "empirique" et "rationnelle" /KODRATOFF 86/. L'approche empirique permet d'assurer que les concepts appris sont bien discriminants alors que l'approche rationnelle permet d'obtenir une description plus riche, offrant de meilleures explications sur les raisons pour lesquelles les exemples appartiennent au même concept. Dans le cas empirique, le nombre total de généralisations possibles sera considérable et dans le cas rationnel, c'est la longueur de la généralisation qui sera importante. Dans les deux cas, on utilise généralement les contre-exemples pour restreindre l'explosion combinatoire.

L'apprentissage de concepts qui est la finalité des deux approches empirique et rationnelle s'apparente à un problème d'acquisition de similarités. Partant d'un ensemble d'exemples et éventuellement de contre-exemples, détecter ce qui est commun aux exemples et absent dans les contre-exemples constitue un travail de recherche des similarités signifiantes entre exemples /KODRATOFF 87/. L'apprentissage de concept consiste à élaborer une caractérisation générale d'un concept à partir de multiples descriptions d'exemples et de contre-exemples, en appliquant généralement des règles d'inférence inductive et/ou déductive. On parle d'instances positives pour désigner les exemples et d'instances négatives pour désigner les contre-exemples. A partir d'une partition de l'ensemble des exemples en classes, on en déduit que les objets d'une classe donnée forment les exemples du concept et que les objets des autres classes forment les contre-exemples.

L'apprentissage d'une description (ou caractérisation) générale de concept déroule les trois étapes principales suivantes :

- l'observation de l'ensemble d'instances positives et négatives (définition en extension) présentant certains descripteurs communs,
- l'extraction des caractéristiques essentielles communes de ces instances,
- la formalisation d'une définition en intention (ou en compréhension) du concept basée sur ces caractéristiques.

La caractérisation inférée à partir d'un ensemble unique d'exemples positifs, relatifs à un concept est souvent appelée description **caractéristique** du concept car elle identifie les propriétés communes des exemples. On exige de cette description qu'elle soit **complète** pour la classe d'exemples. Une description est complète pour une classe quand elle couvre ou reconnait tous les exemples de cette classe. Une description reconnait un exemple quand ce dernier est une instance de la description.

La caractérisation induite est qualifiée de description **discriminante** lorsqu'elle permet de distinguer plusieurs classes d'objets. Cette description doit être non seulement **complète** mais aussi **cohérente.** Une description est cohérente quand elle ne couvre ou ne reconnait aucun contre exemples.

A chacune des méthodes d'apprentissage évoquées précédemment (EBL et SBL) correspondent plusieurs stratégies d'apprentissage décrites ci-après.

2.4.4. Stratégies d'apprentissage

Nous proposons ici de distinguer quatre stratégies d'apprentissage : l'apprentissage par l'action et l'apprentissage par l'espace des versions qui sont du ressort de l'apprentissage par recherche d'explications puis l'apprentissage par détection de régularité empirique et l'apprentissage par classification conceptuelle qui s'inscrivent dans le cadre de l'apprentissage par détection de similarités. Ces stratégies sont détaillées dans les paragraphes suivants.

2.4.4.1. Apprentissage par l'action ou apprentissage par essais et erreurs

Un système qui apprend par l'action contient deux types d'informations :

- le premier type est un ensemble d'opérateurs initiaux donnés au départ sous forme de règles à partir desquelles le système devra engendrer de nouveaux opérateurs, plus efficaces,
- le second type est un ensemble d'heuristiques initiales qui dirigent, parfois de façon maladroite au départ, l'utilisation des opérateurs initiaux. L'apprentissage doit aussi être capable d'améliorer ces heuristiques.

A partir d'un système déjà capable de résoudre certains problèmes et d'heuristiques de contrôle, l'apprentissage opère en faisant exécuter au système la tâche qu'il vient de réussir. Deux cas peuvent se présenter :

- le système repasse par la même succession d'états et il n'apprend rien,
- le système passe par des états différents et on compare chacun de ses mouvements à ceux du chemin solution. Les règles sont modifiées pour obliger le système à passer par le chemin qui a conduit à la solution précédemment trouvée.

On fournit enfin au système des méthodes pour modifier les heuristiques déjà connues. Ces méthodes constituent le cœur de la méthode d'apprentissage par l'action. Un exemple détaillé au jeu du "saute ou glisse" de ce type d'apprentissage, relatif au système SAGE /LANGLEY 83/, est décrit dans /KODRATOFF 87, 89/.

2.4.4.2. Méthode de l'espace des versions

Pour un concept à apprendre, le terme "espace des versions" /MITCHELL 83/, désigne toutes les versions possibles de ce concept. La notion d'espace des versions permet de contrôler la généralisation pour exploiter les succès et les échecs d'un système. Les succès sont utilisés pour généraliser les heuristiques d'utilisation des opérateurs et les échecs sont utilisés pour les particulariser. L'espace des versions se construit à partir de

taxonomies (arbres de relations de généralité), de descripteurs utilisés pour décrire les opérateurs initiaux et d'heuristiques initiales. Dans le cadre de l'apprentissage utilisant l'espace des versions, les règles à acquérir se présentent sous la forme : **SI** condition **ALORS** action A1. On généralise les conditions de plusieurs règles déclenchant la même action A1 et on particularise les conditions des règles qui ne doivent pas déclencher l'action A1. Le système LEX /MITCHELL 83/ qui s'inscrit dans le cadre de l'apprentissage par recherche d'explications et utilise le concept d'espace des versions permet d'apprendre des heuristiques pour le calcul intégral.

2.4.4.3. Apprentissage par classification conceptuelle

La stratégie d'apprentissage par classification ou regroupement conceptuel fait généralement appel à des mécanismes inductifs. Elle est empirique, car elle utilise un grand nombre d'exemples et peu ou pas de connaissance spécifique au domaine. Son but est de construire à partir d'une suite de descriptions individuelles d'objets, des regroupements en classes conceptuelles organisés en une structure hiérarchisée et caractérisés de façon à satisfaire un critère de qualité fixé. Le processus de regroupement d'objets en catégories conceptuelles peut poursuivre plusieurs objectifs /THAYSE 91/ dont les deux principaux sont :

- le premier consiste à augmenter la précision de la classification : le regroupement doit viser à maximiser la similarité intra-classe et à minimiser la similarité interclasse. Autrement dit, les instances d'une même classe doivent être aussi similaires que possible tandis que des instances de classes distinctes doivent être aussi différentes que possible.
- le second objectif vise à assurer le caractère évolutif de la classification : le regroupement par classes doit pouvoir être reconfiguré dynamiquement lorsqu'une nouvelle instance ne s'intègre pas suffisamment dans la classification existante.

Ces objectifs conduisent à définir des critères de qualité ou "fonctions d'évaluation" pour la classification recherchée. L'élaboration de classes répondant à ces objectifs suppose le déroulement de trois tâches :

- une tâche d'agrégation pour identifier un ensemble de classes à partir d'une suite d'objets donnée; ces classes peuvent être disjointes ou se recouvrir partiellement.
- une tâche de caractérisation pour construire une description conceptuelle de chacune des classes.
- une tâche de structuration pour organiser les classes précédemment identifiées et caractérisées en une structure faisant apparaître les liens habituellement hiérarchiques entre concepts plus généraux et concepts plus spécialisés.

2.4.4.4. Apprentissage par détection de régularités empiriques pour la construction des bases de connaissances

A partir d'un ensemble d'exemples, d'un langage de description de ces exemples et de contraintes d'apprentissage, l'objectif de l'apprentissage par détection de régularités empiriques est de construire une base de connaissances exprimée sous forme de règles utilisables par un système expert. A partir d'une collection d'exemples et éventuellement de contre-exemples décrivant une classe de problèmes C, le mécanisme d'apprentissage induit une caractérisation commune aux exemples qui exclut chacun des contre-exemples.

Une fois cette caractérisation G trouvée, on peut engendrer la règle G => C.

Toutes les stratégies d'apprentissage évoquées ci-dessus font appel à des approches et techniques de généralisation détaillées ci-après.

2.4.5. Principe de généralisation

L'opération de généralisation à partir d'exemples nécessite le choix d'une "approche de généralisation" et la mise en œuvre d'une ou plusieurs "techniques de généralisation", présentées dans les paragraphes suivants.

2.4.5.1. Approche ascendante et approche descendante

On distingue deux approches de généralisation selon que l'exploration de l'espace de généralisation s'effectue du spécifique au général pour l'approche ascendante ou inversement du général au spécifique pour l'approche descendante. Une généralisation est ascendante quand elle part des exemples pour les généraliser progressivement jusqu'à obtenir une caractéristique qui couvre tous les exemples. Une généralisation est descendante lorsqu'elle spécialise une description générale jusqu'à ce que cette dernière ne couvre plus de contre-exemples. En matière d'apprentissage automatique, l'approche descendante est souvent associée à l'apprentissage d'arbres de décision (induction descendante) alors que l'approche ascendante s'apparente à l'apprentissage de règles. La méthode de l'espace des versions fait appel, quant à elle, conjointement aux deux approches.

Après avoir évoqué les deux approches de généralisation employées au cours du processus d'apprentissage, nous présentons ci-après quelques techniques de généralisation.

2.4.5.2. Techniques de généralisation

Tous les systèmes d'apprentissage réalisés utilisent une ou plusieurs techniques de généralisation dont l'objectif est de dégager d'une masse de connaissances une caractéristique générale. En s'inspirant notamment des travaux de J.G. GANASCIA, Y. KODRATOFF, C. VRAIN et G. BISSON nous présentons dans ce paragraphe une vue très "générale" de la généralisation en appliquant nous même cette fameuse "règle de l'oubli" sur laquelle repose en partie la généralisation. A partir d'un ensemble d'exemples E1, E2,… En, décrivant une classe d'objet C, la généralisation s'attache à identifier les descripteurs essentiels communs à ces exemples et à élaborer une caractérisation ou généralisation G. Autrement dit, à partir d'une définition en extension d'un concept C, le mécanisme d'apprentissage doit fournir une description en intention.

En logique des propositions, la généralisation consiste à ne retenir que les propriétés communes des objets. Si les exemples sont décrits sous la forme de conjonctions de propositions, il suffit de prendre l'intersection des différentes conjonctions. Par exemple les conjonctions "Rouge & Carré" et "Rouge & Rectangle" se généralisent en "ROUGE". Cette technique est la plus connue des règles de généralisation et on la qualifie par la règle "d'oubli de conjonctions" ou "d'élimination d'un facteur conjonctif". Plus formellement : si P1 et P2 sont deux propositions alors P1 est plus général que P1 & P2 et on le note P1 ≤ P1 & P2.

En logique des prédicats du premier ordre, on adjoint à cette technique la "variabilisation des constantes". Pour explication, considérons les deux exemples suivants :

E1 : ville (Carthage, tunisienne) & près-de (Carthage, plage)
E2 : ville (Sousse, tunisienne) & près-de (Sousse, plage)

La généralisation de E1 et de E2 est : G = ville (x, tunisienne) & près-de (x, plage). Autrement dit P(a, b) et P(a, c) se généralisent en P(a, x). De façon plus formelle : si P est un prédicat, A une constante et X une variable, alors P(X) est plus général que P(A).

Toute la difficulté du processus d'apprentissage consiste non seulement à découvrir le "bon" généralisé **G** mais aussi à restreindre la production des généralisés potentiels. Comme le montre l'exemple de la figure 3.3, il existe de multiples généralisations possibles telles que G1: deux objets dont un triangle, G2 : un objet coloré placé à droite d'un autre objet, ...

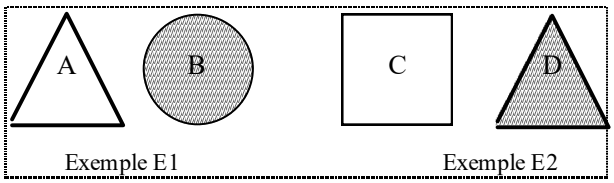


Figure 3.3. : Deux exemples d'apprentissage

Pour limiter l'exploration de l'espace de généralisation aux seules caractérisations pertinentes, on fournit au processus d'apprentissage un ensemble d'informations sémantiques, portant sur les propriétés des objets manipulés et représentant la théorie du domaine généralement acquise auprès de l'expert. Les connaissances utilisées à ce niveau concernent essentiellement des propriétés du langage de description et des contraintes décrites par des heuristiques.

- Les propriétés du langage de description des exemples sont formées des relations entre descripteurs décrivant les objets et explicitées sous forme de règles, théorèmes ou axiomes, d'implications logiques, de taxonomies ou hiérarchies conceptuelles. Ces propriétés sont exploitées au cours de la phase de généralisation. Par exemple, la "remontée dans une hiérarchie conceptuelle" permet de dégager que POLYGONE est plus général que CARRE. De même, la connaissance du théorème utilisé par le système OGUST /VRAIN 87/ : A =>B (où les descriptions A et B sont exprimées en logique du premier ordre) permet de conclure que B est plus général que A. On remarque que ce théorème produit des résultats analogues à ceux obtenus par la règle de "remontée dans une hiérarchie conceptuelle".

- Les contraintes décrites par des heuristiques ou paramètres de généralisation permettent de limiter la combinatoire de la recherche en se propageant dans l'espace de généralisation. Un exemple de contrainte employée par le système CHARADE /GANASCIA 87/ concerne la limitation du nombre de descripteurs par prémisse de la règle.

Revenons maintenant à notre exemple initial : E1 : carré (a) & rouge (a)
E2: rectangle (b) & rouge (b)

La généralisation de ces deux exemples était G = Rouge (x). Si on connait le théorème suivant :
$\forall$ x [carré (x) => rectangle (x)], la généralisation G va pouvoir s'écrire : G : rectangle (x) & rouge (x).

Au cours du processus de généralisation, certains systèmes d'apprentissage utilisent, en tant qu'étape préliminaire, la technique de "mise en appariement structurel". Cette approche, développée par Y. KODRATOFF et J.G. GANASCIA est fondée sur la mise en correspondance des exemples, avant de les généraliser, à l'aide d'une théorie déductive du domaine (théorèmes). L'appariement structurel détecte les similarités entre les exemples par une transformation progressive de ceux-ci, jusqu'à obtention d'une généralisation. Les exemples à généraliser sont exprimés dans la logique des prédicats du premier ordre. Pour illustrer cette définition, prenons l'exemple simple suivant :

E1 :	chat (mimi) & près (mimi, porte)
E2 :	chat (minette) & près (radiateur, minette)
Théorème :	$\forall$ x $\forall$ y [près (x, y) <=> près (y, x)]

A ce niveau, E1 et E2 ne sont pas en correspondance structurelle. Grâce au théorème, on peut transformer E2 en E2' : chat (minette) & près (minette, radiateur). Ainsi, E1 et E2' sont en correspondance structurelle (ils s'apparient structurellement) et on peut alors facilement généraliser E1 et E2 en : G = chat (x) & près (x, y) avec $x \neq y$. On remarque l'introduction des variables intermédiaires (x et y) appelées variables de généralisation, ainsi que des liens ($x \neq y$).
L'algorithme AGAPE /GANASCIA 87/ exploite le principe de généralisation par appariement structurel. Ce principe a été repris et amélioré dans le système OGUST /VRAIN 87/ qui se distingue d'AGAPE par sa capacité à expliquer et justifier le résultat de la généralisation et à gérer le choix entre théorèmes et taxonomies, sans privilégier une représentation par rapport à l'autre. Le système dérivé OGUST + intègre, de plus, la notion d'incrémentalité /VRAIN 88/. Les derniers travaux axés sur le principe d'appariement structurel ont débouché sur le système APOGEE dont la principale amélioration se traduit par une plus grande rapidité de généralisation /BISSON 89/.
En résumé, apprendre c'est généraliser et pour réaliser une généralisation, on doit disposer des éléments suivants :

- une base d'exemples d'apprentissage à généraliser,
- une théorie du domaine pour faciliter la découverte des meilleurs généralisés,
- des contraintes d'apprentissage pour restreindre l'espace de généralisation,
- des techniques de généralisation comme l'élimination d'un facteur conjonctif, la variabilisation d'une constante, la remontée dans une hiérarchie ou l'appariement structurel.

L'examen des systèmes d'apprentissage à partir d'exemples développés actuellement met en évidence que la généralisation s'apparente à l'induction. KODRATOFF /91/ emploie les termes d'"induction symbolique" et "induction numérique" pour évoquer la généralisation.

Après avoir successivement évoqué les données d'entrée, les contraintes et les mécanismes d'un système d'apprentissage, nous abordons dans le paragraphe suivant ses données de sortie.

2.5. Données de sortie d'un processus d'apprentissage

A partir essentiellement d'une théorie du domaine, d'un ensemble expérimental d'exemples et des classes, le processus d'apprentissage s'appuyant sur des mécanismes produit des stratégies de résolution de problèmes tels que l'explication, la prédiction, la classification ou la reconnaissance. Ces connaissances sont généralement exprimées sous forme d'arbres de décision, de règles de production ou de hiérarchies de concepts. Ces trois formats de sortie sont détaillés ci-après.

2.5.1. Arbres de décision

Un arbre de décision est constitué de noeuds correspondant aux attributs des objets sélectionnés et de branches caractérisant les valeurs alternatives de ces attributs. Les feuilles de l'arbre représentent les ensembles d'objets d'une même classe. La construction d'arbres de décision est une approche de généralisation descendante. L'algorithme ID3 /QUINLAN 83/ est un cas type d'approche descendante détaillé notamment dans /QUINLAN 83/, /GANASCIA 85/, /KODRATOFF 85, 86/, /MANAGO 88, 89/, /DUPAS 90/. Rappelons simplement que ce système utilise une stratégie de recherche heuristique, selon la méthode du gradient, en optimisant un critère numérique appelé gain d'information qui a pour base l'entropie de SHANNON.

A partir :

- d'un ensemble de classes exclusives les unes des autres C1, C2, ... CK
- d'un ensemble d'exemples {E1, E2, ... En} représentés sous la forme de couples (Attribut-Valeur) et partitionnés en classes Ci,

ID3 produit un arbre de décision qui permet de reconnaître (ou de classifier) tous les exemples Ei. Cet arbre peut être ensuite utilisé pour engendrer des règles de classification.

La méthode de QUINLAN consiste à tester successivement chaque attribut pour connaître celui à utiliser en premier de façon à optimiser le gain d'information, c'est à dire l'attribut qui permet le mieux de distinguer les exemples de différentes classes. Ce principe a été appliqué à de nombreux cas et a contribué au développement de plusieurs systèmes experts, essentiellement dédiés au diagnostic. Par la suite, des travaux ont été consacrés à l'amélioration du principe de construction de l'arbre de décision. Parmi les résultats obtenus on peut citer :

- le système ASSISTANT /CESTNIK 87/ qui se propose notamment de réduire la taille de l'arbre,
- le système PRISME /CENDROWSKA 87/ qui vise l'amélioration de la stratégie de sélection (qui est basé dans ID3 uniquement sur l'attribut) en proposant une sélection basée à la fois sur le couple (attribut-valeur),
- le système KATE /MANAGO 89/ qui s'attache essentiellement à améliorer le mode de représentation des exemples, en utilisant une représentation à base de schémas (frames),

- ID4, ID5, ID5R /UTGOFF 88, 89/ qui traitent de l'aspect incrémental.

2.5.2. Règles de production

La plupart des systèmes d'apprentissage développés actuellement s'attachent à produire des connaissances sous formes de règles exploitables par un système à base de connaissances. Il s'agit de paires "condition-action" (C => A) où C est un ensemble de conditions et A une séquence d'actions. L'un des précurseurs dans ce domaine est le système INDUCE de MICHALSKI /MICHALSKI 83/ qui repose sur l'approche de généralisation ascendante. Une des applications de cet l'algorithme est le système PLANT /MICHALSKI 80/ utilisé pour la caractérisation des maladies du soja. Pour ce système, l'objectif est d'apprendre des règles permettant de diagnostiquer les maladies du soja à partir de questionnaires remplis par les experts pour chaque cas de maladie rencontré. Ces questionnaires comportaient d'une part un ensemble de descripteurs comme la couleur des feuilles ou la présence de chancre sur les tiges et d'autre part le diagnostic proposé. Chaque questionnaire diagnostiquant une maladie fournissait les exemples pour cette maladie et chaque questionnaire concluant sur une autre maladie fournissait les contre-exemples. Très schématiquement, à partir d'un concept C, d'un ensemble d'exemples E1, E2, ...En et d'un ensemble de contre-exemples CE1, CE2, ...CEn, INDUCE construit une généralisation de E1, ...En en discriminant CE1, ...CEn appelée G puis engendre la règle : G => C.

En matière d'apprentissage de règles, on distingue généralement deux formats de sortie : soit des règles isolées à un seul pas d'inférence comme dans le système INDUCE, soit un système de règles comme dans les systèmes CHARADE /GANASCIA 87/ et KBG /BISSON 91/. CHARADE utilise la logique des propositions comme langage de description des exemples alors que KBG emploie la logique des prédicats.

2.5.3. Hiérarchies de concepts

Dans le cadre de l'apprentissage par classification conceptuelle, les connaissances apprises sont généralement des hiérarchies de concepts. Il s'agit de descriptions taxinomiques qui partagent une classe d'objets en sous-classes. Ces hiérarchies se distinguent, selon que le regroupement des objets en classes conceptuelles s'effectue progressivement au fur et à mesure de l'observation de nouveaux objets (approche incrémentale) ou en une seule fois (approche non incrémentale) à partir de l'ensemble des objets supposés disponibles dès le départ.

Formation incrémentale d'une hiérarchie de concepts : cette approche, qualifiée également d'apprentissage à partir "d'observation", traite les objets au fur et à mesure de leur apparition. Les objets ne sont pas tous disponibles au départ et le regroupement conceptuel est graduellement revu et ajusté à chaque occurrence d'un nouvel objet, sans avoir à traiter à nouveau l'ensemble des objets déjà considérés. Cette approche peut être qualifiée d'incrémentale car elle intègre l'évolutivité des connaissances et à ce titre, elle paraît selon /THAYSE 91/ plus intéressante que l'approche non incrémentale. Les algorithmes COBWEB /FISHER 87/, UNIMEM /LEBOWITZ 87/, CLASSIT /GENNARI 89/ et ADECLU /DECAESTECKER 89/ en sont des exemples types.

Formation non incrémentale d'une hiérarchie de concepts : dans cette approche non incrémentale, les objets à caractériser ou à regrouper conceptuellement par classes sont traités en une seule fois par le processus d'acquisition et doivent par conséquent exister initialement. L'algorithme CLUSTER /MICHALSKI 84/ constitue un exemple représentatif de cette catégorie d'apprentissage. Les classes construites par cet algorithme sont organisées en une structure hiérarchique.

2.6. Intérêt de l'approche proposée pour caractériser un processus d'apprentissage

Nous venons de caractériser un processus d'apprentissage par ses données d'entrée et de sortie, les mécanismes qu'il met en oeuvre et les contraintes qu'il doit respecter. Les principaux concepts définis dans cette caractérisation sont exploités dans le paragraphe suivant pour imposer des propriétés visant à déterminer les systèmes d'apprentissage appropriés au développement de l'outil d'aide à la certification des systèmes de transport automatisés.

3. CHOIX DES SYSTEMES D'APPRENTISSAGE

Le choix d'un système d'apprentissage adapté à une application repose généralement sur l'identification des besoins, des caractéristiques des connaissances disponibles ainsi que sur la définition des performances attendues du processus d'apprentissage.

Pour le problème de la certification, la phase d'acquisition de connaissances a permis de recenser une quarantaine d'exemples de scénarios d'accidents relatifs au risque de collision regroupés dans la Base de Connaissances Historiques des Scénarios. Cette base d'exemples d'apprentissage n'est pas complètement représentative du domaine de la certification et est entachée de bruit. L'objectif de l'étude consiste à opérer par apprentissage sur cette base afin de reproduire les activités de **classification** et d'**évaluation** des scénarios d'accidents impliquées dans la démarche évolutive, intuitive et créative de certification.

En présence d'un nouvel exemple de scénario proposé par le constructeur, l'expert certifieur s'attache à le **classer** dans une famille d'accidents existante avant de s'assurer qu'il prend en compte toutes les pannes envisageables. Pour cerner l'activité de recherche de pannes susceptibles de provoquer une situation d'insécurité, le mécanisme d'apprentissage doit produire une base de **règles** de la forme : "si symptômes alors pannes", exploitable par un moteur d'inférences.

3.1. Propriétés du système d'aide à la certification

Après avoir brièvement rappelé les caractéristiques essentielles du domaine et introduit notre approche pour le développement d'un outil d'aide à la certification, nous justifions maintenant les choix de systèmes d'apprentissage retenus en référence à l'ensemble des propriétés requises par l'outil de certification et à l'offre actuelle perçue au travers de la bibliographie. Une vision synthétique de l'ensemble est proposée dans le tableau de la figure 3.4. et les propriétés recherchées sont ensuite détaillées. Ce tableau mentionne en colonnes les principaux systèmes connus à ce jour et leur comportement au regard des propriétés requises données en lignes.

Propriétés requises par le système d'aide à la certification	Exemples de systèmes d'apprentissage															
	ID3	KATE	PRISM	ASSISTANT	AGAPE	OGUST	APOGEE	INDUCE	AQ11	KBG	CHARADE	COBWEB	UNIMEM	CLASSIT	ADECLU	CLUSTER
Apprentissage par détection de similarités : SBL (Approche synthétique)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Mode d'inférence inductif	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Traitement Symbolique-Numérique		X	X	X						X	X	X	X	X	X	X
Apprentissage par classification	X	X	X	X			?	X	X	X	X	X	X	X	X	X
Apprentissage par détection de régularités empiriques pour la construction de BC exploitables par un SBC		X	X	X			X	X	X	X	X					
Généralisation ascendante					X	X	X	X	X	X	X					
Production incrémentale de descriptions conjonctives de classes																
Production de règles		X	?	?			X	X	X	X	X					
Structuration des règles (système de règles)										X	X					
Description des exemples sous forme de propositions	X	X	X	X				X	X		X	X	X	X	X	X
Apprentissage non monotone												X	X	X	X	
Résistance au "bruit"	X	X	X	X						?	X	X	X	X	X	?
Interactivité Expert/système																

"X" : *propriété acquise*
"Blanc" : *propriété non acquise*
"?" : *pas d'informations suffisantes pour prendre position.*

Figure 3.4. : Tableau de choix d'un système d'apprentissage

Ce tableau n'est pas exhaustif et constitue une première phase de travail qui pourra par la suite être affinée, par exemple, au moyen de questionnaires adressés aux concepteurs de ces systèmes.

Les propriétés imposées à l'outil d'aide à la certification sont présentées ci-dessous.

- **Apprentissage par détection de similarités SBL** : rappelons que la méthode SBL détaillée dans le paragraphe 2.4.3. est caractérisée par la disponibilité d'un grand nombre d'exemples alimentant le système d'apprentissage et par l'absence de connaissances sur le domaine (ou une connaissance faible). Compte tenu des connaissances de certification acquises présentées au chapitre 2 qui sont essentiellement des scénarios d'accidents, le choix de l'apprentissage par détection de similarités SBL est justifié.

- **Mode d'inférence inductif** : pour la certification, les données d'entrée sont des exemples de scénarios historiques d'accidents sur lesquels le mécanisme d'apprentissage opère en vue d'induire des lois (descriptions ou règles) générales.

- **Traitement symbolique-numérique** : le traitement des données choisi est de nature symbolique-numérique. Il combine l'efficacité du traitement numérique qui permet d'opérer en présence des données de certification, bruitées et incomplètes et l'explicabilité du traitement symbolique nécessaire à la compréhension par l'utilisateur des connaissances produites.

- **Apprentissage par classification et Apprentissage par détection de régularités empiriques** : deux stratégies d'apprentissage sont requises pour assurer les deux activités impliquées dans la démarche de certification : la classification des scénarios d'accidents et la détection de régularités empiriques pour construire des bases de connaissances de certification exploitables par un système expert.

- **Généralisation ascendante** : la structure hiérarchisée des risques d'accidents relatifs aux systèmes de transport terrestres automatisés acquise auprès des experts certifieurs et présentée dans la figure 2.3. du chapitre 2 est organisée selon une approche descendante. L'un des objectifs de l'étude est d'exploiter cette hiérarchie par une approche de généralisation ascendante afin de caractériser les accidents selon leur type, risque, fonction de sécurité, sous-fonction de sécurité etc.

- **Production incrémentale de descriptions conjonctives de classes d'objets et production de règles** : l'activité de classification exige l'apprentissage incrémental non monotone de descriptions conjonctives de classes de scénarios d'accidents. L'activité d'évaluation des scénarios requiert la production de règles pour aider à la reconnaissance des pannes.

- **Structuration des règles** : le système d'apprentissage doit générer, non pas des règles isolées à un seul pas d'inférence mais un système de règles structurées qui permet la formation d'un raisonnement déductif tenant compte essentiellement de l'orientation des règles des symptômes vers les causes (pannes).

- **Description des exemples sous forme de propositions** : compte tenu du mode de représentation des scénarios d'accident (fiche descriptive des paramètres) adopté lors de la phase de formalisation des connaissances d'analyse de sécurité, les exemples d'apprentissage sont des propositions.

- **Apprentissage non monotone** : l'incrémentalité est une propriété indispensable pour traiter les connaissances évolutives caractéristiques du domaine de la certification. Elle doit être non monotone afin d'assurer la remise en cause éventuelle des connaissances apprises antérieurement. Pour garantir la non monotonie des connaissances il faut intégrer des moyens permettant de les stabiliser et par conséquent d'assurer la convergence du système.

- **Résistance au bruit** : un système d'apprentissage doit pouvoir traiter la base d'exemples de scénarios d'accidents qui est par nature incomplète et bruitée.

- **Interactivité Expert/Système** : l'apprentissage inductif est par nature incertain et produit des connaissances "plausibles" que l'expert du domaine doit valider. L'intervention de ce dernier ne doit pas se limiter, comme dans la plupart des systèmes d'apprentissage, à la fourniture des exemples d'apprentissage et de validation des résultats mais elle doit également porter sur le contrôle des connaissances apprises. Le système, quant à lui, doit argumenter son raisonnement et ses décisions. Cet apprentissage "interactif" ou "supervisé" favorise l'acquisition de nouvelles connaissances. L'association de l'expert du domaine à chaque étape du processus d'apprentissage nécessite l'élaboration d'une interface Homme-Machine conviviale. Comme on peut le constater sur le tableau de la figure 3.4., l'interactivité Expert/Système telle que nous l'avons définie, n'est prise en compte par aucun des systèmes étudiés.

L'ensemble de ces propriétés s'avère indispensable pour l'application industrielle nouvelle et complexe qu'est la certification. En consultant le tableau récapitulatif (figure 3.4.), on constate qu'aucun des systèmes d'apprentissage considérés ne satisfait, à lui seul, la totalité de ces propriétés. Toutefois, si l'on décompose notre problème en distinguant l'activité de classification de l'activité d'évaluation des scénarios d'accidents, on peut envisager d'utiliser le système CHARADE /GANASCIA 87/ pour la génération d'un système de règles mais on est contraint de développer un nouveau système de classification des scénarios d'accidents. Dans les paragraphes suivants, nous justifions le choix de CHARADE et la nécessité de concevoir le nouveau système de classification des scénarios que nous baptisons CLASCA.

3.2. Justification du choix de CHARADE

Parmi les systèmes d'apprentissage recensés qui produisent des connaissances sous forme de règles, APOGEE et KBG traitent des exemples compatibles avec la logique de premier ordre (logique de prédicats). Compte tenu du mode de représentation des scénarios sous forme de propositions, ils ne peuvent d'emblée être retenus. KATE ainsi qu'INDUCE et ses dérivés (AQ11 et AQ15) utilisent la logique propositionnelle mais ils n'engendrent pas de systèmes de règles structurés. Les règles produites par KATE et INDUCE sont isolées les unes des autres et l'absence d'enchaînement de règles de la forme A => B, B => C, C => D empêche la formation d'un raisonnement déductif complexe. CHARADE permet non seulement d'engendrer un système de règles structuré et exploitable par un moteur d'inférence, mais aussi de compléter la description des exemples fournis par l'expert pour prendre en compte des données éventuellement bruitées comme les exemples de scénarios d'accidents. Il permet d'apprendre simultanément des règles logiques certaines et des règles incertaines modulées par un coefficient de vraisemblance. Enfin, son originalité majeure réside dans sa flexibilité et sa traduction des fonctionnalités du SBC que l'on veut obtenir grâce aux contraintes qu'il met en oeuvre. Tous ces avantages sont obtenus au prix d'un apprentissage qui ne peut être incrémental. Toutefois, au niveau de l'élaboration des règles d'évaluation des scénarios d'accident, la structuration des règles est prioritaire par rapport à l'incrémentalité.

3.3. Nécessité de développer un nouveau système d'apprentissage par classification

L'analyse du tableau de la figure 3.4. en regard des propriétés escomptées pour l'activité de classification révèle des insuffisances. Celles-ci sont commentées sur les quelques systèmes d'apprentissage qui s'approchent le plus de la solution.

- CLUSTER ainsi que ID3 et ses dérivés (KATE, PRISM, ASSISTANT) sont non incrémentaux et exigent que les exemples à classer soient tous disponibles. En pratique et notamment pour la certification, il est difficile d'obtenir une liste exhaustive d'exemples à moins de consacrer un temps considérable à la phase d'acquisition auprès des experts. Ceci est d'autant plus vrai qu'on se trouve en présence d'un domaine évolutif. En revanche, l'avantage de ces systèmes est la prise en compte de la sémantique du domaine d'application, puisque la totalité des exemples qui couvrent le domaine sont présents dès le départ.

- COBWEB, UNIMEM, CLASSIT et ADECLU, traitent incrémentalement les exemples pour former une hiérarchie de concepts qui ne rend pas nécessaire la fourniture dès le départ de tous les exemples du domaine. Cependant, construire une hiérarchie conceptuelle à partir d'exemples observés individuellement, produit nécessairement des résultats non consistants et non maîtrisables sémantiquement. De plus, une hiérarchie de concepts est difficilement compréhensible par l'utilisateur dès que sa taille devient importante. Ce genre de représentation des connaissances produites est peu compatible dans le cas d'un système complexe dans lequel plusieurs modules sont impliqués.

- les mécanismes d'apprentissage internes de la majorité des systèmes de classification ne sont pas accessibles à l'expert du domaine. Concevoir un mécanisme d'apprentissage pour lequel une place prépondérante est laissée à l'expert pour juger, sur le plan sémantique, de la qualité des connaissances produites constitue une avancée intéressante. En effet, un apprentissage supervisé par l'expert est en soi une démarche susceptible de faire émerger des connaissances qui, à l'origine, n'étaient pas forcément évidentes ou même consciemment présentes à l'esprit de l'expert.

3.4. Approche de classification proposée

A l'issue des constatations émises dans le paragraphe précédent, nous proposons de débuter la phase d'apprentissage avec un lot d'exemples pré-classés par l'expert et non représentatifs du domaine, sans obliger l'expert à recenser la totalité des exemples mais en l'impliquant tout au long du processus d'apprentissage pour améliorer les connaissances acquises. De ce fait, la **sémantique** de la connaissance est prise en compte. Puis, on fait évoluer le système à chaque nouvel exemple de scénario fourni par l'expert pour former incrémentalement des descriptions conjonctives de classes d'objets, compréhensibles par l'expert et compatibles avec le système CHARADE. Cette approche, qui se situe entre les systèmes d'apprentissage non incrémentaux exigeant la présence de tous les exemples à classer et ceux traitant **incrémentalement** les exemples un par un, fait l'objet du système CLASCA, conçu et détaillé au chapitre suivant.

CONCLUSION

Ce chapitre a été consacré à l'étude des diverses méthodes, techniques et systèmes d'apprentissage automatique. L'objectif principal de cette étude a été de montrer dans quelle mesure les travaux effectués dans le domaine de l'apprentissage automatique ont pu être d'un recours intéressant pour contribuer à l'élaboration du système d'aide à la certification.

La première partie de ce chapitre a présenté un état de l'art sur le domaine de l'apprentissage qui a débouché sur le problème du choix des systèmes d'apprentissage appropriés à un domaine d'application industrielle.

Pour apporter des éléments de réponse à ce problème, nous nous sommes efforcés dans la deuxième partie de définir et synthétiser les caractéristiques de l'apprentissage automatique en précisant les données d'entrée et de sortie d'un processus d'apprentissage, les contraintes qu'il doit respecter ainsi que les mécanismes employés pour apprendre.

Ce travail de caractérisation, exploité dans la troisième partie, a permis d'imposer à l'outil d'aide à la certification un ensemble de propriétés récapitulées dans un tableau et a débouché sur le choix du système CHARADE pour la production des règles de reconnaissance des pannes et sur la nécessité de concevoir un nouveau système appelé CLASCA pour la classification des scénarios d'accidents. Ces deux systèmes d'apprentissage se complètent pour le développement du système ACASYA d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés présenté dans le chapitre suivant.

Chapitre 4 :

Conception et réalisation du système ACASYA d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés

INTRODUCTION

Dans les chapitres précédents, nous avons présenté la certification, les limites des moyens usuels d'acquisition des connaissances ainsi que la nécessité de recourir à l'apprentissage automatique pour appréhender au mieux le processus de transfert de l'expertise de certification. Le présent chapitre propose les différentes étapes de conception et de réalisation du système ACASYA d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés. Celui-ci repose pour l'essentiel sur l'utilisation conjointe des modules CHARADE et CLASCA identifiés précédemment.

La première partie décrit l'organisation fonctionnelle du système ACASYA qui se décompose en quatre principaux modules. Le premier module est relatif à l'acquisition et à la formalisation des scénarios d'accidents, les trois autres modules sont complémentaires et concernent les systèmes CLASCA, EVALSCA et GENESCA, développés respectivement pour la classification, l'évaluation et la génération des scénarios.

La deuxième partie détaille la conception du système CLASCA de classification des scénarios d'accidents. CLASCA est un système d'apprentissage symbolique-numérique, il est incrémental et s'inscrit dans le cadre de l'apprentissage par détection de similarités (SBL) à partir d'exemples. L'une de ses principales caractéristiques résulte de l'introduction d'une démarche coopérative avec l'expert du domaine pour compléter, contrôler et valider les connaissances apprises.

La troisième partie est consacrée à la présentation du système EVALSCA d'évaluation de la complétude d'un scénario d'accident. EVALSCA s'organise autour de deux modules principaux : un mécanisme d'apprentissage de règles CHARADE et un générateur de systèmes experts IS2. EVALSCA favorise l'aide à la génération de nouveaux scénarios d'accidents par la déduction d'une ou plusieurs pannes résumées (PR) non prises en compte par le constructeur et susceptibles de mettre en cause la sécurité du système de transport automatisé.

Le module GENESCA qui constitue le troisième niveau de traitement de la méthodologie n'est pas développé dans le cadre de cette thèse.

La dernière partie de ce chapitre présente l'architecture de la maquette du système ACASYA. La réalisation de cette maquette a nécessité le développement de trois modules principaux : CLASCA, EVALSCA et l'interface Homme/Machine.

1. CONCEPTION DU SYSTEME ACASYA

Le système ACASYA d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés /HADJ-MABROUK 92a/ repose sur l'utilisation conjointe des techniques d'acquisition de connaissances et d'apprentissage automatique. Les deux principales caractéristiques de cet outil sont la prise en compte de l'incrémentalité nécessaire à une évolution progressive des connaissances apprises par le système et de la coopération Homme/Machine pour permettre à l'expert certifieur de corriger et compléter les connaissances initiales et/ou produites par le système. Contrairement à la majorité des systèmes d'aide à la décision qui s'adressent à un utilisateur non expert, l'outil conçu ici est un système destiné à coopérer avec un expert pour l'assister dans sa décision. Les paragraphes suivants présentent respectivement le principe général et l'organisation fonctionnelle du système ACASYA.

1.1. Principe général du système ACASYA

ACASYA est organisé de façon à reproduire en grande partie la stratégie utilisée par les experts certifieurs. Pour la resituer brièvement, la démarche de certification fait appel à une première phase de reconnaissance qui apparente le scénario étudié à une famille de scénarios connue de l'expert. Cette phase rend nécessaire la définition de classes de scénarios. Dans une seconde phase, le certifieur évalue le scénario afin de dégager d'éventuelles situations d'insécurité non considérées par le constructeur. Ces dernières stimulent l'expert pour la formulation de nouveaux scénarios d'accidents.

1.2. Organisation fonctionnelle du système ACASYA

Comme le montre la figure 4.1, nous retrouvons dans cette organisation quatre modules essentiels /HADJ-MABROUK 92b/. Le premier module de formalisation concerne l'acquisition et la représentation d'un scénario, il est du ressort de la phase d'acquisition de connaissances. Les trois autres modules, CLASCA, EVALSCA et GENESCA, conformément au principe général énoncé précédemment, concernent les problèmes de classification et d'évaluation d'un scénario d'accident ainsi que le problème de l'aide à la génération d'autres scénarios.

Plus formellement, la méthodologie d'aide à la certification est fondée sur la cohabitation des deux modèles précédemment définis : un modèle "générique" de représentation des scénarios d'accidents qui s'articule autour d'une description statique et d'une description dynamique caractérisant un scénario et un modèle du raisonnement "implicite" de l'expert certifieur. Le raisonnement expert met à contribution trois grandes activités : la classification, l'évaluation et la génération des scénarios. L'objectif principal de l'étude est de conjuguer ces deux modèles et d'utiliser des techniques d'apprentissage, en vue de rendre le plus explicite possible le modèle de l'expert et par conséquent permettre de reproduire la démarche experte de certification.

Le paragraphe suivant présente brièvement les trois niveaux de la méthodologie d'aide à la certification impliquée dans le système ACASYA.

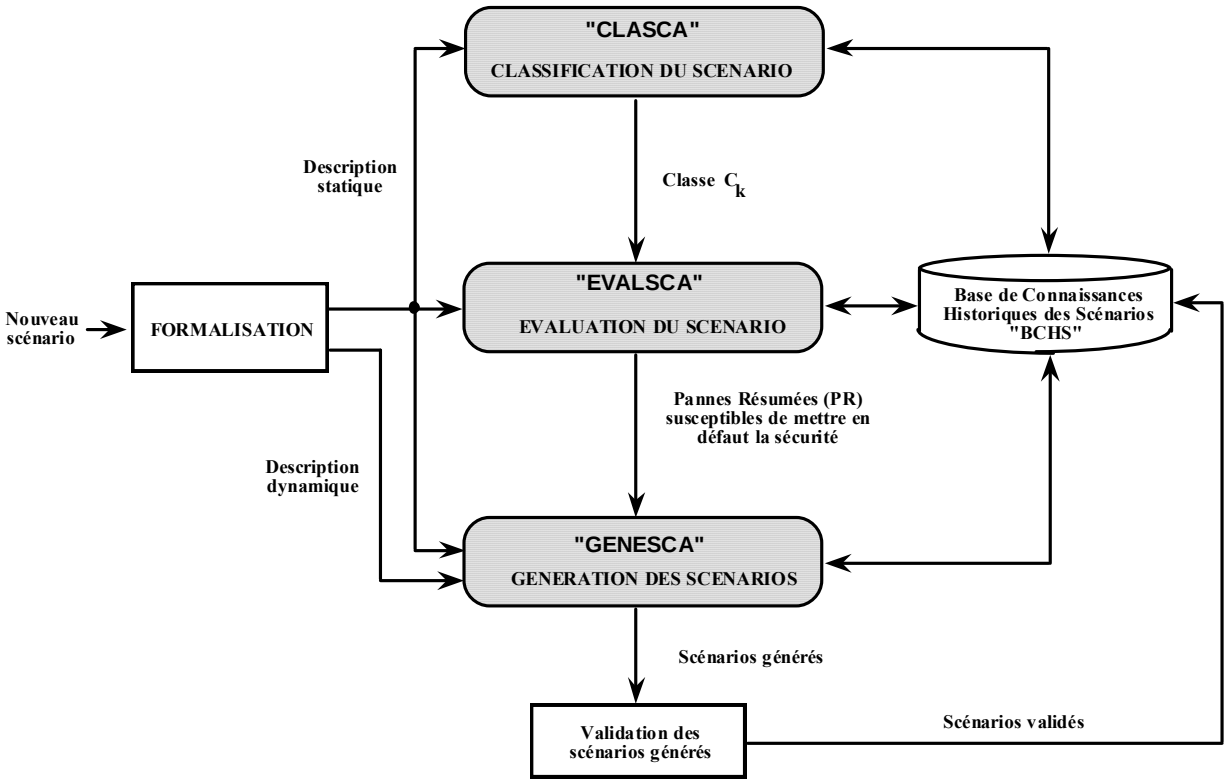


Figure 4.1. : Organisation fonctionnelle du système ACASYA

1.2.1. Classification des scénarios d'accidents : CLASCA

Le premier niveau concerne la recherche de la classe d'appartenance d'un nouveau scénario proposé par le constructeur, dans le but de présenter à l'expert les scénarios historiques partiellement ou totalement similaires à ce dernier. Ce mode de raisonnement est analogue à celui des experts qui tentent de rapprocher les situations décrites par les scénarios du constructeur de certaines situations vécues ou envisagées sur des équipements déjà certifiés ou homologués. La classification d'un nouveau scénario passe par les deux principales étapes suivantes :

- Une étape de caractérisation (ou généralisation) pour la construction d'une description en intention de chacune des classes de scénarios. Cette étape opère par détection de similarités dans un ensemble de scénarios historiques de la BCHS, pré-classés par l'expert du domaine.

- Une étape de déduction (ou classification) pour la recherche de la classe d'appartenance d'un nouveau scénario par évaluation d'un critère d'adéquation. Dans cette étape, les descripteurs du nouveau scénario (description statique) sont confrontés aux descriptions des classes induites précédemment.

Ce premier niveau de traitement permet non seulement de fournir une aide à l'expert certifieur en lui proposant les scénarios similaires au scénario à traiter mais il permet aussi de restreindre l'espace d'évaluation et de génération des nouveaux scénarios en focalisant sur une seule classe Ck de scénarios.

1.2.2 Evaluation des scénarios d'accidents : EVALSCA

Le deuxième niveau de traitement considère la classe d'appartenance déduite par CLASCA pour évaluer le contenu et la cohérence du scénario constructeur. La démarche d'évaluation est centrée sur les Pannes Résumées (PR) impliquées dans le scénario constructeur. L'évaluation d'un scénario constructeur fait intervenir les deux modules suivants :

- Un mécanisme d'apprentissage de règles CHARADE /GANASCIA 87/ qui permet de déduire les fonctions de reconnaissance des PR et donc d'engendrer une base de règles d'évaluation.
- Un moteur d'inférence qui exploite cette base de règles en vue de déduire les PR à considérer dans le scénario constructeur.

Le but du module EVALSCA est de confronter la liste des PR proposées dans un scénario constructeur à la liste des PR historiques archivées pour stimuler la formulation de situations d'insécurité non prévues par le constructeur. Cette démarche d'évaluation permet d'attirer l'attention de l'expert certifieur sur d'éventuelles pannes non prises en compte par le constructeur et susceptibles de mettre en cause la sécurité du système de transport. En ce sens, elle pourra favoriser la génération de nouveaux scénarios d'accidents.

1.2.3. Génération des scénarios d'accidents : GENESCA

En complément aux deux niveaux de traitement précédents qui font intervenir la description statique du scénario (paramètres descriptifs), le troisième niveau fait appel à la description dynamique du scénario (le modèle de Pétri et/ou le tableau de séquencement du marquage). L'aide à la génération d'un nouveau scénario repose sur l'injection d'une PR, déclarée envisageable par le niveau précédent, dans un séquencement particulier d'évolution du marquage du réseau de pétri.

Nos travaux se sont limités à l'étude approfondie des deux premiers niveaux de traitement. Nous en présentons les détails dans la suite.

2. CONCEPTION DU SYSTEME CLASCA

Ce paragraphe présente successivement les propriétés du système d'apprentissage CLASCA, son principe général de fonctionnement, la technique de prévention du bruit utilisée ainsi que ses étapes de développement.

2.1. Principales propriétés de CLASCA

CLASCA est un système d'apprentissage par classification à partir d'exemples /HADJ-MABROUK 92c/. Il est inductif, incrémental et intègre une démarche coopérative Homme/Machine pour contrôler et valider les connaissances apprises, maintenir la vigilance de l'expert et améliorer son aptitude à mieux formaliser et structurer ses connaissances. Notre étude qui porte sur un problème réel de certification auquel sont associés

des enjeux économiques et sécuritaires importants, se caractérise par la nécessité de travailler dans un univers bruité, d'opérer sur des connaissances incomplètes, incertaines et évolutives. Si l'apprentissage dans CLASCA est non monotone au sens de SEBAG /90/, la pérennité et la stabilité des connaissances apprises n'en sont pas moins garanties.

CLASCA s'inscrit dans le cadre d'une approche Symbolique-Numérique /DIDAY 89/, /BOUCHON 90/, /KODRATOFF 89, 91/. Les exemples (scénarios) manipulés sont de nature symbolique. Un exemple est décrit par une conjonction de descripteurs (Attribut/valeur). Un descripteur peut être de type énuméré, multivalué, inconnu (ne prendre aucune valeur), clé, commentaire, minimal ou probable.

La connaissance statique du domaine est décrite par une taxonomie d'accidents impliqués dans les systèmes de transport terrestres automatisés (STA). Les connaissances apprises par CLASCA sont des descriptions caractéristiques en intention de classes d'exemples qui en constituent l'extension. La composante numérique, indispensable pour améliorer l'efficacité de l'apprentissage, met à contribution des critères et fonctions mathématiques tels que le taux d'adéquation ou la fréquence d'apparition des descripteurs.

2.2. Principe général de CLASCA

CLASCA élabore incrémentalement des descriptions conjonctives de classes de scénarios historiques en vue d'une part de caractériser un ensemble de situations d'insécurité et d'autre part de reconnaître et d'identifier un nouveau scénario soumis pour évaluation aux certifieurs. Ces descriptions de classes sont construites à partir :

- d'un langage de description contenant un ensemble de descripteurs (ou paramètres descriptifs d'un scénario), recensés lors de la phase d'extraction des connaissances,
- d'un ensemble d'exemples d'apprentissage non exhaustif décrit dans ce même langage,
- d'une partition de ces exemples en classes préétablies par l'expert,
- d'un choix des paramètres d'apprentissage permettant de régler le couple rapidité de convergence/qualité de la description obtenue.

CLASCA est constitué de quatre modules principaux : acquisition, apprentissage, classification et validation des scénarios (figure 4.2). Le premier module permet l'**acquisition** des exemples d'apprentissage et des exemples à classer auprès des experts du domaine. Les scénarios historiques classés par l'expert de certification constituent les exemples d'apprentissage. Un nouveau scénario soumis à évaluation est un exemple à classer. Le module d'**apprentissage** induit, pour un problème de sécurité donné, une description caractéristique de chaque classe de scénarios historiques. Les connaissances engendrées par ce module sont les descriptions de classes de scénarios qui permettent ensuite au module de **classification** de rechercher la classe d'appartenance d'un nouveau scénario. Cette classification doit être **validée** par l'expert et les résultats obtenus permettent d'enrichir la base d'exemples d'apprentissage. Le module d'apprentissage opère à partir des résultats du dernier apprentissage et d'un nouvel exemple (d'apprentissage ou à classer), pour induire une mise à jour des descriptions de classes. En ce sens, ce processus de classification est incrémental. Au cours de la phase de validation, l'expert peut recourir au système **EVALSCA** pour conforter sa décision.

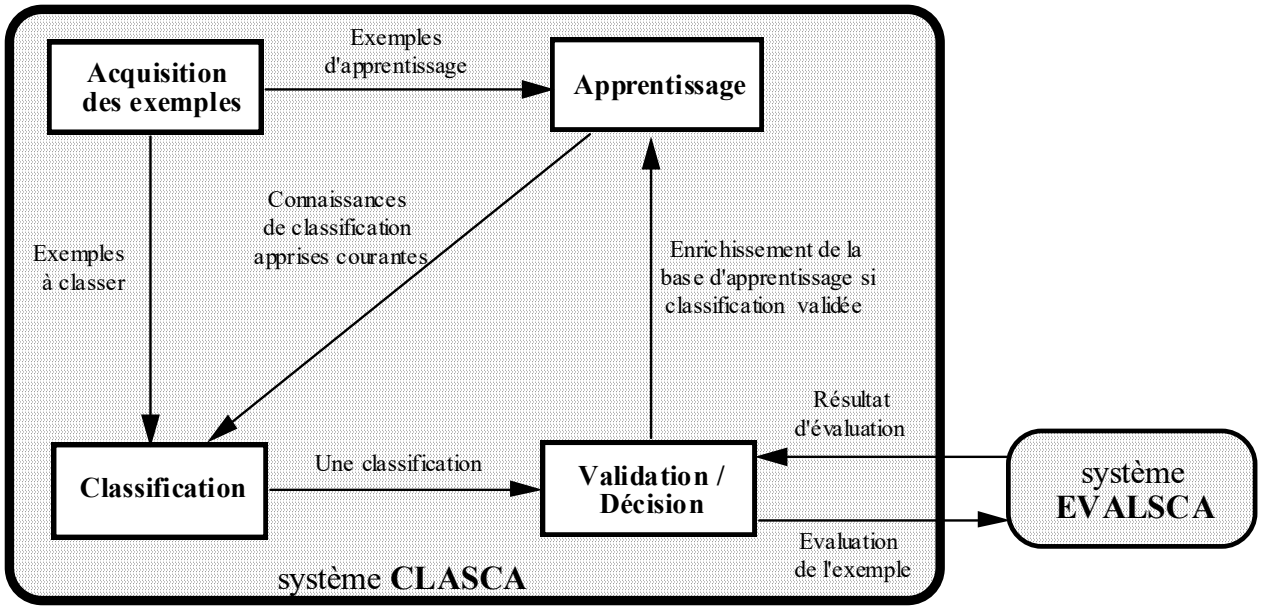


Figure 4.2 : Processus incrémental de classification : CLASCA

2.3. Le traitement du bruit dans CLASCA

La qualité des résultats produits par un système d'apprentissage dépend fortement de la qualité de l'ensemble d'exemples d'apprentissage. La majorité des systèmes d'apprentissage supposent que les données soumises au module d'apprentissage ne sont pas entachées d'erreurs. Ces suppositions sont trop restrictives pour des applications réelles. Pour obtenir une base d'exemples cohérente et complète, il convient de doter le système d'apprentissage d'une capacité à résister au bruit. Il est toujours préférable de prévenir le bruit plutôt que de le traiter. Cette façon de considérer le problème argumente en faveur d'une utilisation conjointe des techniques d'acquisition de connaissances et de l'apprentissage automatique. L'acquisition de connaissances a dans ce cas pour objectif de faire émerger des connaissances de contrôle de la "qualité" des scénarios.

Dans CLASCA, prévenir le bruit revient à :

. Appliquer une approche catégorielle pour la résolution du problème,
. Choisir soigneusement le langage de description des exemples,
. Traiter l'incohérence des données.

2.3.1. Approche catégorielle de résolution du problème

Afin de minimiser l'occurrence de bruit par défaut d'expérience ou par utilisation d'exemples hors contexte, nous avons appliqué une approche catégorielle rigoureuse basée sur une décomposition stricte de l'ensemble des scénarios en sous ensembles associés chacun à un risque particulier tels que la collision qui fait l'objet de notre étude.

2.3.2. Langage de description des exemples

Le langage de description des exemples doit être parfaitement compréhensible par l'expert du domaine et sémantiquement riche pour qu'il puisse véhiculer les concepts de l'expertise. Le langage retenu qui dérive directement du formalisme utilisé pour décrire l'aspect statique des scénarios d'accident est basé sur une représentation classique par des couples (attribut/valeur). Pour caractériser ce langage, nous avons défini sept types de descripteurs (attribut/valeur) :

- **Descripteur énuméré** : dans la description d'un exemple de scénario, ce type de descripteur prend une seule valeur parmi un ensemble de valeurs possibles (domaine de définition). Par exemple, l'attribut "Principe de Cantonnement" (PC) prend soit la valeur "canton mobile" soit la valeur "canton fixe".

- **Descripteur multivalué** tel que (FRR = suivi des trains ET initialisation ET accostage). Ce type de descripteur peut prendre plusieurs valeurs à la fois parmi l'ensemble des valeurs possibles. Il s'exprime sous forme d'une conjonction (&) de valeurs.

- **Descripteur inconnu** qui ne prend aucune valeur dans la description d'un exemple (valeur manquante d'attribut) pour une situation particulière.

- **Descripteur clé** pour une classe d'exemples donnée. Les descripteurs "clés" sont les descripteurs jugés par l'expert les plus pertinents pour caractériser une classe de scénarios d'accidents.

- **Descripteurs minimaux** pour décrire un exemple de scénario d'accident. Contrairement aux descripteurs clés qui caractérisent une classe de scénarios d'accidents, les descripteurs "minimaux" sont constitués des attributs indispensables pour caractériser un exemple de scénario d'accident. Ils définissent les conditions nécessaires (mais non suffisantes) pour qu'un exemple soit recevable. Quatre descripteurs "minimaux" sont identifiés lors de la phase d'acquisition de connaissances pour décrire un scénario d'accident : Risque (R), Fonction en Rapport avec le Risque (FRR), Zones Géographiques (ZG) et Pannes Résumées (PR). Un exemple de scénario est déclaré complet si :

  - les quatre descripteurs "minimaux" définis par l'expert sont présents dans la description du scénario,
  - parmi les descripteurs non "minimaux" (non pertinents) pour décrire un exemple de scénario, certains ne prennent aucune valeur, on considère qu'ils n'ont pas d'intérêt pour caractériser un exemple. Au lieu de leur attribuer n'importe quelle valeur possible comme dans l'algorithme PSG /ROGER 91/, on les considère comme dans /MANAGO 89/ en leur associant la valeur "inconnu". En revanche, on ne tolère pas de valeur manquante pour les attributs pertinents (descripteurs "minimaux").

- **Descripteur probable** tel que (ZG = terminus OU ligne OU limite de tronçon). On appelle descripteur "probable" un descripteur qui peut prendre plusieurs valeurs, déclarées probables par l'expert, parmi l'ensemble des valeurs possibles formant le domaine de définition. Dans l'exemple ci-dessus, un

scénario d'accident peut se passer dans plusieurs Zones Géographiques (ZG), il peut se passer en terminus ou en ligne ou en limite de tronçon. Il s'agit d'une disjonction de valeurs par opposition à un descripteur de type multivalué.

. **Descripteur commentaire** : à ce type de descripteur est associée une valeur exprimée sous forme d'un commentaire. Dans le cadre de la certification, ce type de descripteur est employé pour décrire les Pannes Résumées (PR) et les Solutions Adoptées (SA). Par exemple :

PR25 = Pénétration d'une rame sur un canton par recul

SA36 = Interdire la commutation de mode E/S tant qu'un mode est en cours

2.3.3. Traitement de l'incohérence des données

Le dernier critère de prévention du bruit dans CLASCA est relatif au traitement de l'incohérence des données. Le traitement consiste à extraire auprès de l'expert les valeurs d'un attribut "Ai" qui ne doivent pas être présentes dans l'attribut "Aj" lors de la description d'un exemple. Si ces valeurs sont présentes dans "Aj" alors l'exemple est incohérent et demande à être vérifié. En matière de certification, nous avons effectué le recensement exhaustif des Fonctions en Rapport avec le Risque (FRR) dont voici un exemple (figure 4.3.) :

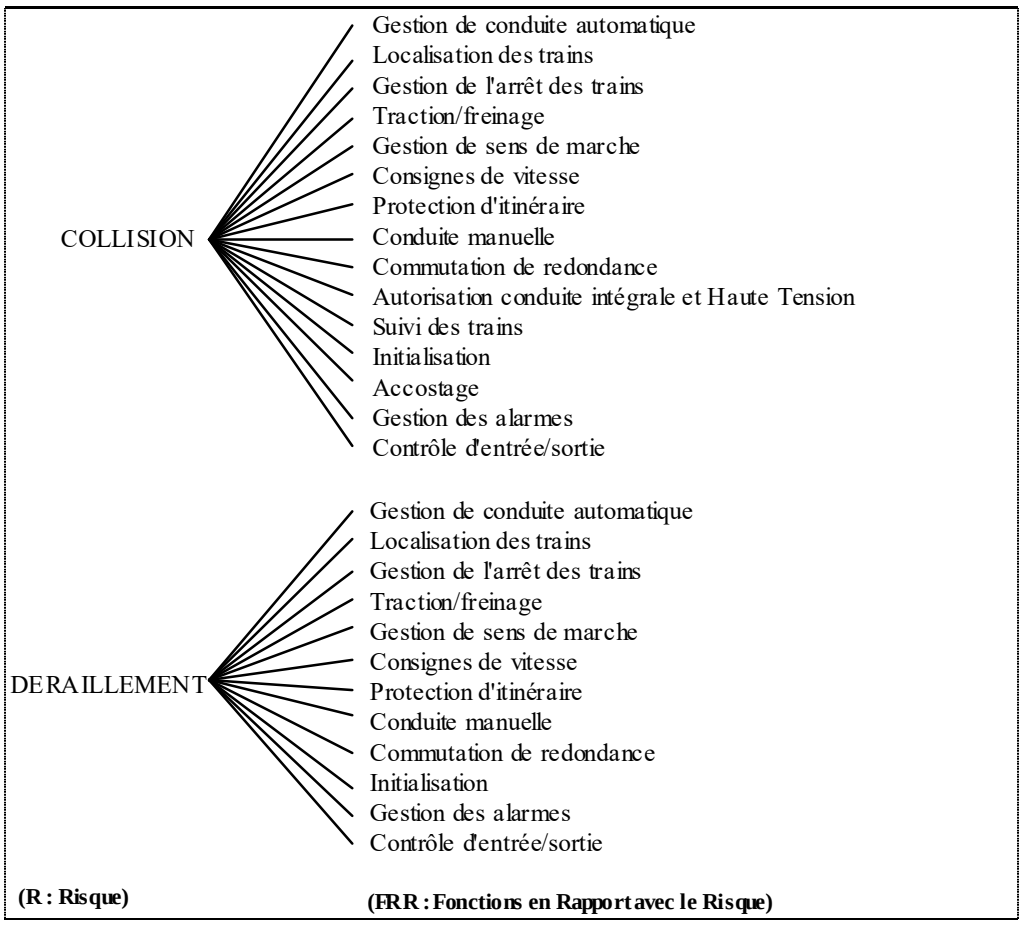


Figure 4.3. : Risque et Fonctions en Rapport avec le Risque

De ces deux taxonomies on peut déduire que les Fonctions en Rapport avec le Risque : "accostage", "suivi des trains" et "autorisation conduite intégrale et haute tension" ne doivent pas être impliquées dans la description d'un scénario d'accident relatif au risque de déraillement.

L'ensemble de ces moyens préventifs permet de réduire sensiblement le bruit sans pour cela le supprimer totalement. Pour minimiser les effets du bruit résiduel, nous utilisons un traitement statistique qui met en évidence des critères tels que la probabilité d'apparition de tel ou tel descripteur lors de la phase d'induction de descriptions de classes d'exemples. Cette technique permet de bien discriminer les classes d'objets construites et d'éviter ainsi tout chevauchement entre les classes. Dans CLASCA, une fonction d'évaluation similaire à une distance est utilisée pour prédire efficacement les décisions de classification qui sont , par ailleurs, systématiquement contrôlées par l'expert.

Après avoir présenté le traitement du bruit dans CLASCA, nous détaillons ci-après son organisation interne.

2.4. Organisation de CLASCA

Dans CLASCA, l'opération de classification se déroule en six étapes consécutives détaillées ci-après /HADJ-MABROUK 92d/ :

- Agrégation
- Induction des descriptions conjonctives des classes d'exemples
- Classification d'un nouvel exemple
- Evaluation des connaissances apprises
- Evolution et amélioration des connaissances produites
- Etude de convergence et de stabilité des connaissances apprises

2.4.1. Etape d'agrégation

La tâche d'agrégation consiste à partitionner l'ensemble des scénarios collectés en un ensemble de classes de scénarios d'accidents. Cette partition en classes est réalisée par l'expert du domaine par l'application d'une technique de tri voisine du tri conceptuel. Chaque groupe d'exemples décrit implicitement un problème de sécurité particulier que l'expert cherche à caractériser. Le but est d'élaborer les définitions en extension des classes d'objets. L'avantage majeur de cette méthode de regroupement réside dans la prise en compte de l'aspect sémantique des connaissances du domaine avant de procéder à la phase d'apprentissage automatique.

2.4.2. Induction des descriptions conjonctives des classes d'exemples

Cette étape opère par généralisation sur les classes préétablies en vue d'induire pour chaque classe une description en intention qui d'une part caractérise la partition effectuée par l'expert et d'autre part permet d'identifier la classe d'appartenance d'un nouvel exemple. Chaque description apprise est caractérisée par un ensemble de conjonctions de Triplets : (< Attribut> < Valeur> <Fréquence>). Une fréquence d'apparition est calculée pour chaque descripteur (attribut/valeur) afin de limiter la perte d'information.

Ce calcul s'apparente à celui utilisé dans l'algorithme COBWEB /FISHER 87/.

Généraliser nécessite de choisir une technique. Nous avons écarté a priori la généralisation par "élimination d'un facteur conjonctif" qui ne prend en compte que les descripteurs communs aux exemples et en cela n'est pas compatible avec le caractère évolutif et incomplet de la base d'exemples de scénarios. Nous avons donc opté pour la technique de généralisation qualifiée d'"induction statistique" /KODRATOFF 91/ qui permet d'induire une caractérisation de la classe tenant compte de la totalité des descripteurs en évaluant leur pertinence en fonction de leur fréquence d'apparition dans la classe. Ainsi, contrairement à la généralisation par "élimination d'un facteur conjonctif", cette technique évite toute perte d'information. Les paragraphes suivants présentent les deux phases de construction de la description d'une classe.

2.4.2.1. Calcul des fréquences d'apparition des descripteurs

L'objectif est de déterminer la fréquence d'apparition τ dans une classe d'exemples choisie C_k, de l'attribut A de rang n avec une valeur donnée Vm : $\tau_m^n (C_k)$

Soit ξ un ensemble de N exemples du domaine considéré : $\xi = \{ E_i , i \in \{ 1, 2, \ldots N\}\}$.

Soient des classes C_k , au nombre de K, formant une partition de ξ : $C_k C \xi$ pour k =1, 2, …K;

$C_k \cap C_{k'} = \varnothing \ \forall \ k \neq k'$ et $\bigcup_{k=1}^{K} C_k = \xi$.

On désigne les éléments de la classe C_k par e_p^k où 1 Š p Š Card C_k : Avec p rang de l'exemple dans la classe.

Une classe C_k est représentée par une description formée d'une conjonction de descripteurs. Un descripteur est un couple (attribut/valeur) noté (A_n, V_m^n), où A_n désigne un attribut et V_m^n désigne une valeur prise par l'attribut A_n. En matière de certification des systèmes de transport à conduite automatique, un descripteur peut prendre plusieurs valeurs à la fois (multivalué).

A tout descripteur (A_n, V_m^n), on associe une application D_m^n définie de la manière suivante :

$D_m^n : \xi \rightarrow \{0, 1\}$

$$E_i \rightarrow D_m^n (E_i) = \begin{cases} 1 \text{ si le descripteur } (A_n, V_m^n) \text{ prend la valeur } V_m^n \text{ dans l'exemple } E_i \\ 0 \text{ sinon} \end{cases}$$

$\tau_m^n (C_k)$ Désigne la probabilité que l'attribut A_n prenne la valeur V_m^n dans l'exemple e_p^k. et correspond à la **fréquence d'apparition** de la valeur V_m^n dans la classe C_k :

$$\tau_m^n (C_k) = \frac{\sum_{p=1}^{\text{Card } C_k} D_m^n (e_p^k)}{\text{Card } C_k} \qquad (1)$$

Cas particulier :

Parmi l'ensemble des descripteurs recensés pour caractériser un exemple de scénario d'accident, le descripteur noté (A_4, V_m^4) est particulier. A_4 désigne l'attribut "Zone Géographiques (ZG)" et V_m^4 désigne le domaine de définition de cet attribut. La particularité résulte du fait que, pour un exemple de scénario donné, cet attribut A_4 peut prendre une ou plusieurs valeurs considérées par l'expert comme les plus **probables**. Par exemple : ZG = terminus ou ligne ou station. Cela peut être interprété par l'expert comme suit : le scénario peut se situer en terminus ou en ligne ou encore en station. Pour résoudre ce cas particulier, deux solutions sont envisageables :

- on duplique autant de fois l'exemple qu'il y a de valeurs prises par l'attribut. Cette première solution est prise en compte dans le système CHARADE /GANASCIA 87/.
- on calcule la probabilité relative à chaque valeur prise par l'attribut. Compte tenu des expérimentations effectuées, cette deuxième solution nous semble plus juste. Nous proposons donc d'améliorer l'application définie ci-dessus comme suit.

Pour le descripteur (A_4, V_m^4), on associe l'application D_m^4 suivante :

$D_m^4 : \xi \rightarrow \{ 0, 1/5, 1/4, 1/3, 1/2, \}$

$$E_i \rightarrow D_m^4(E_i) = \begin{cases} \frac{1}{Hi} & \text{Si l'attribut } A_4 \text{ peut prendre la valeur } V_m^4 \text{ dans l'exemple } E_i \\ 0 & \text{Sinon} \end{cases}$$

Hi désigne le nombre de valeurs que prend l'attribut A4 dans l'exemple Ei

Le domaine de définition du descripteur Zone Géographique est composé des cinq valeurs : {terminus, station, ligne, zone_d'injection_de_rames, limite_de_tronçon}. La valeur de Hi varie entre 1 et 5.

2.4.2.2. Discrimination : recherche de la description courante d'une classe

Cette phase distingue les descripteurs significatifs caractérisant une classe C_k de ceux qui sont non significatifs car relevant du bruit. Une opération de filtrage des descripteurs selon un seuil de discrimination (sd) prédéfini est nécessaire pour scinder la description de la classe C_k en deux descriptions (figure 4.4) dites "caractéristique" et "non significative". En effet, dans la base d'exemples d'apprentissage acquise et incomplète par nature, une information non significative à un moment donné peut devenir caractéristique du domaine, compte tenu de l'acquisition de nouveaux exemples. On appelle descripteur caractéristique d'une classe Ck, un descripteur (A_n, V_m^n) tel que $\tau_m^n(C_k) \geq$ sd. Un descripteur (A_n, V_m^n) tel que $\tau_m^n(C_k) <$ sd est appelé descripteur non significatif.

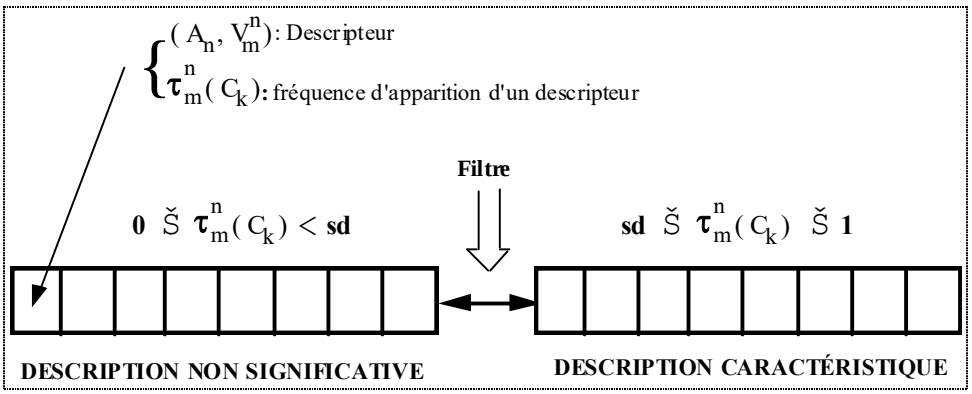


Figure 4.4 : principe de construction d'une description de classe d'exemple

Pour la déduction de la classe d'appartenance d'un nouvel exemple, seule la description caractéristique est prise en compte. Pour exemple, la description caractéristique de la classe "séquence d'initialisation" induite par CLASCA sous forme de conjonctions (<Attribut> <Valeur> <Fréquence>) est la suivante :

	(<PC> <canton mobile>	<0.71>)
&	(<FRR> <initialisation>	<1.00>)
&	(<ZG> <ligne>	<0.86>)
&	(<AI> <opérateur au PCC>	<0.86>)
&	(<AI> <PA sans redondance>	<0.86>)
&	(<AI> <2 rames>	<1.00>)
&	(<FI> <consignes>	<0.86>).

La description d'une classe est de plus enrichie par la prise en compte des Pannes Résumées (PR) impliquées. Ces PR seront exploités par la suite par le système EVALSCA.

PR impliquées dans la classe "séquence d'initialisation :

PR1 : Pénétration d'une rame sur un canton occupé par recul
& PR2 : Rame en panne d'émetteur anticollision
& PR3 : Masquage d'alarme par initialisation
& PR9 : Pénétration d'une rame sur un canton occupé
& PR10 : Rétablissement erroné de FS/HT
& PR11 : Elément invisible sur la zone de conduite automatique intégrale
& PR19 : Elément muet
& PR20 : Désaccouplement intempestif
& PR27 : Poussage intempestif

Nous venons de présenter l'étape d'apprentissage de descriptions de classes d'exemples. Ces connaissances induites sont ensuite exploitées pour classifier un nouvel exemple.

2.4.3. Classification d'un nouvel exemple

Dans cette phase, un nouvel exemple de scénario est intégré au sein d'une classe C_k existante. Cette opération nécessite la définition d'un paramètre de classification appelé "taux d'adéquation" qui mesure le degré de ressemblance entre le nouvel exemple et chacune des classes préexistantes. Ce taux d'adéquation est proche du critère d'adéquation utilisé par l'algorithme ADECLU /DECAESTECKER 89/ et est caractérisé de la façon suivante :

$$\text{Taux d'adéquation entre exemple } E_i \text{ et classe } C_k = \frac{\text{La somme des fréquences d'apparition des descripteurs caractéristiques de la classe } C_k \text{ présents dans l'exemple } E_i}{\text{La somme des fréquences d'apparition des descripteurs caractéristiques de la classe } C_k}$$

Ce taux d'adéquation T_{ad} entre le nouvel exemple E_i et une classe C_k donnée est défini de la manière suivante (formule 2) :

$$T_{ad\,\mathbf{1}}(E_i, C_k) = \frac{\sum_{(m,n) / \tau_m^n(C_k) \neq 0} D_m^n(E_i) \times \tau_m^n(C_k)}{\sum_{(m,n) / \tau_m^n(C_k) \neq 0} \tau_m^n(C_k)} \qquad (2)$$

Ce taux d'adéquation basé sur des calculs statistiques est purement **numérique**. Nous proposons de l'affiner pour tenir compte de la sémantique du domaine d'application. L'idée consiste à extraire de l'ensemble des descripteurs identifiés auprès des experts, la liste des descripteurs pertinents pour caractériser chaque classe d'exemples. Les descripteurs acquis et spécifiques à chaque classe sont appelés "descripteurs clés". Par exemple, pour la classe "séquence d'initialisation", trois descripteurs clés ont été définis par l'expert : (FRR = localisation des trains), (FRR = initialisation), (FI = consignes). Ce point de vue permet de définir un deuxième taux d'adéquation qui reflète la **sémantique** des connaissances :

$$T_{ad\,\mathbf{2}}(E_i, C_k) = \frac{\sum_{(m,n) / (A_n, V_m^n) \text{ soit "clé" de } C_k} D_m^n(E_i) \times \tau_m^n(C_k)}{\sum_{(m,n) / (A_n, V_m^n) \text{ soit "clé" de } C_k} \tau_m^n(C_k)} \qquad (3)$$

La combinaison de ces deux taux d'adéquation (2) et (3) aboutit finalement à la définition d'un taux permettant de mesurer l'adéquation entre un nouvel exemple Ei et une classe Ck, en tenant compte à la fois des aspects statistique et sémantique des données :

$$T_{ad}(E_i, C_k) = \lambda \quad T_{ad}\mathbf{1}(E_i, C_k) + (1 - \lambda)\, T_{ad}\mathbf{2}(E_i, C_k) \qquad (4)$$

Avec $0 \leq \lambda \leq 1$

"λ" est un coefficient de lissage qui peut être ajusté expérimentalement ou proposé par l'expert du domaine pour tenir compte de ses convictions profondes. Il permet d'accorder plus ou moins d'importance au traitement statistique ou sémantique. Par exemple, si $\lambda = 0$ alors le taux d'adéquation est purement sémantique et si $\lambda = 1$, il est purement statistique. Les deux types de traitements sont pris en compte équitablement dans le cas où $\lambda = 0{,}5$.

2.4.3.1. Critère de classification d'un nouvel exemple

Le taux d'adéquation (formule 4) est utilisé pour rechercher la classe candidate, susceptible de recevoir le nouvel exemple E_i. A cette fin, nous définissons un seuil de similarité "ss" tel que $T_{ad}(E_i, C_k) \geq ss$ pour que l'exemple E_i soit admissible dans la classe C_k. Ce critère de classification traduit l'exigence de similarité requise pour que le nouvel exemple appartienne à la classe. Dans un premier temps, nous considérons la valeur de ce seuil de similarité "ss" fixée arbitrairement et par la suite nous proposons un ajustement automatique de ce seuil en fonction du cardinal de la classe C_k.

Quatre cas de classement sont pris en compte dans l'algorithme :

. Scénario bien classé : si $Tad(Ei, Ck) \geq ss$ pour une seule valeur de $k \in \{1, ...K\}$,

. Scénario simultanément admissible dans plusieurs classes : si $T_{ad}(E_i, C_k) \geq ss$ pour deux valeurs ou plus de $k \in \{1, ...K\}$. En effet, s'il existe plusieurs classes candidates à l'intégration du nouveau scénario, on considère que E_i appartient à celle dont le taux d'adéquation est maximum :

$\max T_{ad}(E_i, C_k)$,
$k / Tad(Ei, Ck) \geq ss$

. Scénario tangent (cas litigieux) : si $\max T_{ad}(E_i, C_k) = Tad(E_i, C_{k1}) = Tad(E_i, C_{k2})$
$k / Tad(Ei, Ck) \geq ss$

avec $k1 \neq k2$.

En présence de deux classes C_{k1} et C_{k2}, également candidates à l'intégration du nouvel exemple de scénario E_i, il convient de définir une stratégie de sélection de la classe la plus similaire, basée sur l'emploi des descripteurs non significatifs "bruits". Dans ce cas, la formule relative au taux d'adéquation reste applicable mais la somme est étendue aux indices (m, n) correspondant à un descripteur (A_n, V_m^n) non significatif de la classe C_k tel que $\tau_m^n (C_k) < sd$.

. Scénario non classé : ∀k, si Tad (Ei, Ck) < ss alors la décision sera prise par l'expert du domaine (création d'une nouvelle classe, modification des paramètres d'apprentissage,...). Ce point sera développé ultérieurement.

2.4.3.2. Recherche des exemples les plus similaires

Après avoir sélectionné la classe d'appartenance Ck du nouvel exemple, le système propose à l'utilisateur la description de cette classe, la liste des PR et des scénarios impliqués. Lorsque le cardinal de la classe Ck est important, il est judicieux de présenter les descriptions des scénarios similaires au nouvel exemple, par ordre de similarité décroissant. A cette fin, nous introduisons une nouvelle mesure du taux d'adéquation (formule 5) entre Ei et l'ensemble des exemples e_p^k impliqués dans la classe Ck engendrée précédemment. Le max Tad (Ei , e_p^k) nous renseigne sur le ou les exemples les plus similaires au nouvel exemple E_i.

$$Tad(Ei, e_p^k) = \frac{\text{nombre de descripteurs communs à Ei et } e_p^k}{\text{nombre de descripteurs de Ei ou } e_p^k} \quad (5)$$

Après avoir présenté la procédure de recherche de la classe d'appartenance d'un nouvel exemple, il convient d'évaluer le résultat de cette classification.

2.4.4. Evaluation des connaissances apprises

La classe probable d'appartenance du nouveau scénario identifiée par apprentissage doit être validée par l'expert du domaine. CLASCA permet de gérer trois cas distincts :

- L'expert **confirme** le classement obtenu et la description de la classe Ck en question est réactualisée par le système,
- L'expert désire **évaluer** le scénario et il peut lancer le système EVALSCA,
- L'expert **conteste** la proposition du système et diverses possibilités s'offrent alors :

    - Rejeter définitivement le scénario,
    - Affiner la description du scénario (modifier, compléter, ...),
    - Modifier les paramètres d'apprentissage (sd, ss, λ, ...),
    - Créer une nouvelle classe intégrant le nouvel exemple,

. Différer sa décision et mettre en "attente" le scénario pour un examen ultérieur. Le scénario pourra être réévalué après acquisition de nouveaux scénarios historiques.

Les décisions de l'expert entrainent l'évolution des connaissances apprises antérieurement et par conséquent leur mise à jour par le système CLASCA.

2.4.5. Evolution et amélioration des connaissances produites

Dans le cas où le système CLASCA a permis d'identifier la classe d'appartenance du nouvel exemple de scénario, il est nécessaire de mettre à jour les connaissances relatives à cette classe. Une des caractéristiques fondamentales imposées au système d'apprentissage est l'incrémentalité qui permet de prendre en compte l'arrivée d'un nouvel exemple sans pour cela remettre profondément en cause les acquis. On rappelle que la description de la classe est constituée de deux types de descripteurs (caractéristiques et non-significatifs "bruit") et qu'à chaque descripteur de la classe est associé un taux τ_m^n qui mesure sa fréquence d'apparition au sein de la classe d'exemples. Tous les descripteurs dont la fréquence est supérieure à un seuil de discrimination (sd) fixé sont rangés dans la description caractéristique, les autres constituent la description non significative. La figure 4.5 schématise la démarche de mise à jour des connaissances. Lorsqu'arrive un nouvel exemple E_i de classe C_k, quatre cas peuvent se présenter :

- Phénomène de "particularisation" des descripteurs : des descripteurs jugés caractéristiques à l'instant t peuvent devenir non significatifs à l'instant t+1.
- Phénomène de "généralisation" des descripteurs : des descripteurs de la classe jugés non significatifs "bruit" peuvent devenir caractéristiques.
- Phénomènes simultanés de "généralisation" et de "particularisation".
- Apprentissage de nouveaux descripteurs.

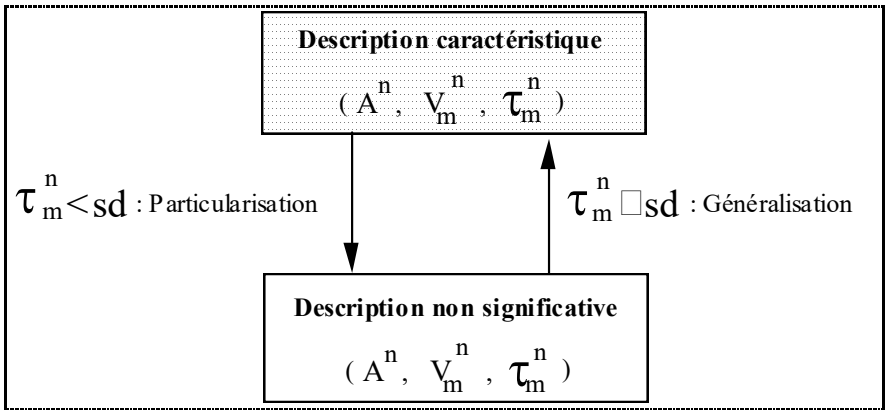


Figure 4.5 : Aspect incrémental dans l'élaboration des descriptions des classes d'exemples

Le dispositif de mise à jour des connaissances doit gérer ce phénomène de mutation de descripteurs, par la prise en compte de leur fréquence d'apparition dans la classe qui évolue au gré de l'arrivée des nouveaux scénarios. L'intérêt d'un tel traitement réside non seulement dans la distinction entre information caractéristique et information non significative mais aussi dans la classification d'un nouvel exemple, même si l'ensemble d'exemples d'apprentissage n'est pas exhaustif. Pour effectuer la mise à jour de la description de la classe considérée, nous proposons la méthode de réactualisation des taux associés aux descripteurs .

Supposons que l'exemple E_i appartienne à la classe C_k, on appelle momentanément la classe C'_k la classe obtenue après adjonction de E_i à C_k. Dans ce contexte E_i sera désigné par : $E_i = e^k_{(\text{Card } C_k)+1}$

$$C'_k = C_k \cup \{e^k_{(\text{Card } C_k)+1}\}$$

A partir de (1) on définit la fréquence d'apparition : $$\tau^n_m(C'_k) = \frac{\sum_{p=1}^{(\text{Card } C_k)+1} D^n_m(e^k_p)}{\text{Card } C'_k}$$

Ou encore :

$$\tau^n_m(C'_k) = \frac{\sum_{p=1}^{\text{Card } C_k} D^n_m(e^k_p) + D^n_m(e^k_{\text{Card } C_k+1})}{\text{Card } C_k + 1}$$

Or (1) peut s'écrire de la façon suivante : $$\sum_{p=1}^{\text{Card } C_k} D^n_m(e^k_p) = (\text{Card } C_k)\,\tau^n_m(C_k)$$

Finalement, la relation obtenue entre la nouvelle et l'ancienne fréquence d'apparition de descripteur (A_n, V^n_m) dans la classe C_k réactualisée (C'_k) s'écrit :

$$\tau^n_m(C'_k) = \frac{(\text{Card } C_k)\,\tau^n_m(C_k) + D^n_m(e^k_{\text{Card } C_k+1})}{(\text{Card } C_k)+1} \qquad (6)$$

L'application de (6) à l'ensemble des descripteurs d'une classe permet de réactualiser les taux associés à ces derniers pour la mise à jour de la description de la classe.

Le phénomène de mutation des descripteurs précédemment évoqué soulève le problème de la stabilité des descriptions des classes. Le paragraphe suivant présente une approche pour gérer la non monotonie des connaissances et tendre vers la convergence du système.

2.4.6. Etude de convergence et de stabilité

Avec l'augmentation du nombre d'exemples, la description d'une classe doit normalement converger vers une caractérisation stable et pertinente. Nous appelons cette propriété la convergence du système. Bien que la maquette soit suffisamment démonstrative pour assurer la faisabilité du système CLASCA, elle a révélé un problème concernant la convergence du système. Comme le confirme SEBAG /90/, l'incrémentalité non monotone de l'apprentissage est souhaitable mais ne garantit pas la convergence du système. Pour notre application à la certification, l'ensemble des scénarios s'enrichit au fil du temps. L'intégration d'un nouvel exemple dans une classe C_k provoque la réactualisation des fréquences d'apparition des descripteurs, ce qui entraine une éventuelle mutation des descripteurs (généralisation et/ou particularisation). Dans ce contexte, la présence inévitable du "bruit" rend nécessaire un apprentissage non monotone tel que le $\tau_m^n(C_k)$ puisse croître ou décroître en fonction de l'influence des nouveaux exemples de scénarios sur la consistance de la classe. Afin de résoudre le problème de la convergence lié à l'utilisation d'un mécanisme d'apprentissage incrémental non monotone, nous proposons l'approche suivante.

Au début de la phase d'apprentissage, lors de l'intégration de nouveaux exemples de scénarios, la description de la classe considérée n'est pas encore stable, représentative et pertinente. Un seuil de similarité "ss" faible peut donc être **toléré** pour favoriser l'acquisition et l'apprentissage des nouveaux exemples. En revanche, lorsque l'ensemble d'exemples acquis devient de plus en plus représentatif du domaine (les descriptions des classes sont de plus en plus stables et pertinentes), il convient **d'exiger** un "ss" fort pour la classification d'un nouvel exemple. Pour tenir compte de ce point de vue, nous avons convenu de faire évoluer la valeur du "ss" tout au long du cycle de classement, de façon à être de plus en plus "exigeants" au fur et à mesure de l'accroissement du cardinal de la classe considérée. L'évolution du "ss" doit donc être corrélée à celle du nombre d'exemples acquis.

Notre point de vue sur l'évolution du seuil de similarité "ss" a débouché sur la définition de deux types de convergence /HADJ-MABROUK 92b/ : la convergence "interne" qui vise la stabilité des connaissances au sein d'une classe et la convergence "globale" qui assure la stabilité des connaissances pour l'ensemble des classes et par conséquent pour l'ensemble des exemples du domaine. Ces deux types de convergence sont englobés dans une définition plus large appelée "convergence interne améliorée".

2.4.6.1 Convergence interne d'une classe

On recherche la formule exprimant la valeur du seuil de similarité ss (C_k) en fonction du Card (C_k). La forme que nous avons retenue est la suivante :

$$ss(C_k) = 1 - \alpha\ e^{1 - Card\, C_K} \quad (7)$$

avec α positif, afin que ss (C_k) tende vers 1 quand Card (C_k) tend vers l'infini et croisse en fonction de Card (C_k).

Le coefficient α est déterminé par la condition initiale. Par exemple, on cherche α de façon à ce que ss (C_k) = 0,5 avec Card (C_k) = 1. Autrement dit, lorsqu'une classe C_k vient d'être créée (au début de la phase d'apprentissage) contenant un seul exemple (Card (C_k) = 1), on peut tolérer par exemple un ss = 0,5. Par la suite ss augmente avec le cardinal de la classe.

Le choix de ss permet de déduire la valeur correspondante de α. Pour le cas où ss = 0,5, il résulte que 0,5 = 1 - α, d'où α = 0,5. Par conséquent, ss (C_k) = 1 - 0,5 α $e^{1- \text{Card Ck}}$. Numériquement, on constate qu'avec cette formule, le ss converge "trop" vite vers 1. Cela nous incite à faire intervenir dans l'exposant un deuxième coefficient β permettant d'atténuer la variation de ss. Finalement, nous proposons la formule suivante :

$$ss\ (C_k) = 1 - \alpha\ e^{\beta\ (1 - \text{Card}\, C_K)} \qquad (8)$$

Avec :

SS : seuil de similarité, avec $0 < ss < 1$

Card (C_k) : Cardinal de la classe C_k à un instant t

α : un coefficient déterminé par la condition initiale. $\alpha > 0$

β : un coefficient agissant sur le temps d'apprentissage et par conséquent sur la rapidité de convergence. Il s'agit d'un facteur d'atténuation. $\beta > 0$

La formule (8) permet d'assurer la stabilité des descriptions de classes d'exemples et favorise la convergence du système. Le seuil de similarité ss tend vers 1 de façon monotone, compte tenu de l'évolution du cardinal de la classe C_k. A chaque adjonction d'un exemple dans une classe C_k, le seuil de similarité ss est automatiquement mis à jour.

Ajustement de la rapidité de convergence par le choix du paramètre "β"

On appelle μ le seuil de similarité "critique" pour lequel l'intégration d'un nouvel exemple au sein d'une classe C_k donnée devient sélective. Il convient de fixer un μ élevé lorsque l'expert du domaine estime que l'ensemble d'exemples appris est représentatif de la classe C_k. Dans ce contexte, la description de la classe C_k est considérée comme stable et pertinente et seuls les exemples présentant une "forte" similarité peuvent être admis. Fixer la valeur du seuil de similarité critique (μ) permet de déterminer le paramètre d'apprentissage β. A partir de la formule (8) : $\mu = 1 - \alpha\ e^{\beta\ (1- \text{Card Ck})}$ et avec $0 < \mu < 1$, on déduit la formule relative au paramètre β :

$$\beta = \frac{\ln \dfrac{1 - \mu}{\alpha}}{1 - \text{Card}\, C_K)} \qquad (9)$$

Exemple de calcul de β

On rappelle que α est un coefficient déterminé par la condition initiale (Card (C_k) = 1) : on tolère par exemple un ss "faible" égal à 0,5 et on trouve α = 0,5 à partir de la formule (7). En revanche, si l'échantillon de la classe C_K est représentatif (égal par exemple à 50), on exige un ss critique (par exemple μ = 0,95) et on déduit la valeur de β = 0,047 de la formule (9). A partir de la formule (8), nous démontrons numériquement que la convergence du système est étroitement liée au paramètre β : lorsque β est faible par rapport à 1, on ralentit la convergence et on augmente le temps d'apprentissage de nouveaux exemples et inversement lorsque β se rapproche de 1. Le paramètre β représente donc un coefficient d'atténuation de la convergence.

2.4.6.2 Convergence "globale" du système

Nous venons de présenter une méthode permettant d'assurer la stabilité des connaissances par classes d'exemples. La convergence du système est prise en compte uniquement au sein d'une classe C_k. Nous pouvons compléter cette approche par l'introduction de la notion de "convergence globale" du système qui s'intéresse non plus à une classe C_k mais à la totalité des classes et par conséquent à l'ensemble des exemples du domaine considéré.

Soit ξ l'ensemble des n exemples couvrant le domaine : $\xi = \{E_i, i \in \{1, 2, ... n\}\}$
Soit ξ_o un ensemble de n_o exemples d'apprentissage : $\xi_o = \{E_i, i \in \{1, 2, ... n_o\}\}$
Parmi les n exemples couvrant le domaine, on dispose pour amorcer la phase d'apprentissage de n_o exemples dits "exemples d'apprentissage". On cherche une formule de ss, fonction du couple (n_o, n) analogue à celle établie pour la "convergence interne". Cette formule du seuil de similarité relative à la "convergence globale" du système s'écrit comme suit :

$$ss(n) = 1 - \gamma\, e^{\delta (n_o - n)} \qquad (10)$$

Avec γ > 0, δ > 0 et Lim ss (n) = 1 (en croissant)
n ∅ ∞
γ est un coefficient déterminé par la condition initiale. Par exemple pour n = n_o, ss (n_o) = 1-γ .

δ est un paramètre d'apprentissage agissant sur la rapidité de convergence du système. Par analogie avec β, δ est déterminé par la formule suivante :

$$\delta = \frac{\ln \frac{1 - \mu}{\gamma}}{n_o - n} \qquad (11)$$

2.4.6.3 Convergence "interne améliorée"

Le seuil de similarité ss croît avec le cardinal de la classe dans le cadre de la convergence interne. En revanche, il croît avec le nombre d'exemples total (n) du domaine dans le contexte de la convergence globale du système. Afin d'optimiser cette notion de convergence, en tenant compte à la fois de l'évolution des connaissances par classe et de l'ensemble des exemples du domaine, nous introduisons la notion de "convergence interne améliorée" qui consiste à conjuguer les deux approches de convergence (interne et globale). La convergence interne améliorée est décrite par la combinaison de deux formules (8) et (10) :

$$ss(C_K, n) = (1 - \alpha\, e^{\beta\,(1 - \mathrm{Card}\, C_K)})\,(1 - \gamma\, e^{\delta\,(n_o - n)}) \qquad (12)$$

ss (C_k, n) : seuil de similarité, croît de façon monotone en fonction de Card C_k et de n et tend vers 1. Il est mis à jour à chaque adjonction d'un exemple dans une classe C_k

Card C_k : cardinal de la classe C_k à un instant t qui évolue au gré de l'adjonction des nouveaux exemples dans une classe C_k

n_o : ensemble d'exemples d'apprentissage fourni au système dès le départ

n : nombre d'exemples total qui couvrent le domaine considéré

$\alpha, \beta, \gamma, \delta$: quatre coefficients calculés par quatre conditions initiales à déterminer.
$\alpha > 0, \beta > 0, \gamma > 0, \delta > 0$.

On notera que les valeurs de $\alpha, \beta, \gamma, \delta$ peuvent être fixées différemment d'une classe à l'autre. β et δ agissent sur le temps d'apprentissage et par conséquent sur la rapidité de convergence. Il s'agit de deux facteurs d'atténuation de la convergence. Le caractère modulable de β et δ permet à l'utilisateur de faire évoluer à volonté son système et d'assurer sa convergence (sous l'hypothèse que Card (C_k) ne soit pas stationnaire). A partir de la formule (12), on peut considérer le déroulement de deux phases dans l'évolution du système :

- Une phase d'apprentissage, avec par exemple : $0{,}50 < ss(C_k, n) < 0{,}80$
- Une phase de stabilité, avec par exemple : $0{,}80 < ss(C_k, n) < 1$

2.5. Architecture générale du système CLASCA

Ce paragraphe donne une vision d'ensemble du système CLASCA (figure 4.6). Il retrace le déroulement des étapes de l'algorithme en s'appuyant notamment sur les circuits utilisés par les exemples de scénarios. Au cours du processus d'apprentissage, un exemple peut passer par les 14 états suivants :

0. exemple nouveau : exemple à saisir pour vérifier sa recevabilité avant de le traiter.
1. exemple pré-classé : exemple historique vécu par les experts, archivé et dont la classe d'appartenance est connue.
2. exemple à classer : nouvel exemple à classer et/ou à évaluer.
3. exemple appris : exemple pré-classé puis traité par le module d'apprentissage (induction).
4. exemple classé : exemple appartenant à une classe Ck déduite par le module de classification.
5. exemple non classé : exemple qui n'appartient à aucune classe existante.
6. exemple bien classé : sa classification est validée par l'expert (classification correcte).
7. exemple mal classé : sa classification est contredite par l'expert (classification incorrecte).
8. exemple mis en attente : sa classification est différée.
9. exemple à intégrer dans une classe existante : sa classe d'appartenance existe.
10. exemple à intégrer dans une nouvelle classe : sa classe d'appartenance doit être créée par l'expert.
11. exemple à modifier : exemple jugé incomplet par l'expert.
12. exemple à évaluer : exemple mal classé ou non classé qui, complété de sa classe d'appartenance et de la liste des PR impliquées dans cette classe, est transmis à EVALSCA pour évaluation.
13. exemple évalué : sa complétude a été évaluée par le système EVALSCA (développé dans le paragraphe suivant) pour aider l'expert à mieux fonder sa décision.

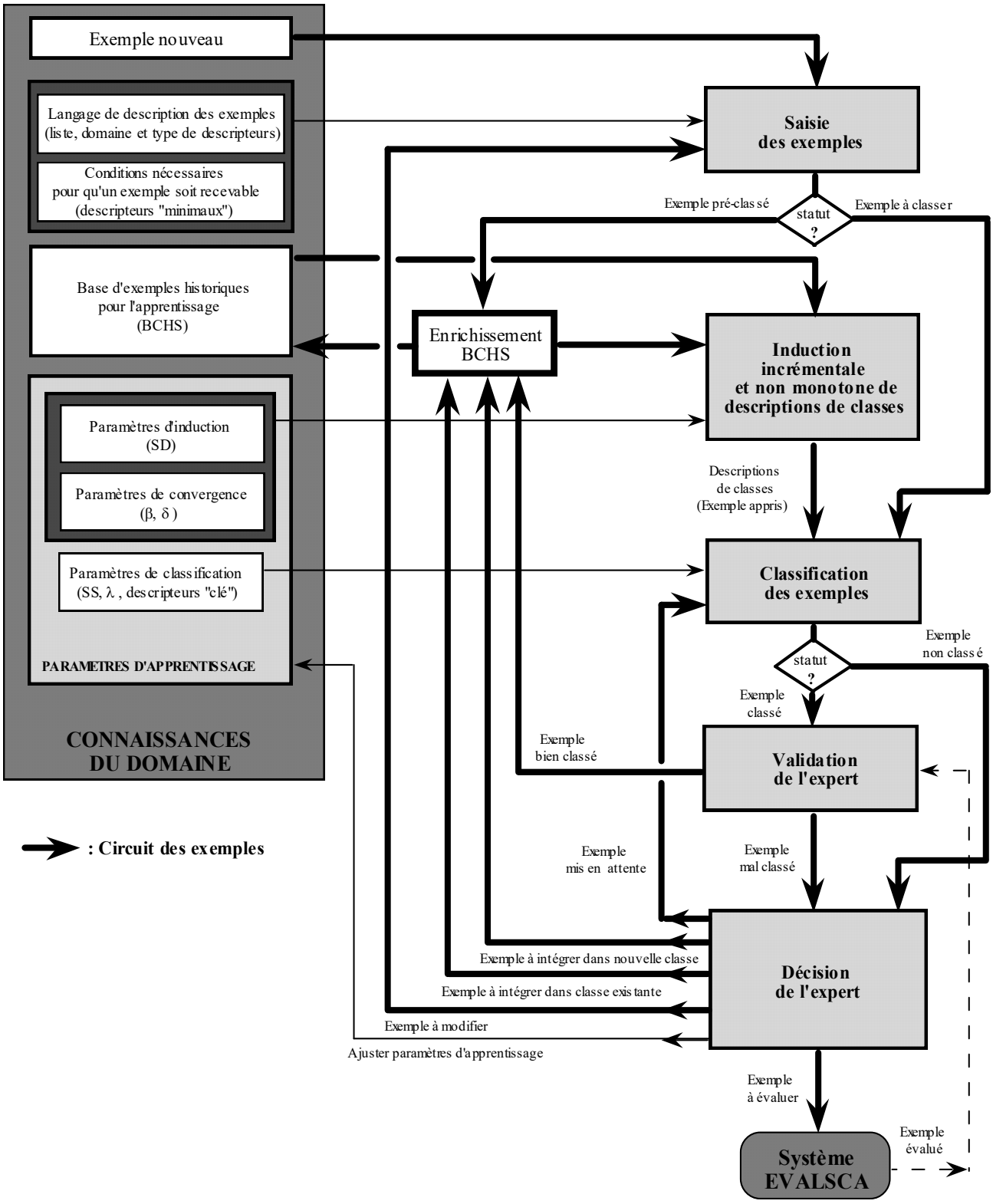


Figure 4.6. : Architecture générale de CLASCA

Nous venons de présenter la conception du système CLASCA dont les résultats (classe d'appartenance du nouvel exemple de scénario, PR historiques impliquées dans une classe, exemples de scénarios historiques appris) sont transmis pour exploitation au deuxième niveau de traitement assuré par le système EVALSCA détaillé ci-après.

3. CONCEPTION DU SYSTEME EVALSCA

Ce paragraphe décrit dans un premier temps le principe général du système EVALSCA développé pour l'aide à l'évaluation des scénarios d'accidents. Dans un deuxième temps, il détaille les modules impliqués dans EVALSCA.

3.1. Principe général du système EVALSCA

EVALSCA /HADJ-MABROUK 92e/ est un système expert à apprentissage pour l'aide à l'évaluation des scénarios d'accidents. Le but de cet outil est de confronter les Pannes Résumées proposées dans un scénario constructeur à la liste des PR historiques suggérées par le système CLASCA. Cette démarche d'évaluation favorise la génération de nouveaux scénarios par la déduction d'une ou plusieurs PR susceptibles de mettre en cause la sécurité du système de transport. En ce sens, elle permet d'aider les experts de certification à juger de l'exhaustivité du dossier de sécurité du constructeur.

EVALSCA s'organise autour des quatre modules suivants présentés dans la figure 4.7 :

- une base d'exemples d'apprentissage issue du système CLASCA,
- un système d'apprentissage de règles CHARADE /GANASCIA 87/ qui exploite cette base d'exemples pour produire des fonctions de reconnaissance de Pannes Résumées (PR),
- un générateur de systèmes experts "Intelligence Service II" (IS2),
- une base de connaissances d'évaluation des scénarios d'accidents exploitée par le moteur d'inférence d'IS2 en vue de déduire les PR à considérer dans le scénario proposé par le constructeur.

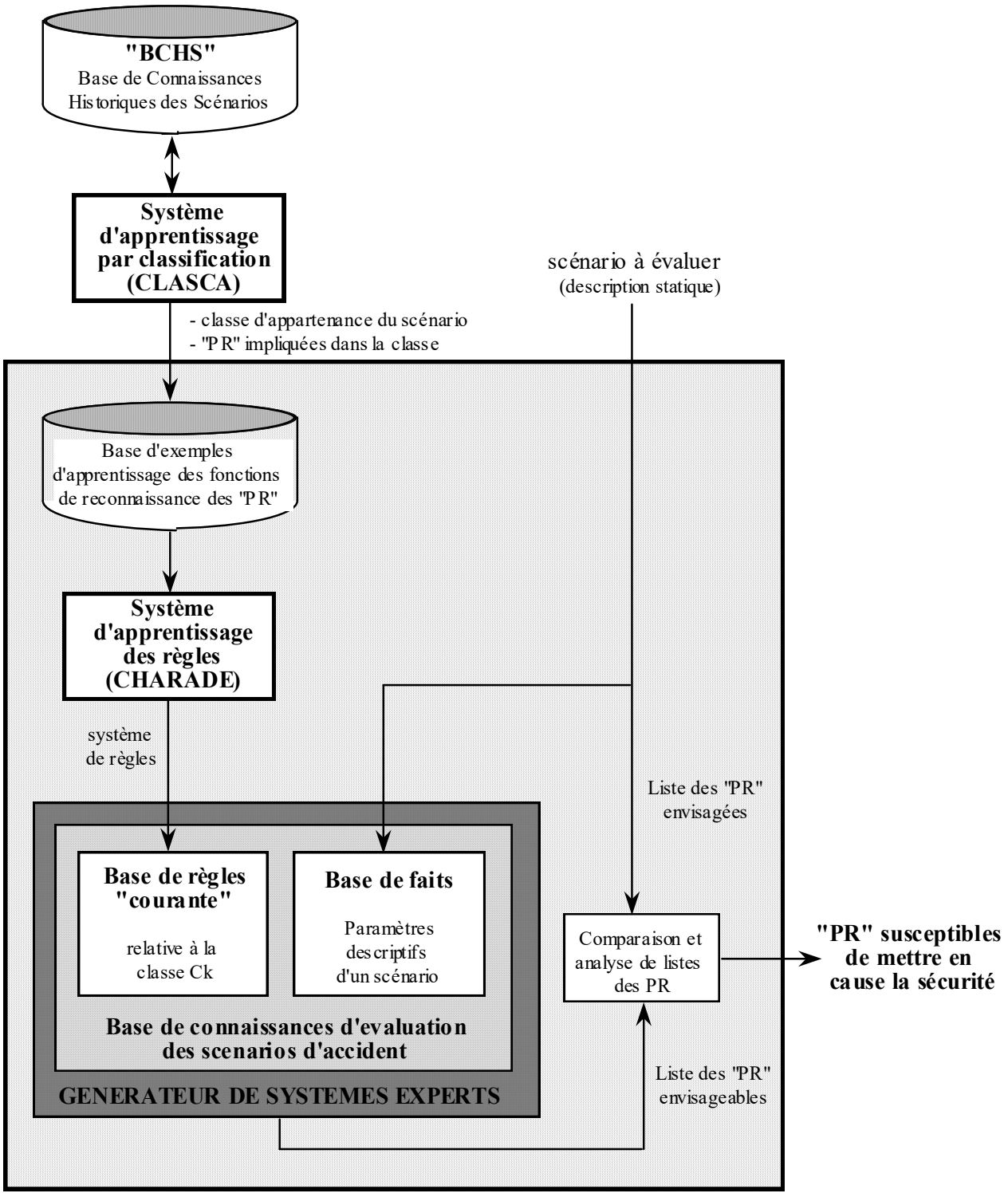


Figure 4.7 : Architecture fonctionnelle du système EVALSCA

3.2. Différents modules constituant EVALSCA

EVALSCA est constitué des quatre modules suivants, détaillés dans la suite :

. une base d'exemples d'apprentissage,
. un système d'apprentissage de règles,
. un générateur de systèmes experts,
. une base de connaissances d'évaluation.

3.2.1. Base d'exemples d'apprentissage

La base d'exemples d'apprentissage relative à une classe est obtenue en regroupant les scénarios de la BCHS dont la description fait intervenir les PR de cette classe. Elle est élaborée à partir des résultats de classification fournis par CLASCA et exploitée par un système d'apprentissage de règles qui construit une base de connaissances pour l'évaluation des scénarios d'accidents. Cette base est organisée dans un format compatible avec celui requis par le mécanisme d'apprentissage CHARADE. Elle est rafraîchie à chaque mise à jour des classes proposées par le système CLASCA. En résumé, CLASCA permet de repérer la classe d'un nouveau scénario qui, après évaluation au sein de cette classe par EVALSCA et l'expert, peut devenir un exemple d'apprentissage qui permettra d'enrichir la BCHS.

3.2.2 Système d'apprentissage de règles

A partir de la base d'exemples constituée précédemment, la phase d'apprentissage des règles s'attache à engendrer un système de règles traduisant les fonctions de reconnaissance des PR /HADJ-MABROUK 91a/. Cette étape a pour but d'induire une fonction de reconnaissance de chacune des PR impliquées dans une classe. La fonction de reconnaissance d'une PR est une règle de production qui établit une relation entre un ensemble de faits (paramètres descriptifs d'un scénario ou descripteurs) et le fait PR. Il s'agit d'une relation de dépendance logique que l'on peut mettre sous la forme :

> **SI** les faits (ou descripteurs) : PC, R, FRR, ZG, AI, FI sont vérifiés
> **ALORS** ils ont pour conséquence le seul fait (ou descripteur) : PR

Pour chaque classe de scénarios on peut induire une base de règles d'évaluation. Toute règle engendrée doit contenir le descripteur ou fait PR dans sa conclusion. Le recours à une méthode d'apprentissage permettant d'engendrer, à partir d'un ensemble d'exemples (ou scénarios) historiques, des règles de production s'avère indispensable. Dans le chapitre 3, la spécification des propriétés requises par le système d'apprentissage ainsi que l'analyse de l'existant ont orienté le choix vers le mécanisme CHARADE. L'induction automatique d'un système de règles et non pas des règles isolées ainsi que la possibilité de structurer les règles pour élaborer des fonctions de reconnaissance des PR confèrent à CHARADE un intérêt indéniable. En effet, pour notre application, le but est d'engendrer des règles contenant dans leur conclusion les PR et dans leur prémisse les autres paramètres descriptifs du scénario.

Le système CHARADE

CHARADE /GANASCIA 87, 88, 91/ est un système d'apprentissage destiné à construire automatiquement des bases de connaissances à partir d'exemples. Il permet d'engendrer un système de règles doté de propriétés particulières et repose sur l'utilisation d'une structure d'hypercube dont les propriétés sont exploitées pour permettre une exploration "intelligente" de l'espace de description et une représentation sous forme d'heuristiques de généralisation des propriétés du SBC à construire. Pour élaborer automatiquement des bases de connaissances, CHARADE requiert :

- un langage de description contenant un ensemble d'attributs et d'axiomes exprimant la sémantique du domaine
- un ensemble d'exemples décrits dans ce langage de description
- la description des fonctionnalités du SBC dont on veut élaborer la base de connaissances.

Les exemples de CHARADE sont définis par des conjonctions de triplets (<attribut> <sélecteur> <valeur>). Les axiomes sont mis à profit pour compléter les descriptions éventuellement incomplètes des exemples et sont exprimés par des règles de production. La génération des règles dans CHARADE est basée sur la recherche et la découverte des régularités empiriques présentes sur l'ensemble d'apprentissage. Une régularité correspond à une corrélation observée entre descripteurs de la base d'exemples d'apprentissage : si tous les exemples de l'ensemble d'apprentissage qui possèdent le descripteur d1 possèdent aussi le descripteur d2, on peut induire que d1-->d2 sur l'ensemble d'apprentissage. Ce processus d'obtention de règles est fondé sur l'emploi de deux treillis booléens : un treillis des descripteurs et un treillis des exemples. Plus précisément, J.G. GANASCIA définit deux fonctions C et D schématisées dans la figure 4.8. La première fonction C va du treillis des descripteurs vers le treillis des exemples : elle associe à chaque description (di) constituée d'une conjonction de descripteurs, le sous-ensemble (E) des exemples de l'ensemble d'apprentissage couverts par cette description. L'autre fonction D fait correspondre à chaque sous-ensemble (E) de l'ensemble d'apprentissage, la description (dj) formée de la conjonction de descripteurs communs à tous les exemples de E. La fonction composée COD qui associe à la description conjonctive (di) la description conjonctive (dj), engendre la règle di-->dj.

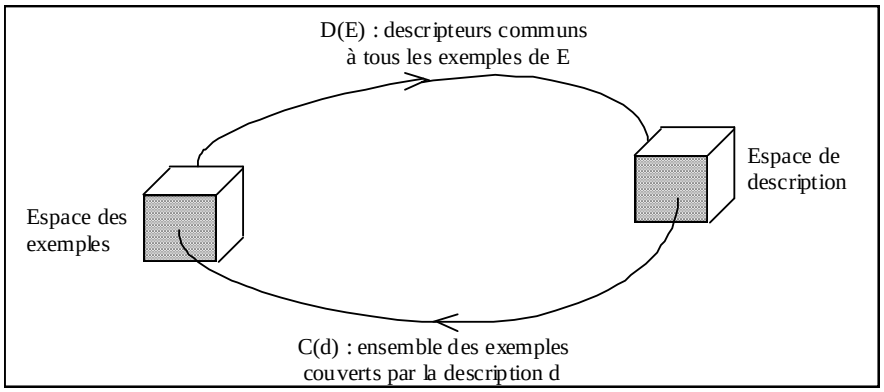


Figure 4.8 : Principe d'obtention d'un système de règles

Pour illustrer ce principe de génération de règles, supposons que l'on dispose d'un ensemble d'apprentissage constitué de trois exemples E1, E2 et E3 :

E1 = d1 & d2 & d3 & d4

E2 = d1 & d2 & d4 & d5

E3 = d1 & d2 & d3 & d4 & d6

CHARADE peut alors détecter une régularité empirique entre la conjonction de descripteurs (d1 & d2) et le descripteur d4. En effet, tous les exemples décrits par d1 & d2 possèdent aussi d4 dans leur description. Cette régularité s'obtient à l'aide des deux fonctions C et D : CoD (d1 & d2) = C ({E1, E2, E3}) = d1 & d2 & d4. Finalement, la règle : d1 & d2-->d4 est obtenue.

En plus de ces règles logiques, CHARADE permet également d'engendrer des règles approximatives qui traduisent les corrélations statistiques observées entre les descripteurs. En effet, la technique qui vient d'être présentée détecte les règles certaines. Cependant, lorsqu'on construit une base de connaissances, il faut aussi détecter et traduire l'incertitude. Cette dernière peut être due soit à la mauvaise qualité de l'ensemble d'apprentissage, soit aux règles qui, par nature, sont incertaines. Pour traduire cette incertitude, GANASCIA utilise les probabilités conditionnelles des descripteurs, calculées à l'aide de leur fréquence d'apparition respective dans les exemples de l'ensemble d'apprentissage. Les probabilités obtenues, comprises par définition entre 0 et 1, sont par la suite ramenées à l'échelle [-1, +1] des coefficients de vraisemblance. Ainsi, si D est une conjonction de descripteurs et d un descripteur, Pr (d/D) est la probabilité que le descripteur d soit présent dans la description d'un exemple lorsque la conjonction D l'est. Finalement, la corrélation statistique entre D et d s'exprime à l'aide de la règle de production suivante : D-->d(Cv) où Cv = 2 Pr(d/D) - 1. Cv désigne le coefficient de vraisemblance. En reprenant l'exemple élémentaire précédent, il est possible d'induire la règle approximative (ou incertaine) suivante : d1 & d2-->d3 (Cv = 2 x 2/3 - 1 = 0,33)

L'exploration exhaustive de la totalité de l'espace de description en vue de détecter toutes les régularités possibles est presque impossible pour deux raisons : d'une part la procédure d'exploration serait exponentielle (2^n conjonctions de descripteurs, avec n le nombre de descripteurs), d'autre part le nombre de règles engendrées serait important et certaines d'entre elles pourraient être redondantes ou inutiles. Pour pallier ces difficultés et adapter le format des règles aux caractéristiques et propriétés recherchées par l'utilisateur d'un SBC, GANASCIA définit des contraintes (paramètres d'induction) dont les principales sont les suivantes :

- Le facteur de bruit. Celui-ci permet de fixer le nombre minimal d'exemples nécessaire pour détecter une régularité. Par exemple, si ce facteur vaut 2 alors seules les régularités présentes dans 2 exemples au moins produisent effectivement une règle. L'avantage de cette contrainte est de limiter l'influence des données bruitées sur les règles, ce qui permet d'engendrer des règles pertinentes.
- La limitation du nombre de descripteurs par prémisse des règles induites. Par exemple, si cette contrainte est fixée à 3, alors toutes les règles générées possèdent au plus 3 descripteurs dans leur prémisse.
- La condition terminale. Il s'agit de définir un ou plusieurs attributs qui ne seront présents dans les prémisses d'aucune règle.

- La structuration du système de règles. Cette contrainte oblige les règles produites à aller d'un ensemble de descripteurs vers un autre ensemble descripteurs. Par exemple, pour un SBC destiné au diagnostic, les règles souhaitées sont orientées des symptômes vers les causes et des causes vers les remèdes.

Nous venons de présenter le principe de fonctionnement du système CHARADE choisi pour la production de systèmes de règles structurés. Pour exploiter ces règles, il est nécessaire de faire appel à un générateur de systèmes experts.

3.2.3. Générateur de systèmes experts

Dans le cadre de l'évaluation des scénarios d'accidents, le système expert exploite les règles produites par CHARADE pour déduire les Pannes Résumées envisageables qui doivent être prises en compte par le constructeur. Les trois principales fonctions que doit assurer ce module sont les suivantes :

- Représentation des connaissances sous forme de règles de production en logique propositionnelle et réalisation d'inférences en chaînage avant, à partir de ces connaissances,
- Structuration de la base de connaissances en plusieurs sous-bases, chacune spécifique à un problème particulier,
- Interfaçage avec le mécanisme d'apprentissage CHARADE.

Intelligence Service II (IS2) est un générateur de systèmes experts permettant d'assurer ces différentes fonctions. Il se compose essentiellement de deux modules :

- Le module "consultant" qui permet à l'utilisateur d'accéder à la connaissance de l'expert en posant des questions, en donnant des informations au système et en lui demandant de faire part de ses déductions.
- Le module "expert" qui permet de formaliser la connaissance de l'expert afin de la rendre exploitable par le module consultant et donc par le moteur d'inférences.

En plus de sa facilité d'utilisation, IS2 dispose d'une interface multi-fenêtres qui assure un bon niveau de convivialité. Son moteur d'inférences d'ordre 0+ fonctionne en logique des propositions et gère des propositions numériques (valeur d'attribut réelle), symboliques (la valeur d'attribut appartient à un domaine de définition discret établi au préalable) ou booléennes (valeur d'attribut vraie ou fausse). IS2 raisonne en logique monotone : dès qu'une proposition ou un fait a reçu une valeur, celle-ci ne peut plus être modifiée, lors de l'application d'une règle. L'existence dans IS2 de procédures externes le rend facilement interfaçable avec CHARADE. La syntaxe des règles utilisée par IS2 est la suivante : SI (prémisse 1) *...* (prémisse n) ALORS (conclusion 1) ET ...ET (conclusion p). Le signe "*" désigne l'opérateur logique "ET" ou l'opérateur logique "OU".

Les règles engendrées par CHARADE sont transmises au système expert (IS2) pour construire une base de connaissances d'évaluation des scénarios.

3.2.4. Base de connaissances d'évaluation des scénarios d'accidents

La base de connaissances du module d'évaluation des scénarios d'accidents contient :

- La base de règles "courante" qui est chargée avec le système de règles induit par CHARADE pour chacune des classes de scénarios proposée par CLASCA,
- La base de faits qui contient les paramètres descriptifs du scénario constructeur à évaluer et qui s'enrichit, au fil de l'inférence, des faits ou descripteurs déduits.

Cette base de connaissances d'évaluation des scénarios (base de faits et base de règles), exploitée en chaînage avant par le moteur d'inférence d'IS2, permet d'engendrer les Pannes Résumées (PR) possibles que doit contenir le scénario constructeur. Par comparaison avec les PR réellement envisagées par le constructeur, on déduit des PR pouvant mettre en cause la sécurité du système de transport. Cette suggestion est à même d'aider à la génération de situations d'insécurité non prévues par le constructeur.

4. IMPLEMENTATION DU SYSTEME ACASYA

Les paragraphes précédents ont détaillé les différentes phases de conception du système ACASYA L'étude de faisabilité de ce système, appliquée au domaine de la certification des systèmes de transport terrestres automatisés, a débouché sur la réalisation d'une maquette /MEJRI 91/, /DERENTY 91/, /GABER 92/, /HADJ MABROUK 92f/ dont les outils et langages ainsi que l'architecture fonctionnelle sont abordés ci-dessous.

4.1. Outils et langages de développement

Dans le cadre de l'étude de faisabilité du système ACASYA, notre premier souci a été de valider la démarche proposée pour l'aide à la certification plutôt que de privilégier une étude approfondie et coûteuse des outils et langages de développement nécessaires au développement d'un prototype. De plus, contrairement à un logiciel commercialisé, nous avons tenu à offrir à l'utilisateur et/ou à l'expert de certification une certaine "transparence" dans l'articulation des différents niveaux de traitement de la méthodologie afin de favoriser la coopération Homme-Machine. La maquette a donc été implémentée comme suit :

- CLASCA est programmé en Turbo Pascal sous Windows et fonctionne sur compatible IBM PC sous MS DOS.
- Une version prototype de CHARADE a été aimablement mise à notre disposition par J-G. GANASCIA. CHARADE est écrit en Le-lisp 15.2 et fonctionne sur un Macintosh Classic pour des raisons de mémoire et de logiciel disponible. Des versions commerciales écrites en C++ sur PC ou station de travail devraient bientôt voir le jour.
- L'interface Homme-Machine qui utilise le concept de menu déroulant pour aborder toutes les étapes de la méthodologie d'aide à la certification est développée en Turbo Pascal sous Windows.
- Le générateur de systèmes experts Intelligence Service II (IS2) fonctionne sur compatible IBM PC.

Après cette description des principaux outils utilisés pour le développement du système ACASYA, nous présentons son architecture fonctionnelle.

4.2. Architecture de la maquette du système ACASYA

Les connaissances extraites et formalisées au cours de la phase d'acquisition de connaissances ont été implémentées grâce aux outils matériels et logiciels choisis précédemment. L'architecture du système ACASYA finalement conçu est présentée en figure 4.9. La maquette d'ACASYA est constituée des trois principaux modules suivants dont le contenu est détaillé ci-après :

. Une interface Homme-Machine
. Un système pour la classification des scénarios d'accidents : CLASCA
. Un système pour l'évaluation des scénarios d'accidents : EVALSCA

4.2.1. L'interface Homme-Machine

L'interface Homme-Machine permet d'assurer le dialogue avec les utilisateurs et/ou l'expert de certification. Cette interface assure les deux grandes fonctions suivantes :

. La première concerne la saisie et la mise à jour des connaissances d'analyse de sécurité des systèmes de transport automatisés. Les connaissances impliquées sont :

- un langage de description des scénarios d'accidents, composé d'un ensemble de descripteurs (ou paramètres descriptifs d'un scénario) auquels sont associés un type et un domaine.
- la "BCHS", constituée à ce jour d'une quarantaine de scénarios historiques relatifs au risque de collision. Ces scénarios sont formalisés selon une description statique puis regroupés en classes par l'expert de certification.
- des scénarios d'accidents historiques et pré-classés par l'expert pour enrichir la BCHS ainsi que des nouveaux scénarios proposés par le constructeur dont les certifieurs cherchent à évaluer la consistance.
- des paramètres d'apprentissage (paramètres d'induction, de classification et de convergence) et des contraintes de recevabilité d'un scénario.

. La seconde fonction permet la consultation des différentes connaissances produites par le système ACASYA et notamment :

- les résultats de classification engendrés par CLASCA tels que la classe d'appartenance d'un nouveau scénario, la description d'une classe de scénarios, la liste des Pannes Résumées (PR) impliquées dans une classe, les paramètres et contraintes d'apprentissage.
- les résultats d'évaluation obtenus par le système EVALSCA. L'utilisateur peut, par exemple, consulter les règles générées par le système d'apprentissage CHARADE, le contenu de la base de connaissances d'évaluation ainsi que les résultats de déduction du SBC (par exemple, les PR non considérées par le constructeur).
- les connaissances historiques d'analyse de sécurité telles que la liste des classes, des scénarios d'accidents appris, des Pannes Résumées (PR), des Solutions Adoptées (SA), des scénarios en attente.

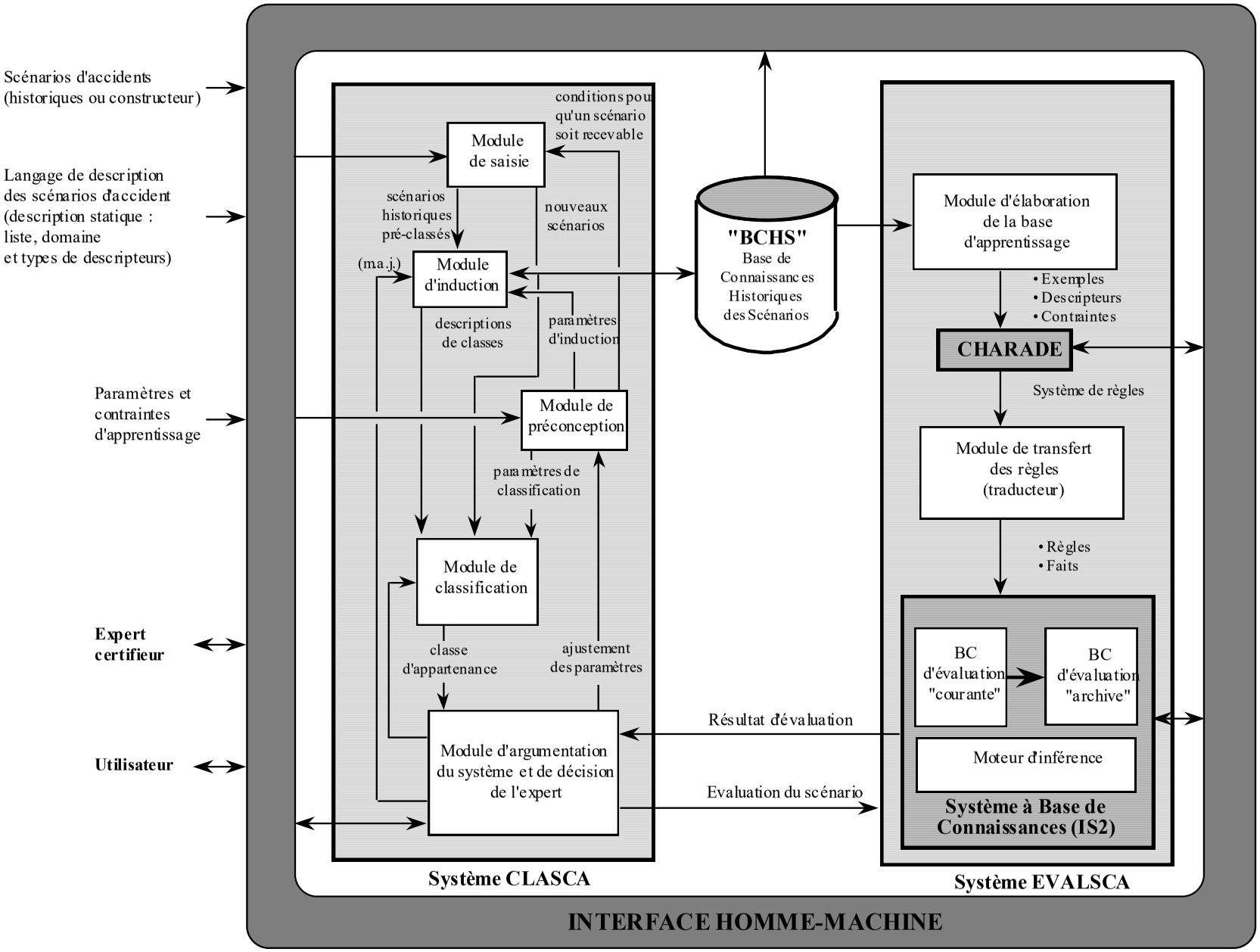


Figure 4.9. : Architecture de la maquette du système ACASYA

4.2.2. Le système CLASCA

CLASCA dont le but est la recherche de la classe d'appartenance d'un nouveau scénario d'accident, est constitué des cinq modules suivants :

- Un module de saisie des scénarios d'accidents.
- Un module de préconception, au sens où l'a défini GANASCIA /90/, utilisé pour fixer les différentes valeurs des paramètres et contraintes d'apprentissage requis par le système. Ces paramètres agissent principalement sur la pertinence et la qualité des connaissances de classification apprises ainsi que sur la rapidité de convergence du système.
- Un module d'induction pour l'apprentissage des descriptions conjonctives de classes de scénarios.
- Un module de classification dont l'objectif est la déduction de la classe d'appartenance d'un nouveau scénario à partir des descriptions de classes induites précédemment et en référence à un taux d'adéquation.
- Un module de dialogue pour l'argumentation du système et la décision de l'expert. Dans l'argumentation ou la justification, le système conserve une trace de la phase de déduction pour construire son explication. Suite à cette phase de justification des décisions de classification, l'expert de certification décide, soit d'accepter la classification proposée et dans ce cas le scénario sera appris par CLASCA, soit de rejeter la classification. Dans le deuxième cas il appartient à l'expert de décider de la suite à donner. Il peut par exemple décider d'évaluer le scénario à l'aide du système EVALSCA que nous détaillons ci-après.

4.2.3. Le système EVALSCA

EVALSCA évalue le contenu et la consistance du scénario en se basant essentiellement sur l'analyse des Pannes Résumées (PR). Il permet de révéler à l'expert une ou plusieurs PR non considérées par le scénario constructeur à évaluer. EVALSCA s'organise autour des quatre modules suivants :

- Un module d'élaboration de la base d'apprentissage qui permet la création de la base d'exemples pour l'apprentissage des fonctions de reconnaissance des PR. Cette base regroupe les scénarios de la "BCHS" dont les PR appartiennent à la liste des PR relatives à la classe courante engendrée par CLASCA.
- Le système CHARADE d'apprentissage de règles qui exploite la base d'apprentissage précédemment préparée pour engendrer un système de règles traduisant les régularités présentes sur les exemples d'apprentissage et permet de reconnaître l'existence des PR.
- Un module de traduction/transfert des règles. Les règles produites par CHARADE sont écrites suivant une syntaxe particulière. Elles sont traduites pour être compatibles avec le générateur de systèmes experts IS2 qui ne traite pas les descripteurs multivalués mais uniquement les descripteurs énumérés

ou booléens. Elles sont ensuite transférées automatiquement vers la base de connaissances du SBC.

- Un système à base de connaissances (SBC), constitué d'un moteur d'inférence et d'une base de connaissances (BC) d'évaluation. La BC est scindée en deux parties : une BC d'évaluation "courante" qui contient les règles induites par CHARADE relatives à une classe proposée par CLASCA à l'instant t et une BC d'évaluation "archive", composée de la liste des bases de règles historiques. Après évaluation, une BC "courante" devient une BC "archive". La BC "courante" est exploitée par le moteur d'inférence pour déduire les PR à considérer dans le scénario constructeur. Suite à cette phase d'évaluation du contenu du scénario, l'utilisateur peut revenir au module d'argumentation et de décision (système CLASCA) pour confirmer la consistance du scénario qui sera alors appris par CLASCA pour devenir historique. Il peut également opter pour d'autres décisions telles que la mise en attente du scénario à évaluer ou le choix d'une autre classe pour évaluer à nouveau le scénario.

L'ensemble des modules précédemment évoqués est aujourd'hui opérationnel dans le cadre de la maquette du système ACASYA.

CONCLUSION

A partir de l'application des moyens d'acquisition de connaissances pour l'élaboration d'une base de connaissances de certification présentée dans le chapitre 2 et en nous appuyant sur les récents travaux de recherche menés en matière d'apprentissage automatique présentés dans le chapitre 3, nous avons acquis un savoir faire qui a abouti à l'élaboration d'une méthodologie d'aide à la certification des systèmes de transport terrestres automatisés. L'originalité principale de cette méthodologie réside dans l'utilisation conjointe des techniques d'acquisition de connaissances et d'apprentissage automatique pour renforcer les méthodes usuelles d'analyse de sécurité. Au contraire d'une démarche de diagnostic, cette méthodologie a pour finalité la prévention des pannes. Elle s'organise en trois niveaux complémentaires permettant d'aborder progressivement la résolution du problème de certification (classification, évaluation et génération des scénarios d'accidents) tout en conservant une forte interactivité avec l'expert du domaine.

Ce chapitre a présenté l'organisation fonctionnelle du système ACASYA, système d'Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés qui constitue l'outil support de la méthodologie, opérationnel à ce jour pour les deux premiers niveaux de traitement : CLASCA et EVALSCA.

CLASCA est un système d'apprentissage à partir d'exemples, inductif, incrémental et dédié à la classification des scénarios d'accidents. L'apprentissage dans CLASCA est d'une part non monotone pour prendre en compte les données bruitées et incomplètes relatives aux scénarios d'accidents et il est d'autre part interactif (supervisé) pour contrôler les connaissances produites par le système et aider l'expert à mieux formuler son expertise. EVALSCA est un système d'évaluation des scénarios d'accidents basé sur la production de règles de reconnaissance des pannes par le système d'apprentissage CHARADE. Ces règles sont exploitées par un système à base de connaissances pour déduire les pannes susceptibles de mettre en défaut la sécurité d'un nouveau système de transport. Les deux modules CLASCA et EVALSCA sont complémentaires et interactifs pour aider à la certification. Ils ont nécessité la mise en oeuvre d'interfaces conviviales pour assurer la communication Système/Système et Expert/Système.

L'évaluation de la maquette de faisabilité du système ACASYA ainsi que les perspectives envisagées pour l'améliorer sont développées dans le chapitre suivant.

Chapitre 5 :

Evaluation du système ACASYA et perspectives

INTRODUCTION

Après avoir présenté dans le chapitre précédent les différentes phases de conception et de réalisation du système ACASYA, nous dressons dans ce chapitre un premier bilan des performances d'ACASYA. Quelques perspectives d'évolution de ce système et des travaux qui s'y rattachent sont envisagées. Ce chapitre est scindé en trois parties. La première partie est consacrée à la présentation d'un exemple de fonctionnement du système illustrant les différentes étapes de la démarche d'aide à la certification. Une première évaluation globale de la maquette de faisabilité du système ACASYA, à partir de l'ensemble des scénarios formalisés à ce jour, est complétée dans la deuxième partie par une évaluation interne du module CLASCA basée sur un lot de scénarios tests. Chacune de ces évaluations introduit une discussion sur les intérêts, limites et extensions possibles. L'évaluation interne du module EVALSCA n'est pas nécessaire puisqu'il met à contribution des logiciels existants. Son intérêt par rapport à la méthodologie est pris en compte dans l'évaluation globale d'ACASYA. Dans la troisième partie, nous proposons une réflexion sur la complémentarité de l'acquisition de connaissances et de l'apprentissage automatique mise à contribution dans notre démarche de conception d'un outil d'aide à la certification.

1. EVALUATION DE LA MAQUETTE DE FAISABILITE DU SYSTEME ACASYA

A ce jour et suite à une trentaine de sessions d'extraction de connaissances, nous avons constitué une base de connaissances historiques d'analyse de sécurité des systèmes de transport automatisés déjà certifiés et/ou homologués. Ces connaissances ont été formalisées selon le modèle "générique" de représentation (description statique et dynamique) que nous avons élaboré. Dans l'état actuel des travaux, les deux modules CLASCA et EVALSCA sont implémentés et constituent la maquette sur laquelle porte l'étude de faisabilité du système ACASYA. Ce paragraphe présente l'évaluation d'ACASYA en regard de son objectif d'aide à la certification ainsi que ses perspectives d'évolution.

1.1. Exemple d'application du système ACASYA

L'utilisation d'ACASYA suppose au préalable l'extraction et la formalisation de scénarios. Pour évaluer ACASYA nous avons testé ses performances sur les quarante trois scénarios d'accidents relatifs au risque de collision archivés à ce jour. La mise en oeuvre d'ACASYA implique le déroulement des neuf étapes suivantes /HADJ-MABROUK 92f/ :

1. Préconception
2. Saisie des scénarios d'accident (nouveaux scénarios à évaluer et scénarios historiques à archiver)
3. Apprentissage des descriptions conjonctives de classes de scénarios
4. Classification d'un nouveau scénario saisi
5. Constitution de la base d'apprentissage centrée sur les PR impliquées dans la description de la classe d'appartenance du nouveau scénario
6. Apprentissage des fonctions de reconnaissance des PR
7. Déduction des PR à considérer dans le scénario à évaluer
8. Evaluation de la complétude du scénario
9. Mise à jour des connaissances produites antérieurement.

Les quatre premières étapes ainsi que la neuvième sont du ressort du système CLASCA alors que les quatre autres étapes concernent le système EVALSCA. Nous proposons dans la suite de présenter le système ACASYA au travers du déroulement d'une session de consultation. Au préalable, une présentation du menu principal est nécessaire (figure 5.1.).

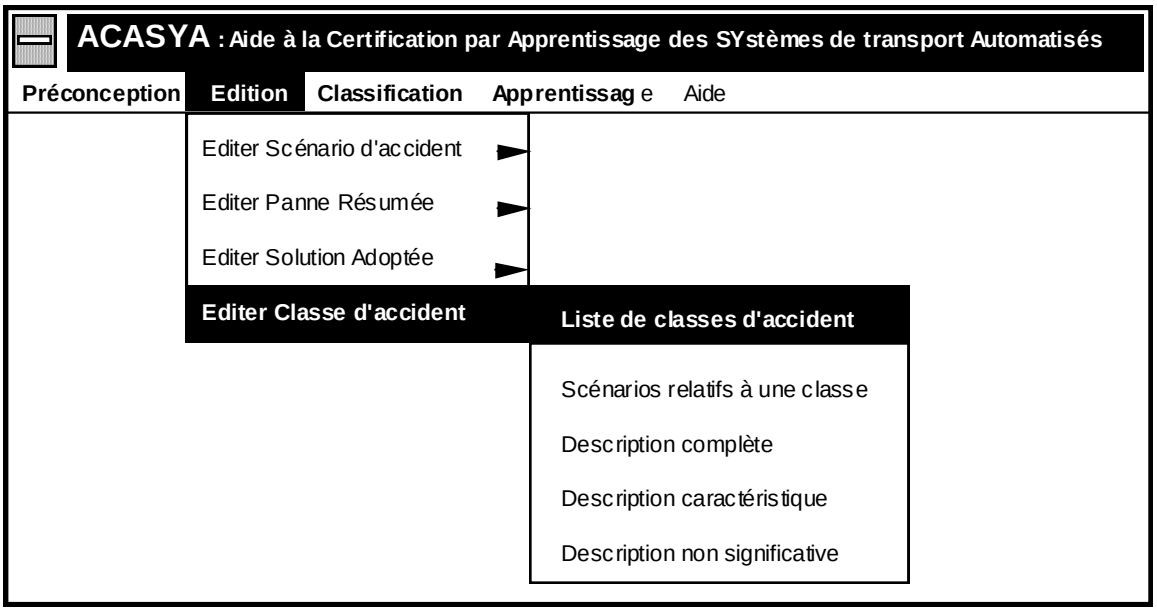
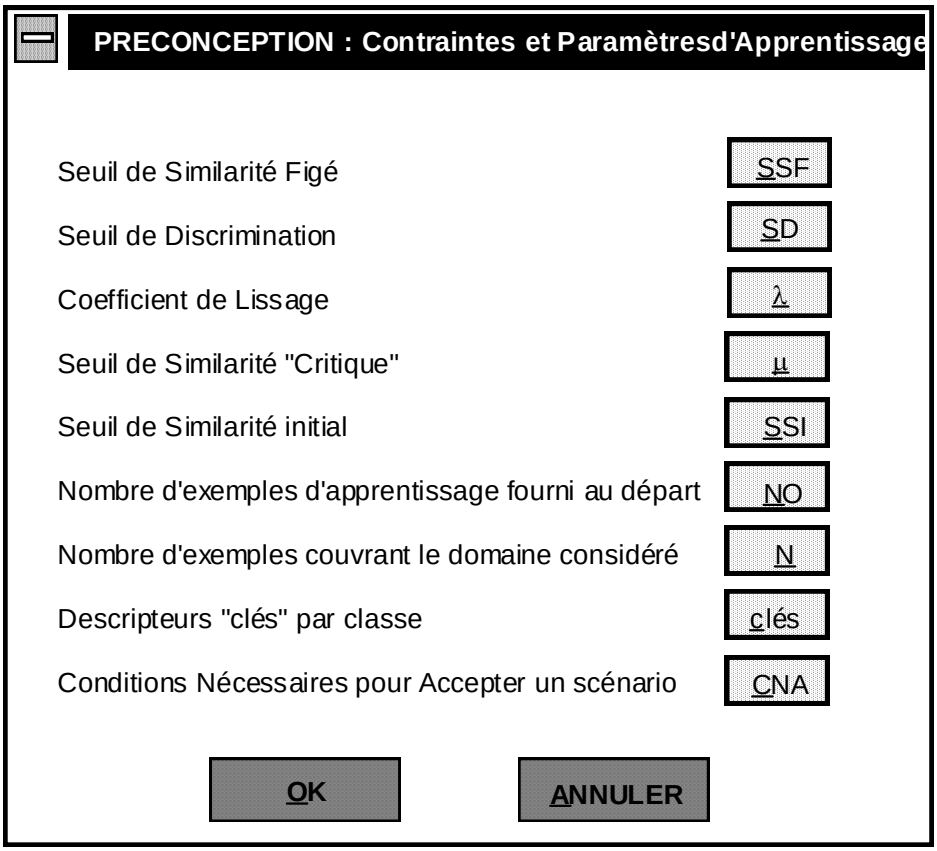


Figure 5.1. : Menu principal d'ACASYA

Dans ce menu, les commandes de "préconception", d'"édition" et d'"apprentissage" correspondent aux trois premières étapes précédemment citées. La commande de "classification" englobe les six étapes suivantes qui s'enchainent successivement. Enfin, une commande d'"aide" est prévue pour guider l'utilisateur au cours de l'exploitation du système. Les paragraphes suivants détaillent le déroulement des neuf étapes impliquées dans ACASYA.

1.1.1. Préconception

L'utilisation d'ACASYA nécessite en premier lieu une étape de préconception (figure 5.2.) au cours de laquelle l'utilisateur définit les valeurs des paramètres et contraintes utiles à l'apprentissage. Dans cette étape, l'utilisateur choisit de fonctionner avec un seuil de similarité figé (SSF) ou avec un seuil de similarité évolutif (SSE). Dans le premier cas, il doit fixer la valeur du seuil de similarité alors que dans le second cas, le seuil s'ajuste automatiquement en fonction du cardinal de la classe et agit sur la rapidité de convergence du système. Le fonctionnement avec un SSE rend nécessaire la configuration des paramètres de convergence par la définition du seuil de similarité initial (SSI), du seuil de similarité critique (μ), du nombre d'exemples d'apprentissage fourni au départ au système (N0) et du nombre d'exemples couvrant le domaine de certification (N). Dans un deuxième temps, l'utilisateur doit indiquer les valeurs relatives au seuil de discrimination des descripteurs (SD) et au coefficient de lissage (λ) des taux d'adéquation "statistique" et "sémantique". Enfin, il doit saisir la liste des descripteurs "clés" par classe de scénarios ainsi que les conditions nécessaires d'acceptabilité d'un scénario (CNA) définies par la liste des "descripteurs minimaux"



Ecran 2

Figure 5.2. : Menu de préconception

A titre d'exemple, pour un fonctionnement avec un seuil de similarité évolutif (SSE), voici les valeurs des différents paramètres :

. Seuil de similarité évolutif, fonction de SSI, μ, N0 et N

- SSI = 0,5 : en début d'apprentissage, on tolère une similarité minimale de 50% pour intégrer un scénario dans une classe
- μ = 0,8 : lorsque les descriptions des classes tendent à être stables, on exige une similarité minimale de 80%
- N0 = 43 : la base d'exemples d'apprentissage contient 43 scénarios d'accidents
- N = 100 : l'expert a estimé que le nombre de scénarios relatifs au risque de collision est d'environs une centaine

. Seuil de discrimination : SD = 0,5

Les descripteurs dont la fréquence d'apparition est supérieure ou égale à 0,5 sont considérés comme caractéristiques pour décrire une classe de scénarios.

. Coefficient de lissage : λ = 0,5

Le taux d'adéquation entre un nouveau scénario et les classes préexistantes prend en compte de façon équivalente l'aspect sémantique et l'aspect statistique de la classification (chapitre 4, § 2.4.3.)

. Conditions nécessaires pour accepter un scénario (CNA) : "Risque", "Fonction en Rapport avec le Risque", "Zone Géographique" et "Pannes Résumées".

. Descripteurs "clés" par classe : la liste des descripteurs "clés" par classe de scénarios d'accidents est présentée dans le tableau ci-après (figure 5.3.).

CLASSES DE SCENARIOS d'ACCIDENT	DESCRIPTEURS "CLES" : (Attribut / Valeur)
C1 : Commutation de redondance	(FRR = commutation de redondance) (AI = PA avec redondance)
C2 : Séquence d'initialisation	(FRR = localisation des trains) (FRR = initialisation) (FI = consignes)
C3 : Localisation des trains	(FRR = localisation des trains) (ZG =ligne)
C4 : Gestion du freinage d'urgence	(FRR = gestion de l'arrêt des trains) (FRR = sécurité Quai-Voie) (FRR = traction / freinage) (ZG = limite de tronçon)
C5 : Accostage	(FRR = consigne de vitesse) (FRR = accostage) (ZG = ligne) (FI = consignes)
C6 : Gestion de sens de marche	(FRR = gestion de sens de marche) (FRR = consigne de vitesse) (ZG = limite de tronçon)
C7 : Contrôle d'Entrée/sortie	(FRR = contrôle d'entrée / sortie) (FRR = localisation des trains) (FRR = suivi des trains) (ZG = limite de tronçon)
C8 : Suivi de l'ordre des trains	(FRR = localisation des trains) (FRR = suivi des trains) (ZG = limite de tronçon)
C9 : Conduite manuelle	(FRR = conduite manuelle) (FRR = consigne de vitesse) (FI = consignes) (FI = communication)

Figure 5.3. : Descripteurs "clés" par classe de scénarios d'accidents
(Expertise acquise auprès de l'expert)

A ce jour, les deux paramètres "CNA" et "clés" sont figés et spécifiques au problème de la collision, en revanche les autres paramètres sont modifiables par l'utilisateur.

Après avoir saisi l'ensemble des paramètres et contraintes d'apprentissage, l'utilisateur doit procéder à l'acquisition des scénarios d'accidents.

1.1.2. Acquisition des scénarios d'accidents

Cette deuxième étape correspond à la commande "édition" du menu de la figure 5.1. De façon générale, elle permet de saisir la description statique d'un scénario pour pouvoir ensuite l'inclure dans la base d'apprentissage s'il est de type historique ou de le classer s'il est de type constructeur. L'écran 3 de la figure 5.4. permet la saisie des paramètres descriptifs d'un scénario. A chaque paramètre (ou attribut) visualisé sur écran est associé un bouton de demande d'affichage des valeurs possibles (domaine). Le choix d'une ou plusieurs valeurs se fait par boîtes à cocher et la saisie est contrôlée. Par exemple, pour un attribut de type énuméré, l'utilisateur ne peut cocher qu'une seule case parmi l'ensemble des cases disponibles (figure 5.4., écran 4) alors que pour un attribut de type multivalué, cocher simultanément plusieurs cases est possible (figure 5.4., écran 5).

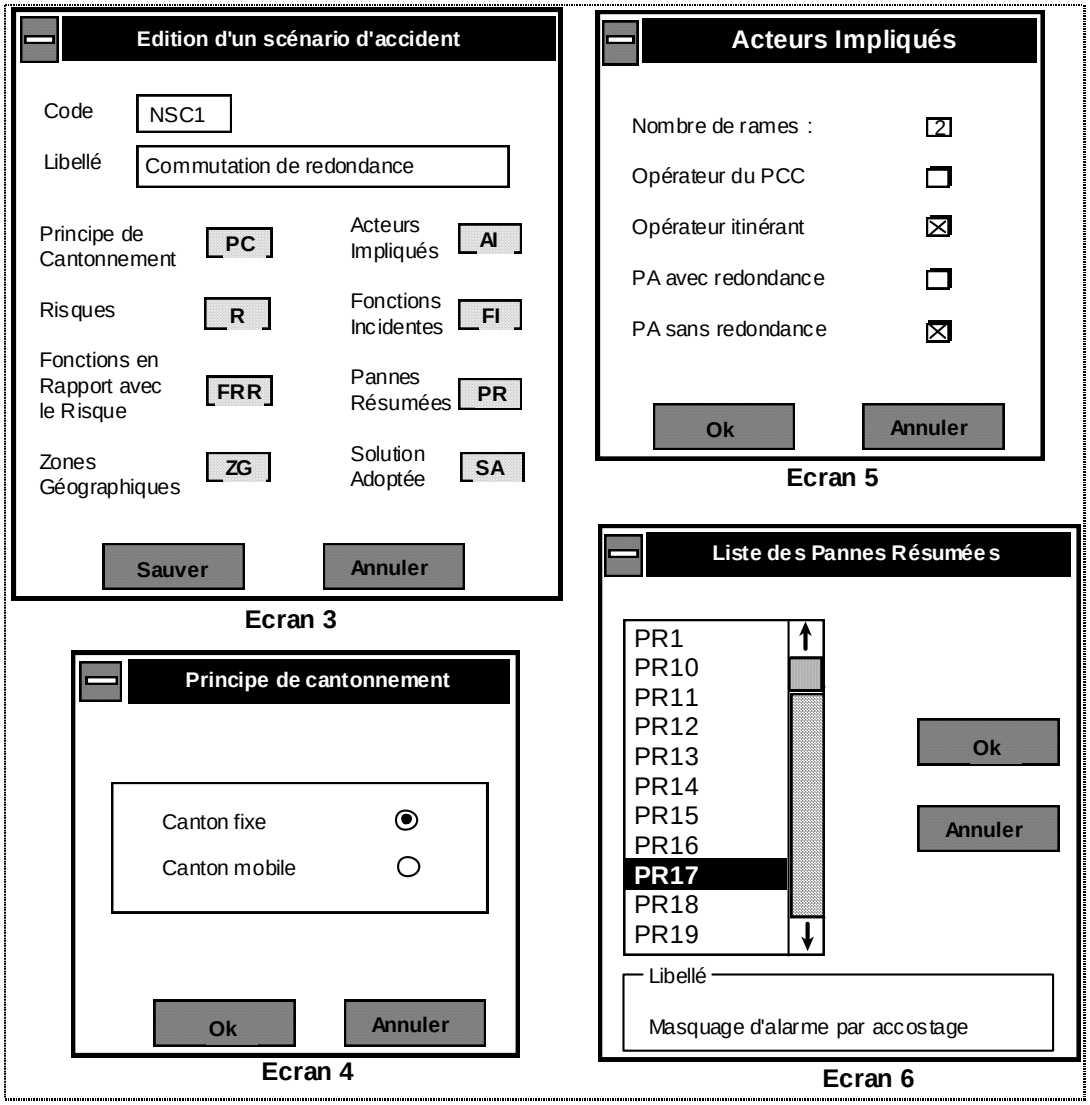


Figure 5.4. : Menus de saisie de la description statique d'un scénario d'accident

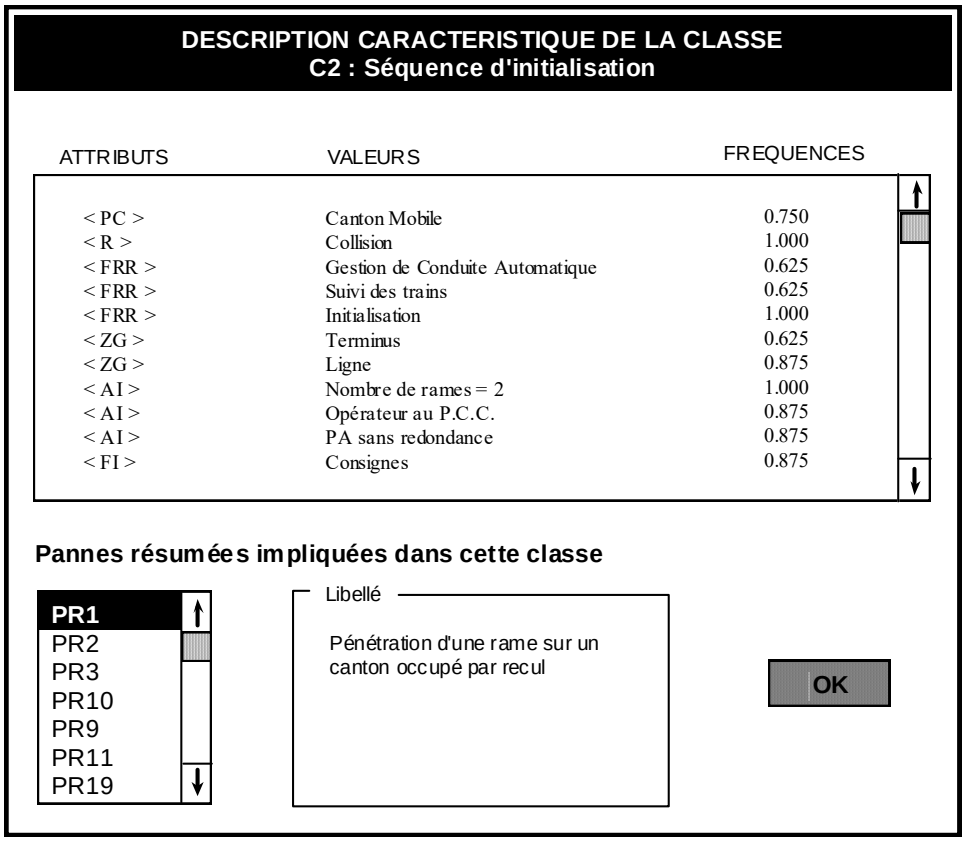
1.1.3. Apprentissage de descriptions conjonctives de classes de scénarios

Cette troisième étape qui correspond à la commande "apprentissage" du menu de la figure 5.1. assure l'induction incrémentale des descriptions conjonctives de classes de scénarios, à partir de l'ensemble des scénarios historiques précédemment appris. Le but est à la fois de considérer séquentiellement les nouveaux scénarios et d'apprendre progressivement des descriptions de classes. Les concepts appris sont décrits pour une partie par des conjonctions de descripteurs (attribut/valeur) traduisant les propriétés communes à tous les scénarios; chaque descripteur est de plus assorti d'une fréquence d'apparition. Pour l'autre partie, une conjonction de PR impliquées est associée à chaque concept. Considérons les sept exemples de scénarios repérés S4, S5, S6, S21, S22, S23, S34 appartenant à la classe C2 relative à la "séquence d'initialisation" (figure 5.5.).

GROUPES DE SCENARIOS	IDENTIFICATION
S1, S8, S29, S30,	C1 : Commutation de redondance
S4, S5, S6, S21, S22, S23, S34	C2 : Séquence d'initialisation
S3, S7, S17, S19, S25, S38	C3 : Localisation des trains
S9, S10, S20, S24	C4 : Gestion du freinage d'urgence
S13, S14, S15, S32, S33, S35, S40	C5 : Accostage
S2, S12,S16, S18	C6 : Gestion de sens de marche
S11, S31, S36, S37, S39	C7 : Contrôle d'Entrée/sortie
S26, S41, S42	C8 : Suivi de l'ordre des trains
S27, S28, S43	C9 : Conduite manuelle

Figure 5.5. : Regroupement de scénarios d'accidents historiques
(Expertise acquise auprès de l'expert)

L'étape d'induction permet d'engendrer la description caractéristique de la classe C2 (figure 5.6.) :



Ecran 7

Figure 5.6. : Description caractéristique de la classe C2

1.1.4. Classification d'un nouveau scénario

Au cours de cette quatrième étape de "classification" (menu principal d'ACASYA, figure 5.1.), le système utilise la connaissance apprise précédemment pour déduire la classe d'appartenance d'un nouveau scénario d'accident. Pour prolonger l'exemple précédent nous avons choisi de classer le nouveau scénario suivant proposé par le constructeur :

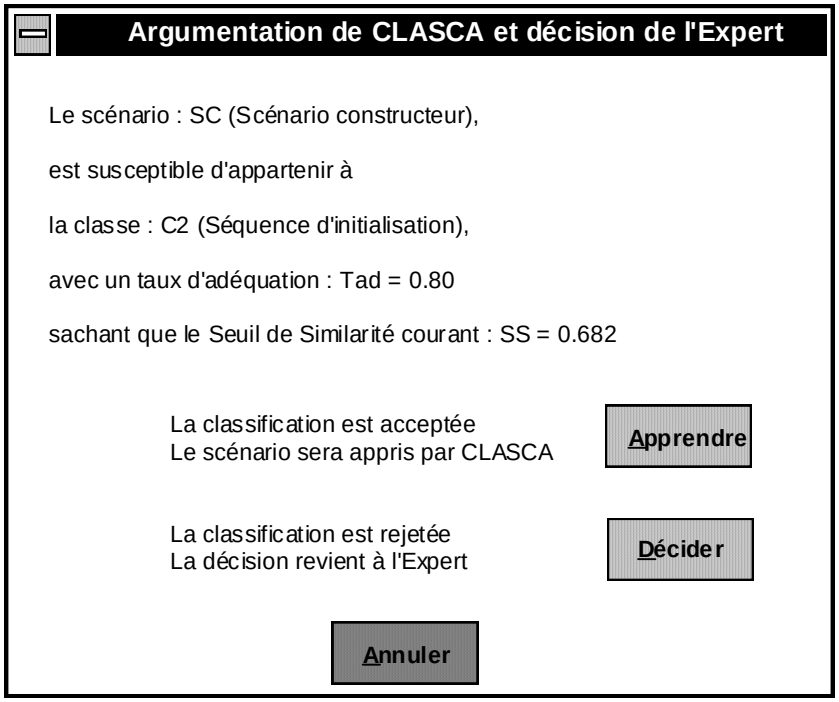
code : SC

libellé : Scénario Constructeur

PC = canton mobile
R = collision
FRR = gestion de conduite automatique
FRR = suivi des trains
FRR = initialisation
ZG = terminus
AI = opérateur au PCC
AI = PA sans redondance
FI = consignes
PR = PR 20

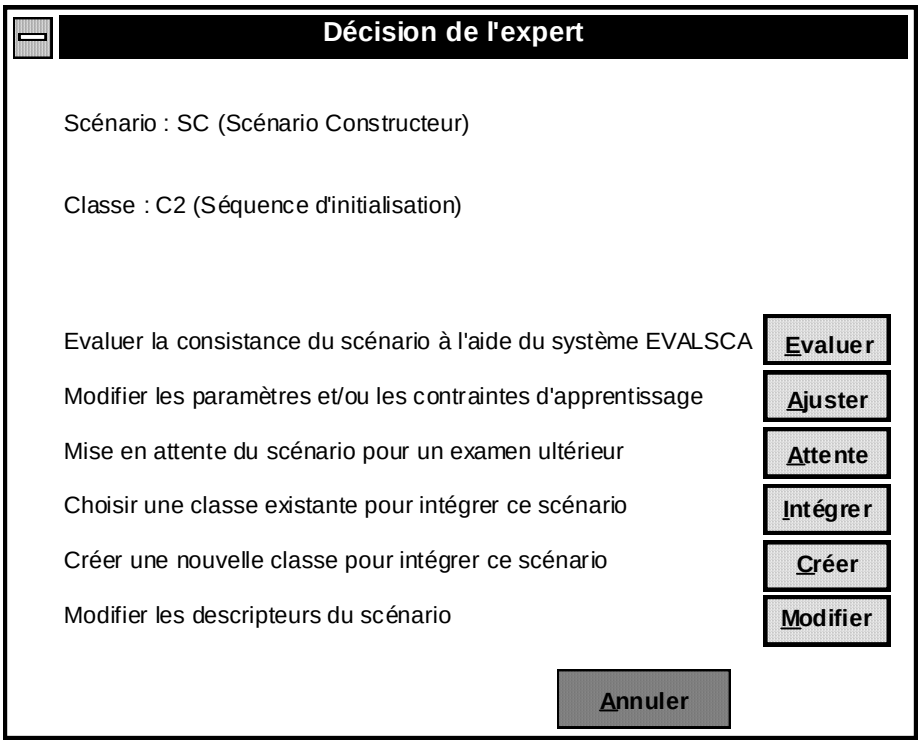
Après la saisie de la description de ce scénario, la phase de "classification" peut être lancée. Le résultat de la classification est présenté en figure 5.7. Deux parties se distinguent sur cet écran : la première pour l'argumentation et la seconde pour la prise de décision. Dans la partie argumentation, CLASCA relate le

cheminement de la phase déductive afin de construire une explication ou justification de la classification. Dans la deuxième partie de l'écran, l'expert choisit d'accepter ou de rejeter la classification proposée. S'il la rejette, un menu (écran 9, figure 5.8.) s'affiche pour préciser les suites envisageables à son refus.



Ecran 8

Figure 5.7. : Menu d'argumentation du système et de décision de l'expert



Ecran 9

Figure 5.8. : Menu des décisions de l'expert

Dans la suite de notre exemple, on suppose que l'expert a décidé d'évaluer la consistance du scénario à l'aide du système EVALSCA en cochant la case "évaluer" qui correspond à l'étape 8 de la démarche décrite au paragraphe 1.1.8.

1.1.5. Constitution de la base d'apprentissage centrée sur les PR impliquées dans la description de la classe d'appartenance du nouveau scénario

Cette cinquième étape de la démarche est réalisée lorsque l'expert décide de faire évaluer par EVALSCA la complétude du scénario avant d'en valider la classification proposée par CLASCA. Au préalable, il est nécessaire d'élaborer la base d'exemples d'apprentissage à partir des connaissances précédemment induites par CLASCA. Elle est obtenue en sélectionnant pour chacune des PR engendrées, les scénarios historiques appris dont la description fait intervenir ces PR. Sur le plan technique, cette opération a nécessité de faire le lien entre les deux systèmes CLASCA et CHARADE. Une conversion de format est indispensable puisque CLASCA est programmé en Turbo Pascal et CHARADE en Le-lisp. Le programme de transfert des données engendre deux fichiers qui constituent les données d'entrée de CHARADE : un fichier de descripteurs (classe.des) et un fichier d'exemples (classe.ex). Nous avons automatisé cette tâche longue et fastidieuse d'acquisition des descripteurs et des exemples afin d'éviter les erreurs de saisie (bruit blanc).
Le paragraphe suivant présente un échantillon de la base d'apprentissage relative à la classe C2 "séquence d'initialisation". Cette base regroupe un fichier de 9 descripteurs et un fichier de 36 exemples de scénarios. Chaque descripteur est décrit par son nom (ou attribut), son type (énuméré, multivalué,...) et son domaine de définition (ensemble des valeurs possibles).
Chaque exemple est décrit par une conjonction de descripteurs (attribut/valeur).

```
;---------------------
;;;; descripteurs ;;;;
;---------------------

(setq descripteurs 9)

(putprop 'descr1
'principe_de_cantonnement
'nom)

(putprop 'descr1
'enumere
'type)

(putprop 'descr1
'(canton_fixe canton_mobile)
'domaine)

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
(putprop 'descr2
'fonctions_en_rapport_avec_le_risque
'nom)

(putprop 'descr2
'multivalue
'type)

(putprop 'descr2
'(gestion_de_conduite_automatique localisation_des_trains controle_entree_sortie suivi_des_trains
gestion_de_sens_de_marche autorisation_ci_ht initialisation conduite_manuelle gestion_des_alarmes evacuation accostage)
```

```
'domaine)

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
(putprop 'descr3
'zones_geographiques
'nom)

(putprop 'descr3
'enumere
'type)

(putprop 'descr3
'(terminus ligne zone_injection_de_rame limite_de_troncon)
'domaine)

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
(putprop 'descr4
'pannes_resumees
'nom)

(putprop 'descr4
'multivalue
'type)

(putprop 'descr4
'( PR1 PR2 PR3 PR10 PR9 PR11 PR19 PR20 PR27)
'domaine)

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
;------------------
;;;; exemples ;;;;;
;------------------
(setq exemples 37)

;;;;;;;;;;description;;;;;;;;;;;;

;; Scenario :   SH4
(putprop 'ex1
'((Principe_de_cantonnement = canton_fixe)
(risque = collision)
(fonctions_en_rapport_avec_le_risque = initialisation)
(zones_geographiques = ligne)
(rames = Oui)
(acteurs_impliques = operateur_au_pcc)
(acteurs_impliques = pa_sans_redondance)
(pannes_resumees = PR1)
(pannes_resumees = PR2)
(pannes_resumees = PR3))
'description)

;; Scenario :   SH5
(putprop 'ex2
'((Principe_de_cantonnement = canton_fixe)
(risque = collision)
(fonctions_en_rapport_avec_le_risque = localisation_des_trains)
(fonctions_en_rapport_avec_le_risque = autorisation_ci_ht)
(fonctions_en_rapport_avec_le_risque = initialisation)
(fonctions_en_rapport_avec_le_risque = gestion_des_alarmes)
(zones_geographiques = ligne)
(rames = Oui)
(acteurs_impliques = operateur_au_pcc)
(acteurs_impliques = pa_sans_redondance)
(fonctions_incidentes = consignes)
(pannes_resumees = PR10)
(pannes_resumees = PR9))
'description)
```

```
;; Scenario :  SH6
(putprop 'ex3
'((Principe_de_cantonnement = canton_mobile)
(risque = collision)
(fonctions_en_rapport_avec_le_risque = localisation_des_trains)
(fonctions_en_rapport_avec_le_risque = initialisation)
(zones_geographiques = terminus)
(rames = Oui)
(acteurs_impliques = operateur_au_pcc)
(acteurs_impliques = operateur_itinerant)
(acteurs_impliques = pa_avec_redondance)
(fonctions_incidentes = consignes)
(pannes_resumees = PR11))
'description)

;; Scenario :  SH21
(putprop 'ex4
'((Principe_de_cantonnement = canton_mobile)
(risque = collision)
(fonctions_en_rapport_avec_le_risque = gestion_de_conduite_automatique)
(fonctions_en_rapport_avec_le_risque = suivi_des_trains)
(fonctions_en_rapport_avec_le_risque = initialisation)
(zones_geographiques = terminus)
(rames = Oui)
(acteurs_impliques = operateur_au_pcc)
(acteurs_impliques = pa_sans_redondance)
(fonctions_incidentes = consignes)
(pannes_resumees = PR19)
(pannes_resumees = PR20))
'description)
```

1.1.6. Apprentissage des fonctions de reconnaissance des PR

L'objectif de cette sixième étape est de générer un système de règles traduisant les fonctions de reconnaissance des PR d'une classe de scénarios, à partir de la base d'apprentissage précédemment construite. L'emploi de CHARADE pour cette étape exige de l'utilisateur la définition des différentes contraintes d'apprentissage. Dans le cadre de notre exemple, quatre contraintes sont prises en compte:

- La condition terminale constituée par un ou plusieurs attributs non présents dans les prémisses des règles,
- La couverture définie par le nombre maximum de règles qui couvrent chaque exemple de l'ensemble d'apprentissage et qui concluent sur la condition terminale,
- Le nombre maximum de descripteurs par prémisse,
- Le "bruit" dont la quantification permet d'éviter une généralisation abusive est défini par le nombre minimal d'exemples qui doivent vérifier une relation pour que celle-ci soit généralisée en règle. Lorsque le bruit vaut 0, CHARADE détecte toutes les régularités observables sur la base d'exemples et engendre les règles. Cette valeur peut être augmentée par la suite pour éliminer le bruit ou engendrer rapidement les régularités les plus fréquentes, sans se soucier des phénomènes ponctuels. La contrainte "bruit" joue un rôle prépondérant dans l'évaluation du contenu du scénario constructeur, étroitement liée au degré de généralisation des règles obtenues par CHARADE. Plus le facteur de bruit (FB) est élevé plus les règles concluant sur les PR sont pertinentes et ont un fort pouvoir de décision. On peut, par exemple, envisager les trois niveaux d'évaluation suivants :

. Faible degré d'évaluation : FB = 0.

. Moyen degré d'évaluation : FB = 3.

. Fort degré d'évaluation : FB = 5.

Dans le premier cas, il n'y a pas de perte d'information. Pour les deux derniers cas l'induction est plus fiable car une règle n'est émise que si elle est vérifiée par au moins 3 ou 5 exemples. Voici, pour la suite de notre exemple, une partie du système de règles engendré par CHARADE, sous les contraintes d'induction suivantes :

- Condition terminale = Pannes Résumées (les règles d'évaluation d'un scénario doivent contenir dans leur conclusion le descripteur Pannes Résumées)
- Bruit = 0 (on considère un faible degré d'évaluation pour éviter la perte d'informations)
- Nombre maximum de descripteurs par prémisse = 33 (les descripteurs des scénarios recensés sont au nombre de 33)

R1 :
Si acteurs_impliques = operateur_itinerant,
fonctions_incidentes = consignes,
acteurs_impliques = operateur_au_pcc
Alors pannes_resumees = **pr11**,
acteurs_impliques = pa_avec_redondance,
fonctions_en_rapport_avec_le_risque = localisation_des_trains,
zones_geographiques = terminus. [0]

R2 :
Si principe_de_cantonnement = canton_fixe,
fonctions_en_rapport_avec_le_risque = initialisation,
fonctions_incidentes = consignes
Alors pannes_resumees = **pr10**,
fonctions_en_rapport_avec_le_risque = autorisation_ci_ht,
fonctions_en_rapport_avec_le_risque = gestion_des_alarmes,
fonctions_en_rapport_avec_le_risque = localisation_des_trains. [0]

R3 :
Si fonctions_en_rapport_avec_le_risque = localisation_des_trains,
acteurs_impliques = pa_sans_redondance
Alors pannes_resumees = **pr9**,
zones_geographiques = ligne,
principe_de_cantonnement = canton_fixe. [0]

R4 :
Si zones_geographiques = zone_injection_de_rame,
fonctions_en_rapport_avec_le_risque = accostage,
principe_de_cantonnement = canton_mobile
Alors pannes_resumees = **pr20**. [0]

R5 :
Si principe_de_cantonnement = canton_fixe,
fonctions_en_rapport_avec_le_risque = gestion_de_conduite_automatique
Alors pannes_resumees = **pr2**. [0]

R6 :
Si principe_de_cantonnement = canton_mobile,
acteurs_impliques = pa_sans_redondance
Alors pannes_resumees = **pr19**. [0]

Pour les 36 exemples de scénarios d'accidents qui constituent la base d'exemples d'apprentissage relative à la classe C2, CHARADE a induit un système de 59 règles. Une première évaluation par les experts certifieurs a montré, d'une part la cohérence des règles engendrées et d'autre part leur intérêt pour stimuler la recherche de nouvelles situations d'insécurité.

1.1.7. Déduction des PR à considérer dans le scénario à évaluer

A l'étape précédente, CHARADE a créé un système de règles à partir de la base d'exemples d'apprentissage courante relative à la classe Ck proposée par le système CLASCA. L'étape de déduction des PR nécessite au préalable une phase de transfert des données. Celle-ci s'attache à traduire le système de règles induit par CHARADE pour le rendre compatible avec le générateur de systèmes experts IS2 avant de le transférer automatiquement vers la base de connaissances d'IS2. La phase de traduction transforme les descripteurs multivalués non exploitables par IS2 en variables booléennes, en supprimant les occurences de type "Attribut =" et en préservant la valeur de l'attribut. La phase de transfert consiste à compiler automatiquement le système de règles traduit (classe.rl) en règles compréhensibles par IS2 (règles.rxt) et à remplir le dictionnaire d'IS2 par l'ensemble des faits (fait.dxt). Nous présentons ci-après un exemple des faits et règles obtenus après traduction.

Exemples de faits

```
~~~~~~~~~~~~~~~~~
@fait 1
litteral = "canton_fixe"
proprietes = booleen,affichable,fixe

@fait 2
litteral = "collision"
proprietes = booleen,affichable,fixe

@fait 3
litteral = "deraillement"
proprietes = booleen,affichable,fixe

@fait 4
litteral = "consigne_de_vitesse"
proprietes = booleen,affichable,fixe

@fait 5
litteral = "gestion_des_alarmes"
proprietes = booleen,affichable,fixe

@fait 6
litteral = "consignes"
proprietes = booleen,affichable,fixe

@fait 7
litteral = "pannes_resumees"
proprietes = symbole,affichable,compteur
domaine = "PR1","PR2","PR3","PR9","PR10","PR11","PR19","PR20","PR27"
```

Exemple de règles
~~~~~~~~~~~~~~~~~~~

@REGLE
R 1.
SI operateur_itinerant ET consignes ET operateur_au_pcc
ALORS pannes_resumees="pr11" ET pa_avec_redondance ET localisation_des_trains ET terminus

@REGLE
R 2.
SI canton_fixe ET initialisation ET consignes
ALORS pannes_resumees="pr10" ET autorisation_ci_ht ET gestion_des_alarmes ET localisation_des_trains

@REGLE
R 3.
SI localisation_des_trains ET pa_sans_redondance
ALORS pannes_resumees="pr9" ET ligne ET canton_fixe

@REGLE
R 4.
SI gestion_des_alarmes ET gestion_de_conduite_automatique
ALORS conduite_manuelle ET limite_de_troncon

A partir des paramètres descriptifs du scénario constructeur hors PR et des règles induites par CHARADE puis transformées pour IS2, le système expert fonctionnant en chaînage avant déduit les PR qui doivent être impliquées dans la description du scénario constructeur à évaluer. Pour notre exemple, le système expert a déduit la PR19. La trace de la déduction est présentée ci-après.

@@ 20 / 07 / 1992
. canton_mobile
. collision
. gestion_de_conduite_automatique
. suivi_des_trains
. initialisation
. terminus
. opérateur_au_pcc
. pa_sans_redondance
. consignes
DEDUCTION : Panne_résumée = pr19
~~~~~~~~~~~~~~~~~~~

1.1.8. Evaluation de la complétude du scénario sur la base des PR

Dans cette huitième étape de la démarche, les PR **envisageables** déduites par le système expert sont analysés et comparées aux PR **envisagées** par le constructeur. Cette confrontation peut engendrer trois cas :

- Une ou plusieurs PR non prises en compte par le constructeur lors de la conception des équipements de protection sont susceptibles de mettre en cause la sécurité du système de transport.
- Il y a plus de PR envisagées que de PR envisageables : on peut admettre que l'analyse de sécurité a été effectuée en profondeur et que même des risques d'accident à faible probabilité d'occurrence ont été envisagés. Ce cas peut aussi se rencontrer lors de la prise en compte de l'évolution technologique des systèmes de transport automatisés.
- les PR envisageables et envisagées sont identiques. Le scénario constructeur est donc complet et l'analyse de sécurité effectuée semble exhaustive.

Notre exemple traduit simultanément les deux premiers cas puisque le constructeur a considéré la PR20 non déduite par le système et que ce dernier a déduit la PR19 non envisagée par le constructeur. La déduction de la PR19 incite l'expert certifieur à formuler de nouveaux scénarios d'accidents susceptibles de mettre en défaut l'analyse de sécurité proposée par le constructeur.

1.1.9. Mise à jour des connaissances produites antérieurement

Après l'étape d'analyse et d'évaluation du scénario constructeur, l'utilisateur du système ACASYA a le choix entre deux possibilités :

. Il peut revenir au menu "argumentation et décision" (figure 5.7.) :
- pour confirmer la complétude du scénario en sélectionnant l'option "apprendre". Dans ce cas, le nouveau scénario constructeur est appris par CLASCA et devient un scénario historique qui enrichit la BCHS.
- pour reconsidérer la démarche (figure 5.8.) en ajustant les paramètres d'apprentissage dans le menu de préconception (figure 5.2.) ou en complétant ou modifiant la description du scénario.

. Il peut retourner à l'étape d'apprentissage des fonctions de reconnaissance des PR, c'est à dire à l'étape de génération des règles par CHARADE. Cette option est sélectionnée lorsque les déductions réalisées par le SBC n'aboutissent à aucune PR, en cas de non représentativité de la base d'apprentissage ou de niveau trop élevé des contraintes d'apprentissage. En effet, si ces dernières sont trop fortes, certaines régularités présentes dans la base d'apprentissage ne sont pas traduites sous forme de règles car elles ne satisfont pas le seuil minimal requis par la contrainte. Dans ce cas, l'utilisateur peut décider d'affaiblir les valeurs de certaines contraintes et notamment celles relatives au "bruit", de manière à obtenir des règles supplémentaires. Ces règles sont susceptibles de permettre des déductions qui aboutissent à des PR non formulées précédemment. D'une façon générale, la méthode de détermination de ces contraintes consiste à débuter le cycle d'apprentissage avec des contraintes très fortes puis à les relâcher au cours des cycles

suivants. En résumé, ce processus d'obtention du système de règles final est un processus itératif et interactif entre le système CHARADE et l'utilisateur. Son efficacité dépend, en grande partie, des appréciations subjectives portées par ce dernier /HADJ MABROUK 89/ et /DUPAS 90/.

1.2. Intérêts et extensions du système ACASYA

Le paragraphe précédent a présenté une première évaluation d'ACASYA sur la base d'un exemple simple mais représentatif de l'articulation des différents modules impliqués et du comportement de la maquette réalisée. Le présent paragraphe propose un bilan des apports d'ACASYA, en regard du cahier des charges initialement défini, ainsi que ses perspectives d'amélioration et d'extension.

1.2.1. Intérêts d'ACASYA

Le système ACASYA réalisé pour l'aide à la certification des systèmes de transport terrestres automatisés répond aux objectifs de **classification** et d'**évaluation** des scénarios d'accidents.

ACASYA montre la **complémentarité** des techniques d'apprentissage automatique et d'acquisition de connaissances pour faciliter le processus de transfert des connaissances de certification. Il prend en compte l'**interactivité Expert/Système** pour contrôler et compléter les connaissances produites.

Le formalisme de représentation des scénarios d'accidents composé des deux descriptions statique et dynamique permet d'aider les certifieurs à mieux **structurer et conceptualiser** leur savoir et de proposer éventuellement aux constructeurs des systèmes de transport un cadre méthodologique pour une définition plus exhaustive des scénarios.

Renforcer en aval les méthodes classiques d'analyse de sécurité constitue l'un des apports originaux d'ACASYA /HADJ-MABROUK 92e/. En effet, la démarche proposée exploite par des techniques d'apprentissage les connaissances historiques obtenues à l'aide des méthodes d'analyse de sécurité telles que la MCPR, l'AMDEC, l'APD présentées dans le premier chapitre. L'objectif est de dégager des situations d'insécurité non prévues par le constructeur afin d'améliorer l'exhaustivité de l'analyse des risques.

Contrairement aux systèmes d'aide au diagnostic, ACASYA se présente comme un outil d'**aide à la prévention des pannes**. Lors de la conception d'un nouveau système, le constructeur s'engage à respecter les objectifs de sécurité définis. Il doit démontrer que le système est réalisé de telle sorte que l'ensemble des risques est couvert et que le système ne peut, en aucun cas, être un danger pour les utilisateurs. A l'opposé, les experts de certification visent à montrer que le système n'est pas sécuritaire et dans ce cas à déceler les causes d'insécurité. Construit dans cette seconde optique, ACASYA est un outil potentiellement précieux d'aide à la certification qui évalue la complétude de l'analyse proposée par le constructeur.

1.2.2. Améliorations et extensions d'ACASYA

ACASYA est à ce jour une maquette dont la première validation montre l'intérêt de la méthode d'aide à la certification proposée et qui de ce fait requiert certaines améliorations et extensions. Les connaissances d'analyse de sécurité acquises à ce jour sont loin d'être représentatives du domaine de la certification et nécessitent, d'une part, d'être complétées par d'autres scénarios relatifs au risque de collision et d'autre part, d'être étendues à plusieurs autres risques d'accident (déraillement, électrocution ...). Dans un premier temps, la conception d'une version intégrée d'un prototype d'ACASYA est nécessaire pour finaliser les résultats de la maquette et tenir compte des remarques formulées. Ce prototype s'attachera à :

- Assurer une unicité d'environnement de développement Système/Système,
- Homogénéiser les interfaces Homme/Machine,
- Intégrer la version C++ de CHARADE,
- Répondre aux aspirations de la phase d'aide à la génération des scénarios qui n'a pas encore été développée dans le cadre du présent travail.

L'objectif est de concrétiser dans le système ACASYA la totalité des concepts définis dans la méthodologie d'aide à la certification (chapitre 4) avant d'en envisager l'extension à la sûreté de fonctionnement d'autres systèmes de transport terrestres, maritimes, aériens ou même à celle d'édifices sensibles tels que les centrales nucléaires.

1.3. Perspective d'évolution d'ACASYA : une approche d'aide à la génération des scénarios d'accidents

La méthodologie élaborée pour aider à la certification repose essentiellement sur l'utilisation des techniques d'apprentissage et d'un modèle "générique" de représentation des scénarios d'accidents faisant intervenir les deux types de description : statique et dynamique. Compte tenu de l'ampleur du problème, notre contribution s'est essentiellement axée sur les activités de classification et d'évaluation impliquées dans les deux premiers niveaux de la méthodologie et concrétisées dans les modules logiciels CLASCA et EVALSCA. Ces derniers font appel uniquement à la description statique d'un scénario. Pour compléter ces travaux, il convient maintenant de poser le problème de la génération de nouveaux scénarios d'accidents qui fait appel à la description dynamique ainsi qu'aux résultats produits par CLASCA et EVALSCA. C'est l'objet du module GENESCA évoqué dans la méthodologie et détaillé dans /MEJRI 92/.

La démarche d'aide à la génération prévue repose sur l'injection d'une (ou plusieurs) Pannes Résumées (PR) dans le Tableau de Séquencement du Marquage (TSM) afin de "casser" les protections prévues par le constructeur et par conséquent de mettre en défaut la sécurité du système de transport. On rappelle qu'un TSM présente la chronologie des évènements conduisant à une situation d'insécurité critique (la collision par exemple). Une protection peut être définie comme une fonction de sécurité ou une fonction en rapport avec le risque, adoptée par le constructeur pour se prémunir contre le danger traduit par l'occurrence d'une ou plusieurs PR.

Dans cette approche on envisage d'injecter aléatoirement la PR, précédemment déduite, à chaque étape du TSM afin de voir si l'on aboutit à une situation dangereuse non considérée par le constructeur. L'objectif est

d'observer la réaction des fonctions de protection pour déceler les faiblesses de conception sur le plan de la sécurité. L'avantage de cette approche résulte de l'exploration exhaustive de tout l'espace potentiel de génération des scénarios. Cependant, une telle exploration est sujette à une explosion combinatoire. Une solution pour rendre cette exploration "intelligente" et restreindre l'espace de recherche potentiel, consiste à déterminer l'instant propice d'injection de la PR dans le TSM. Réaliser une telle opération nécessite de disposer de connaissances sur les conditions d'applicabilité des PR ou en d'autres termes sur la manière d'altérer les fonctions de protection prévues par le constructeur. Si ce domaine est vaste, on peut par exemple recourir à une technique d'apprentissage automatique permettant d'opérer par généralisation sur l'ensemble des conditions d'applicabilité et d'engendrer les règles consistantes d'applicabilité d'une PR.

Néanmoins, cette approche de génération qui repose essentiellement sur la description dynamique d'un scénario d'accident est non exhaustive car l'opération d'injection des PR dans le TSM est relative à **un seul scénario**. De ce fait, la découverte de nouvelles situations d'insécurité est limitée et peut parfois déboucher sur des situations non significatives. Une amélioration de cette approche consiste à élaborer non plus un modèle de représentation d'un scénario mais un **modèle de représentation d'une classe de scénarios**. L'opération d'injection de PR s'effectue alors dans un TSM relatif à une classe de scénarios. L'élaboration d'un modèle de représentation des classes de scénarios nécessite le recours aux deux types de connaissances suivants :

- Le premier regroupe les classes de scénarios ainsi que leurs descriptions caractéristiques produites par CLASCA. On rappelle que la description caractéristique d'une classe de scénarios est formée d'une conjonction de descripteurs les plus pertinents caractérisant la classe. Cette description caractéristique peut être considérée comme le **modèle statique de la classe**.
- Le second type de connaissances concerne les descriptions dynamiques (RdP et TSM) de chaque scénario d'accident. Pour chaque classe de scénarios repérée par CLASCA, on peut recenser la liste des descriptions dynamiques relatives aux scénarios de la classe. La tâche délicate à réaliser consiste à trouver une technique permettant de combiner l'ensemble de ces descriptions dynamiques pour former un **modèle dynamique par classe** de scénarios.

2. EVALUATION DU SYSTEME CLASCA ET PERSPECTIVES

L'application d'ACASYA au sous problème de la collision constitue une première évaluation globale de la maquette réalisée. Le présent paragraphe propose une évaluation focalisée sur le module d'apprentissage CLASCA ainsi que les perspectives envisagées pour son amélioration.

2.1. Protocole d'évaluation de CLASCA

La qualité d'un apprentissage se mesure en évaluant la **validité** et l'**utilité** des connaissances apprises. Une connaissance est valide si elle est adéquate et cohérente par rapport à ce que l'on sait déjà du domaine de certification. Elle est utile si elle contribue à réaliser les objectifs définis. Le contrôle de ces deux critères est du ressort des experts certifieurs. Pour tester la capacité de prédiction du système et par conséquent la représentativité des résultats obtenus par apprentissage, il existe deux solutions /SEBAG 90/ : soit l'expert dispose d'une liste de cas témoins constituant la base d'exemples test, soit l'ensemble des données initiales disponibles est partagé en base d'apprentissage et base de test selon une loi de répartition donnée par l'expert.

Initialement estimé à plusieurs centaines, le nombre d'exemples acquis à ce jour est de quarante trois pour la base d'apprentissage et de treize pour les scénarios de test (de validation). Si cette révision à la baisse du nombre de scénarios n'entame en rien la qualité de la base d'exemples dédiée au risque de collision, le problème de l'évaluation des performances du module d'apprentissage CLASCA se pose. Le tableau de la figure 5.9. présente les résultats d'évaluation obtenus pour les treize scénarios test.

Scénarios de validation (scénarios test)	Prédiction de CLASCA		Décision de l'expert	Comparaison des résultats
	Classe	Taux d'adéquation		
SV1 : divergence entre contrôle de sortie et contrôle d'entrée	C7	1.00	C7	bien classé
SV2 : enregistrement prématuré d'un élément sur une section	C7	0.77	C3	**mal classé**
SV3 : discrimination tardive mettant en cause l'identification du suiveur	C7	0.92	C7	bien classé
SV4 : défaut de transmission entre deux PA adjacents	C1	0.88	C1	bien classé
SV5 : scindage d'une unité multiple sur un DN	C7	1.00	C7	bien classé
SV6 : train de trois éléments	C7	0.83	C9	**mal classé**
SV7 : levée du FU à vitesse nulle après une demande d'évacuation	**?**	< SS	C4	non classé
SV8 : conduite manuelle sur une section en demande d'évacuation	C9	0.78	C9	bien classé
SV9 : mouvement d'aiguille en commande manuelle	**?**	< SS	**?**	non classé
SV10 : défaut de localisation et mouvement d'aiguille	C3	0.94	C3	bien classé
SV11 : inversement d'ordre d'enregistrement	C8	0.76	C8	bien classé
SV12 : train en conduite manuelle ne respectant pas les consignes de distance	C9	0.91	C9	bien classé
SV13 : cible à contre -sens	C6	0.83	C6	bien classé

Figure 5.9. : Interprétation des résultats de prédiction de CLASCA

Les résultats de classification de ces treize scénarios peuvent être interprétés comme suit :

- 9 scénarios sont "bien classés" : ils sont classés de la même façon par le système et l'expert de certification.
- 2 scénarios sont "non classés" : leur taux d'adéquation avec une classe est inférieur au seuil de similarité courant et le système ne propose aucune classe d'appartenance. Le seuil de similarité étant évolutif, ces scénarios sont momentanément mis en attente et pourront être reclassés ultérieurement. Pour l'un de ces cas, l'expert a proposé une classe et pour l'autre cas, il a abondé dans le sens de la décision du système en ne classifiant pas le scénario.
- 2 scénarios sont "mal classés" : la classification du système n'agrée pas l'expert certifieur. La principale cause de cet échec provient de la non représentativité des descriptions des classes apprises sur la base d'un petit nombre de scénarios.

Bien que cette première évaluation démontre le bien fondé de l'approche implémentée sur la maquette, une évaluation plus fine est nécessaire pour mesurer et à terme optimiser la performance du système. Pour cela, il est impératif d'enrichir la base d'exemples d'apprentissage par des scénarios supplémentaires relatifs au risque de collision ainsi que par des scénarios relatifs à tous les risques d'accidents.

2.2. Limites et perspectives du système CLASCA

La principale originalité de CLASCA réside dans la prise en compte de l'incrémentalité non monotone. Celle-ci soulève le problème de l'influence de l'ordre dans lequel les exemples sont considérés. De plus, il est dès maintenant envisageable d'enrichir les performances de CLASCA par l'adjonction de nouvelles fonctionnalités. Ces points sont développés ci-après.

2.2.1. Sensibilité du système à l'ordre de prise en compte des exemples

L'apprentissage dans CLASCA est non monotone : la description d'une classe apprise à l'instant t est remise en cause à l'instant t+1 et la fréquence d'apparition d'un descripteur est tantôt croissante, tantôt décroissante. Bien que mieux adapté aux données bruitées, l'apprentissage non monotone ne garantit pas la convergence du processus. Dans CLASCA, le problème de convergence est maîtrisé par l'introduction d'un seuil de similarité évolutif. La non monotonie de l'apprentissage influe sur la sensibilité de CLASCA à l'ordre d'arrivée des scénarios. Un exemple qui est intégré dans une classe à l'instant t n'est pas assuré de réintégrer cette même classe à l'instant t+n du fait de l'évolution de la description de la classe. Ceci est d'autant plus vrai en début d'apprentissage, lorsque la description de la classe n'est pas encore bien typée. A ce jour, nous avons admis à un premier degré que l'implication de l'expert dans le contrôle des résultats de classification permet de contourner ce problème. Une étude sur ce point devra toutefois être conduite.

2.2.2. L'apprentissage par recherche d'explications pour trancher en cas de conflit entre l'expert et le système

Le module d'"argumentation du système et de décision de l'expert" contenu dans le système CLASCA (figure 5.7.) assure une interactivité Homme-Machine. Ce dialogue entre l'expert et le système permet de contrôler les connaissances produites. Pour gérer les situations critiques issues d'un échec du système ou d'un rejet de décision par l'expert, plusieurs solutions peuvent être envisagées par l'expert. Ces solutions consistent à :

a. Affiner la description du scénario
b. Modifier les paramètres et contraintes d'apprentissage
c. Evaluer la complétude du scénario à l'aide du système EVALSCA
d. Mettre le scénario en attente pour un examen ultérieur
e. Intégrer le scénario dans une classe existante
f. Créer une nouvelle classe intégrant le scénario

Dans les cas (e) et (f), les connaissances apprises antérieurement peuvent être remises en cause. En effet, intégrer un nouveau scénario dans une classe existante non suggérée par le système signifie que l'expert met en cause, non seulement sa partition (pré-classification) initiale mais aussi toutes ses décisions antérieures puisque l'adjonction d'un nouvel exemple est validée par ses soins. Par ailleurs, créer une nouvelle classe peut amener à faire glisser des exemples classés de façon limite dans une classe vers la nouvelle classe. En fait, au fur et à mesure que le processus d'apprentissage avance, les connaissances produites sont de plus en plus significatives aux yeux de l'expert. Il se familiarise avec le système d'apprentissage, comprend mieux l'utilité de ces connaissances et devient de plus en plus exigeant vis à vis des décisions du système.

Pour traiter plus efficacement ce problème d'incohérence ou de conflit, une amélioration du module d'"argumentation du système et de décision de l'expert" ayant recours aux méthodes d'apprentissage par recherche d'explications peut être envisagée. On propose, dans un premier temps, d'approfondir l'étape d'argumentation du système, en enrichissant les critères de justification de la classification des scénarios et dans un deuxième temps d'intégrer un module d'explication des décisions de l'expert. La cohabitation de ces deux modules doit assurer un apprentissage efficace :

- Elle favorise l'acquisition de nouvelles connaissances en incitant l'expert à mieux formuler son expertise,
- Elle améliore les performances du système, non seulement lorsque celui ci aboutit au classement d'un nouvel exemple mais aussi lorsqu'il échoue dans ses tentatives; l'apprentissage par l'échec et la réussite semble une voie intéressante pour aborder ce point.

2.2.3. Génération des règles de classification dans l'hypothèse de convergence du système CLASCA

Les systèmes d'apprentissage par classification développés actuellement produisent des arbres de décision, des hiérarchies conceptuelles ou des règles de classification. La mise à jour d'une hiérarchie de concept ou d'un arbre de décision est un processus très complexe qui exige un temps de calcul important. L'obtention de

règles de classification pertinentes suppose l'existence d'une base d'exemples d'apprentissage quasi exhaustive. Dans le cadre de notre étude, nous ne disposons pas d'une telle base. C'est pourquoi, dans CLASCA, nous avons opté pour une classification qui repose sur les descriptions des classes induites et sur un critère d'adéquation. La pertinence des décisions de classification se renforce au fur et à mesure que les descriptions des classes s'enrichissent. Lorsque CLASCA converge, les descriptions des classes deviennent stables et pertinentes et les décisions de classification sont fiables. A terme, l'objectif est donc d'améliorer CLASCA en générant des règles de classification dès lors que les descriptions de classes convergent vers des valeurs stables. La description d'une classe est initialement basée sur un traitement **numérique** ou statistique qui distingue les descripteurs caractéristiques des descripteurs "bruit". Dans l'hypothèse de convergence du système, seuls les descripteurs caractéristiques des classes sont exploités pour former des règles de classification **symboliques**. Ces règles peuvent être exploitées par un système expert pour la classification d'un nouveau scénario. Dans ce cas, la prémisse de la règle contient la description caractéristique (conjonctions de descripteurs) et la conclusion de la règle contient le libellé de la classe (accostage, commutation de redondance, séquence d'initialisation, ...). Voici un exemple de règle de classification :

SI PC = canton_mobile ET R = collision ET FRR = commutation_de_redondance ET ZG = limite_de_tronçon ET AI = PA_avec_redondance
ALORS classe commutation_de_redondance

A ce jour, l'approche proposée relative à la génération des règles de classification est en cours d'implémentation pour évaluation.

Nous venons de présenter l'évaluation du système ACASYA qui a montré son intérêt pour l'aide à l'analyse de sécurité des systèmes de transport terrestres automatisés. Pour tendre vers l'exhaustivité de cette analyse, ACASYA requiert l'enrichissement de la base de connaissances historiques, l'amélioration du module d'apprentissage CLASCA ainsi que le développement du module GENESCA dédié à la génération de nouveaux scénarios d'accidents. Nous proposons dans le paragraphe suivant une perspective de recherche relative à l'amélioration du processus de transfert des connaissances.

3. COMPLEMENTARITE DE L'ACQUISITION DE CONNAISSANCES ET DE L'APPRENTISSAGE AUTOMATIQUE POUR AMELIORER LE PROCESSUS DE TRANSFERT DES CONNAISSANCES

La méthodologie d'aide à la certification des systèmes de transport terrestres automatisés développée dans le cadre de ce mémoire repose sur l'utilisation conjointe des techniques d'apprentissage automatique et d'acquisition de connaissances. La complémentarité de ces deux approches nous a permis de cerner le problème de la certification et constitue, de ce fait, une voie de recherche prometteuse que nous proposons d'approfondir pour améliorer le processus de transfert de connaissances. Ce paragraphe évoque les principaux obstacles à l'acquisition de connaissances ainsi que la complémentarité de l'apprentissage et de l'acquisition de connaissances. Il débouche sur la proposition d'un processus itératif d'acquisition de connaissances.

3.1. Obstacles à l'acquisition des connaissances

L'analyse du domaine de la certification a montré que le processus de transfert de connaissances des experts vers la machine est complexe et peu étudié et que le fameux "goulot d'étranglement" du développement d'un SBC ne se limite pas à la seule phase d'extraction de connaissances mais est également lié aux caractéristiques et à la formalisation des connaissances ainsi qu'à la collaboration Expert/Cogniticien.

Le savoir-faire des experts de certification repose sur des connaissances subjectives, empiriques et parfois implicites et ambiguës qui peuvent générer plusieurs interprétations. Il n'existe généralement pas d'explication scientifique pour justifier cette expertise compilée. Ces connaissances ne sont pas toujours conscientes chez l'expert, compréhensibles par un novice ou même exprimables par l'intermédiaire d'un langage. La transcription d'un langage verbal (naturel) en langage formel interprétable par une machine provoque souvent une distorsion de la connaissance experte. Ceci introduit un biais entre le modèle cognitif de l'expert et le modèle implémenté. Ce décalage est dû non seulement au fait que les langages de représentation employés en IA ne sont pas d'une richesse suffisante pour expliciter le fonctionnement cognitif de l'expert mais aussi à l'interprétation subjective du cogniticien.

De plus, l'expert qui est généralement le seul à disposer du savoir-faire permettant de résoudre certains problèmes cruciaux peut faire preuve d'un manque de motivation et de disponibilité. Il peut refuser de coopérer avec le cogniticien, soit parce qu'il ne voit pas ses efforts récompensés, soit par peur d'être remplacé par une machine ou par blessure d'amour-propre lorsqu'il constate que toute sa compétence, pour laquelle on l'estime, se réduit à quelques règles /NASSIET 87/. Une fois que son savoir est figé dans un SBC, l'expert peut avoir le sentiment qu'on le lui a confisqué.

Toutes ces contraintes restreignent le champs d'investigation de l'acquisition de connaissances. L'utilisation conjointe des techniques d'acquisition de connaissances et d'apprentissage automatique est une solution pour affaiblir ces contraintes.

3.2. Complémentarité de l'acquisition de connaissances et de l'apprentissage automatique

Les experts considèrent généralement qu'il est plus simple de décrire des exemples ou des cas expérimentaux plutôt que d'expliciter des processus de prise de décision. L'introduction des systèmes d'apprentissage automatique fonctionnant sur des exemples permet d'engendrer de nouvelles connaissances susceptibles d'aider l'expert à résoudre un problème particulier. L'expertise d'un domaine est non seulement détenue par les experts mais aussi répartie et emmagasinée implicitement dans une masse de données historiques que l'esprit humain éprouve des difficultés à synthétiser. Extraire de cette masse d'informations des connaissances pertinentes dans un but explicatif ou décisionnel constitue l'un des objectifs de l'apprentissage automatique caractérisé par l'amélioration des performances avec l'expérience. L'apprentissage à partir d'exemples est toutefois insuffisant pour acquérir la totalité du savoir faire des experts et nécessite le recours à l'acquisition de connaissances pour identifier le problème à résoudre, extraire et formaliser des connaissances accessibles par les moyens usuels d'acquisition. En ce sens, chacune des deux approches peut combler les faiblesses de l'autre. Pour améliorer le processus de transfert d'expertise, il est donc intéressant de concilier ces deux approches dans le processus itératif d'acquisition de connaissances présenté ci-après.

3.3. Processus itératif d'acquisition des connaissances

Ce paragraphe propose un processus itératif de transfert des connaissances (figure 5.10.) modélisé par un diagramme SADT décrivant les quatre activités d'acquisition, d'apprentissage, d'évaluation et d'enrichissement des connaissances.

A partir des connaissances initiales du domaine (connaissances expertes et historiques), l'acquisition de connaissances permet de construire un modèle du raisonnement de l'expert et un modèle de représentation des exemples et d'obtenir un ensemble d'exemples et de classes d'objets. Ces connaissances **acquises** sont exploitées par apprentissage pour produire de nouvelles connaissances **apprises** qui seront ensuite **évaluées** par l'expert du domaine. La confrontation des connaissances découvertes par l'apprentissage aux connaissances acquises auprès de l'expert permet d'**enrichir** les connaissances initiales du domaine. Il y a toujours un décalage entre les connaissances acquises et les connaissances réellement détenues par l'expert. En effet, on peut rarement extraire du premier coup l'ensemble des connaissances expertes mais lorsqu'on présente à l'expert les connaissances apprises par le système, il est conscient de leur intérêt, repère des contradictions, des "trous" ou des règles pertinentes. Il peut fournir un avis sur le choix des exemples et des descripteurs, interpréter les résultats produits par apprentissage, améliorer le modèle d'expertise acquis préalablement, corriger et compléter le langage de description des exemples et ajuster les paramètres d'apprentissage.

En incitant l'expert à mieux verbaliser son expertise, on contribue donc à l'enrichissement des connaissances du domaine. En cela, la prise en compte de l'expert dans le processus s'avère indispensable pour améliorer le transfert d'expertise. On trouve donc ici un nouveau volet du concept de coopération Expert/Système qui doit dépasser le but de la seule mise en oeuvre d'interfaces de dialogue ergonomiques et s'orienter vers un accroissement des capacités décisionnelles des systèmes afin d'instaurer un "partenariat" avec l'opérateur humain /MILLOT 88, 90/.

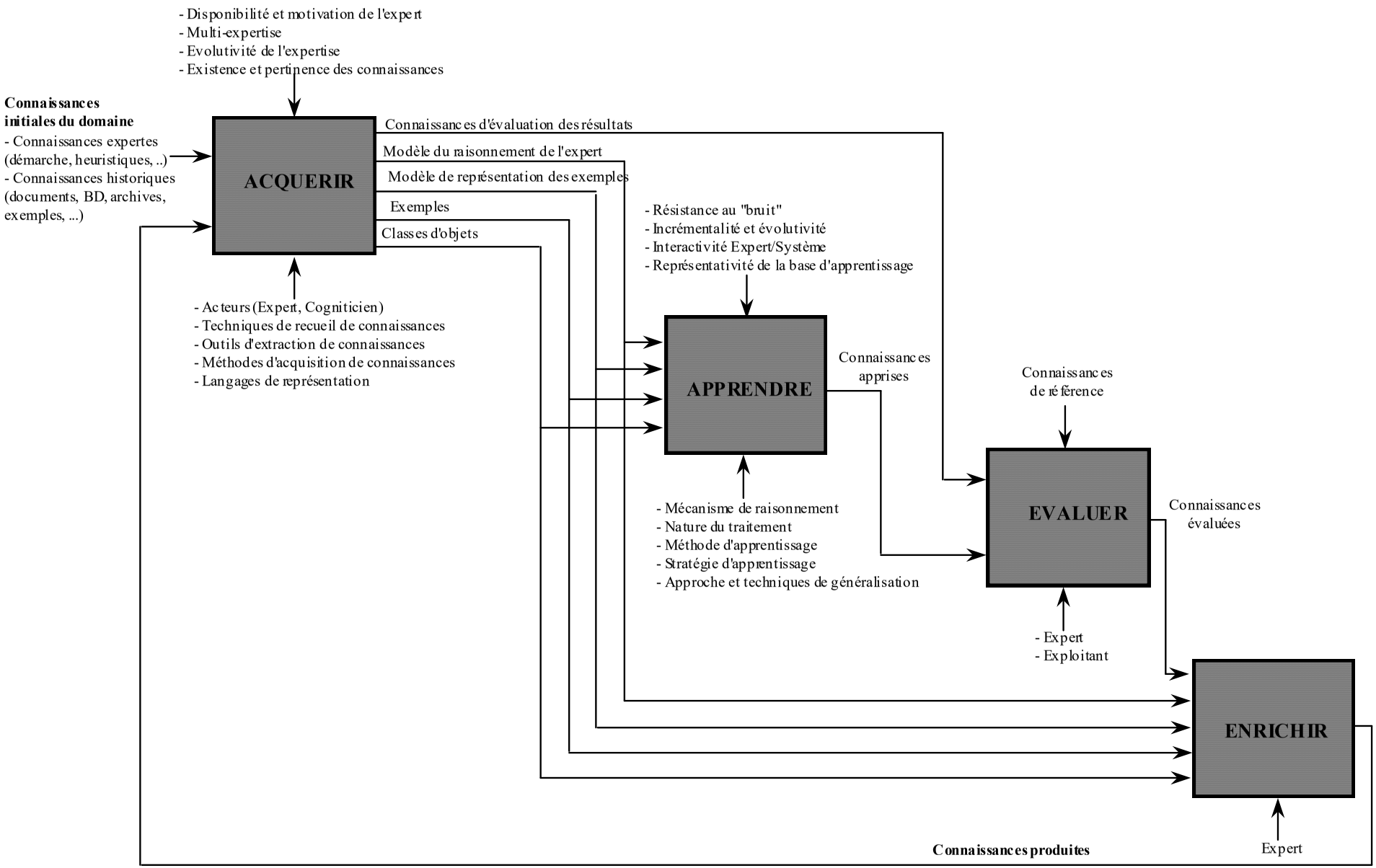


Figure 5.10. : Processus itératif d'acquisition des connaissances

CONCLUSION

Ce dernier chapitre a présenté l'évaluation de la méthodologie d'aide à la certification ainsi que les perspectives de recherche envisagées. La validation de cette méthodologie s'est déroulée en deux phases. La première phase a consisté en une évaluation globale de la maquette de faisabilité du système ACASYA (Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés) sur la base de son application à l'ensemble des scénarios archivés à ce jour pour le risque de collision. La seconde phase s'est focalisée sur une évaluation interne du module de classification CLASCA en s'appuyant sur un lot de scénarios de test fourni par les experts. Cette validation a révélé les intérêts mais aussi les limites actuelles de la maquette réalisée et a débouché sur des perspectives d'amélioration.

Cette maquette atteste de l'intérêt et de la validité de la méthodologie élaborée pour classifier et évaluer les scénarios d'accidents, ce qui constituait notre objectif initial. ACASYA est considéré comme un système qui favorise la recherche de nouvelles situations d'insécurité susceptibles de rendre exhaustive l'analyse des risques d'accidents. Afin que ce système puisse être évalué en milieu industriel, il est nécessaire de prévoir l'enrichissement de la base des connaissances acquises et de spécifier et développer le troisième niveau de la méthodologie relatif au module GENESCA, destiné à la génération des nouveaux scénarios d'accidents. Parallèlement, diverses améliorations du module d'apprentissage CLASCA sont envisagées. Elles portent sur la résolution des conflits potentiels entre l'expert et le système par l'intégration d'un module d'explication, sur la prise en compte de la sensibilité à l'ordre d'arrivée des exemples ainsi que sur la génération de règles de classification. Enfin, la réalisation de cet outil d'aide à la certification des systèmes de transport terrestres automatisés résulte de l'utilisation conjointe des techniques d'acquisition de connaissances et d'apprentissage automatique qui est une voie de recherche intéressante pour contribuer au développement des SBC.

CONCLUSION GENERALE

Ce mémoire a présenté notre contribution à l'amélioration des méthodes usuelles d'analyse de sécurité employées dans le cadre de la certification des systèmes de transport terrestres automatisés (STA). Cette contribution, basée sur l'apprentissage automatique et l'acquisition de connaissances ainsi que sur une analyse pratique des problèmes liés à la sécurité des STA, s'est concrétisée par l'élaboration d'une méthodologie d'aide à la certification, dans un domaine où la connaissance experte est dispersée, évolutive, incomplète et intuitive. ACASYA (Aide à la Certification par Apprentissage des SYstèmes de transport Automatisés) est le système logiciel développé en vue de prouver le bien fondé de cette méthodologie. Il a pour vocation d'aider les experts certifieurs de l'INRETS-CRESTA à apprécier le caractère sécuritaire d'un nouveau STA.

Dans le premier chapitre de ce mémoire, nous avons caractérisé le domaine de la certification et défini les objectifs de notre recherche. Cette étude a mis l'accent sur les différentes méthodes prévisionnelles d'analyse de sécurité employées par les experts certifieurs qui, malgré leur intérêt indéniable, ne permettent pas d'effectuer une analyse exhaustive. Afin de les renforcer, une approche a été proposée, centrée sur les techniques d'intelligence artificielle et notamment sur l'utilisation des systèmes à base de connaissances (SBC).

Le développement d'un SBC pour l'aide à la certification nécessite, en premier lieu, le recours aux techniques, méthodes et outils d'acquisition de connaissances étudiés dans le deuxième chapitre. Cet état de l'art a débouché sur le choix d'une méthode pour l'acquisition de connaissances qui, appliquée au domaine de la certification, s'est révélée efficace pour l'extraction et la formalisation des scénarios d'accidents mais insuffisante pour acquérir la démarche experte de certification, fortement intuitive.

Pour pallier cette difficulté, notre étude s'est orientée vers l'utilisation des techniques d'apprentissage automatique, détaillées dans le troisième chapitre. Le foisonnement des méthodes, techniques et systèmes d'apprentissage ainsi que l'ambiguïté du vocabulaire employé compliquent le choix de l'utilisateur d'un système adapté à son domaine d'étude. Pour simplifier ce choix, nous avons proposé une vision synthétique du domaine d'apprentissage. Cette approche, qui reste encore à améliorer, nous a cependant permis d'imposer à l'outil d'aide à la certification un ensemble de propriétés qui ont guidé le choix du système CHARADE et justifié le développement d'un nouveau système CLASCA. Ces deux systèmes se complètent pour la réalisation de l'outil logiciel ACASYA.

L'acquisition de connaissances et l'apprentissage automatique se sont révélés complémentaires pour l'élaboration d'une base de connaissances de certification. En effet, l'acquisition de connaissances a permis d'extraire et de formaliser les connaissances historiques d'analyse de sécurité, exploitées ensuite par un mécanisme d'apprentissage, en vue de produire des connaissances proches de la démarche de l'expert de certification, à l'origine difficilement exprimable par ce dernier. C'est sur l'utilisation conjointe de ces deux approches que repose notre méthodologie d'aide à la certification des STA, décrite dans le quatrième chapitre. Cette méthodologie est organisée de façon à reproduire, en grande partie, la stratégie utilisée par les experts certifieurs. C'est une démarche à trois niveaux, permettant d'aborder progressivement la résolution du

problème, tout en conservant une forte interactivité avec l'expert. Elle est fondée sur la complémentarité de deux modèles. Le premier modèle "générique" de représentation des scénarios d'accidents est du ressort de la phase d'acquisition de connaissances et s'articule autour de deux descriptions : une description statique qui permet de caractériser le contexte d'occurrence du scénario et une description dynamique, basée sur l'emploi des réseaux de Pétri, qui met en évidence les possibilités d'évolution dans ce contexte. Le deuxième modèle "implicite" de raisonnement de l'expert, qui fait appel aux techniques d'apprentissage automatique, est constitué de trois niveaux de traitement complémentaires : classification (CLASCA), évaluation (EVALSCA) et génération (GENESCA) des scénarios d'accidents. Vu l'ampleur du problème, la conception et la réalisation de la maquette de faisabilité du système ACASYA ont porté sur les deux premiers niveaux de classification et d'évaluation. CLASCA est un système d'apprentissage inductif, incrémental et interactif qui permet de caractériser un ensemble de situations d'insécurité vécues afin d'identifier la classe d'appartenance d'un nouveau scénario soumis à évaluation. EVALSCA a pour objectif de suggérer aux certifieurs d'éventuelles pannes non considérées par le constructeur et susceptibles de mettre en défaut la sécurité d'un nouveau STA. Il fait appel au système d'apprentissage CHARADE pour la production des règles et au générateur de systèmes experts IS2 pour la déduction des pannes.

Le dernier chapitre a présenté les résultats d'une première évaluation de la maquette d'ACASYA ainsi que ses principales perspectives d'évolution. En dépit de certaines limites inhérentes, notamment, à la non représentativité des connaissances acquises à ce jour et à la validité des règles produites par CHARADE, la maquette du système ACASYA a prouvé l'intérêt de la méthodologie pour formaliser, exploiter et pérenniser le savoir faire de l'expert certifieur, en vue de l'aider à rendre exhaustive l'analyse de sécurité. En outre, par opposition aux systèmes d'aide au diagnostic, ACASYA peut être perçu comme un outil d'aide à la prévention des pannes, situé en aval des méthodes prévisionnelles classiques d'analyse de sécurité. Afin d'être facilement adaptable à d'autres domaines tels que, par exemple, la certification de tous les transports en commun ou la sécurité des locaux publics, ACASYA requiert des améliorations et extensions portant, notamment, sur l'enrichissement de la base de connaissances, la génération de nouveaux scénarios d'accidents (module GENESCA) et le renforcement de la coopération Homme-Machine par l'intégration d'un module d'apprentissage par recherche d'explications. Le développement d'ACASYA se poursuit actuellement, dans le cadre d'une autre thèse, centrée sur la conception et la mise en œuvre du module GENESCA pour la génération des scénarios.

References

/AMPELAS 77/ : A. AMPELAS. "Quelques réflexions sur les études de sécurité des systèmes de transport en commun urbains". Recherche et développement, documentation - information, RATP Janvier-Fevrier-Mars 1977

/ANGELETTI 91/ : T. ANGELETTI, L. BRUNEAU, E. CHRISMENT. "Acquisition de connaissances multi-expertes de conception par la méthode KOD : Bilan, critiques et suggestions" - JAC - Sète 1991

/AUSSENAC 88/ : N. AUSSENAC, B. MICHEZ. "MACAO : Application d'un modèle psychologique à la réalisation d'un outil d'aide à l'acquisition des connaissances" - Colloque CRETS - Saint-Etienne - Mai 1988

/AUSSENAC 89/ : N. AUSSENAC. "Conception d'une méthodologie et d'un outil d'acquisition de connaissances expertes" - Thèse de Doctorat en Informatique - Université de Toulouse - 6 Octobre 1989 - 253p

/AUSSENAC 90/ : N. AUSSENAC, J.L. SOUBIE. "Place d'un outil d'acquisition des connaissances dans la conception des systèmes intelligents" - Journées Acquisition de connaissances - JAC 90 - Lannion 27 Avril 1990 - p115-129

/AUSSENAC 90/ : N. AUSSENAC. "Compte rendu EKAW 90 : current trends in Knowledge Acquisition". European Knowledge Acquisition Workshop, Amsterdam 25-29 Juin 1990

/AUSSENAC 91/ : N. AUSSENAC. "Acquisition des connaissances" - Bulletin n°5 de l'AFIA (Association Française pour l'Intelligence Artificielle) - Mars 1991

/AUSSENAC 91/ : N. AUSSENAC, J. P. KRIVINE. "Méthodes et outils pour l'acquisition des connaissances". Cours n° 2, Expert Systems & Their Application, May 27-31, Avignon 1991

/AUSSENAC 91/ : N. AUSSENAC. "Compte-rendu EKAW90 : "Current trends in Knowledge Acquisition" - JAC 91 - Sète 1991

/BARANOWSKI 90/ : F. BARANOWSKI. "Définition des objectifs de sécurité dans les transports terrestres". Rapport INRETS n° 133, Décembre 1990, 106p

/BARBET 88/ : J.F. BARBET, I. BERTRAND, J.G. LIGERON, M. LISSANDRE, R. VOGIN. "Techniques d'analyses fonctionnelles appliquées à la modélisation de la sûreté de fonctionnement des systèmes de production". 6ème colloque international de fiabilité et de maintenabilité,Strasbourg France 1988

/BEAUNE 90/ : P. BEAUNE, D. GRAILLOT. "Un serveur expert utilisant l'apprentissage automatique". ERGO IA, 19-21, Septembre 1990, p 294-303

/BEAUNE 92/ : P. BEAUNE. "Acquisition interactive de la connaissance par un système télématique dans le domaine de l'assainissement de l'eau". Thèse de Doctorat, Université de Montpellier II, Janvier 1992

/BENCHIMOL 86/ : G. BENCHIMOL, P. LEVINE, J.C. POMEROL. "Systèmes experts dans l'entreprise"

Paris, éditions HERMES, Mai 1986

/BENKIRANE 90/ : M. BENKIRANE, B. HOURIEZ, P. MILLOT. "Approche méthodologique d'extraction de connaissances : Application au domaine du diagnostic technique". 2ème congrès européen "Multi-Média, Intelligence Artificielle et formation". Applica 90, 24-26 Septembre 1990

/BENKIRANE 91/ : M. BENKIRANE. "Contribution à la méthodologie d'extraction de connaissances dans le domaine du diagnostic technique". Thèse de Doctorat, Université de Valenciennes, Mars 1991

/BICHINDARITZ 91/ : I. BICHINDARITZ, B. SEROUSSI. "ALEXIA : un système de résolution de problèmes à partir de cas utilisant une mémoire d'exemples méta-indexée par un modèle causal". 6èmes JFA 91, Sète 1991, p 5-25

/BISSON 89/ : G. BISSON. "APOGEE : Une utilisation de la saturation pour généraliser". 4èmes Journées Françaises de l'apprentissage. Saint-Malo 22-24 Mai 1989, p 31-44

/BISSON 89/ : G. BISSON, P. LAUBLET. "A functional model to evaluate Learning Systems". EWSL 1989, Montpellier p 37-48

/BISSON 91/ : G. BISSON. "KBG : a generator of knowledge bases". Proceedings of the International Conference Symbolic-Numeric, Data Analysis and Learning. Paris, september 17-20, 1991, p. 1-15

/BONNET 86/ : A. BONNET, J.P. HATON, J.M. NOC TRUONG. "Systèmes experts : vers la maîtrise technique". Paris-Inter éditions, 1986, 286p

/BOUCHON 90/ : B. BOUCHON-MEUNIER, S. DESPRES, D. DUBOIS, O. GASCUEL, A. GUENOCHE, H. PRADE. "Aspect de l'interface entre symbolique et numérique". Actes des 3èmes Journées Nationales, PRC-GDR, Intelligence Artificielle, 5-7 Mars 1990, CNIT Paris-La-Défense. Editions HERMES, p 89-111

/BOY 88/ : G. BOY, B. FALLER, J. SALLANTIN. "Acquisition et ratification de connaissances". Actes des Journées Nationales PRC-GRECO IA, Toulouse, 1988

/BREUKER 85/ : J. BREUKER, B. WIELINGA. "KADS : Structured Knowledge acquisition for expert systems" - Proceedings of 5th international Workshop on expert systems and their applications. Avignon, France 1985

/BREUKER 87/ : J. BREUKER, B. WIELINGA. "Use of models in interpreting verbal data". Reprint from : Alison L. Kidd (ed). Knowledge Acquisition for Expert Systems, a practical handbook . New York, Plenum Press, chapter 2, p. 17-44, 1987

/CALABRETTA 90/ : S. CALABRETTA. "ACKnowledge : vers l'amélioration du processus d'acquisition des connaissances" - Journées Acquisition des connaissances - JAC 90 - Lannion 27 Avril 1990 - p19-32

/CALVEZ 90/ : J.P. CALVEZ. "spécification et conception des systèmes : une méthodologie". Massons, Paris 1990, 614 p

/CORDIER 87/ : M.O. CORDIER. "Les systèmes experts" La recherche en Intelligence Artificielle, Edition SEUIL 1987 - p 177-210

/DAVID 88/ : Y. DAVID. "L'évolution des méthodes de certification de la sécurité face au développement des applications de la microinformatique dans les transports terrestres". Annales des Ponts et Chaussées - 1er trimestre 1988

/DAVID 90/ : Y. DAVID. "Impact des Microprocessus sur la philosophie de la sécurité des transports terrestres". EAED : European Association for Electricol Driving. Bolzano 14-17 Mai 1990, Italy

/DECAESTECKER 89/ : C. DECAESTECKER. "Formation incrémentale de concepts par un critère d'adéquation". 4èmes Journées Françaises de l'apprentissage. Saint-Malo 22-24 Mai 1989, p 63-77

/DECAESTECKER 89/ : C. DECAESTECKER. "Incremental Concept Formation with attribute selection". EWSL : European Working Session on Learning 1989, p 49-58

/DE CARVALHO 90/ : F.A.T. DE CARVALHO, J. LEBBE, R. VIGNES, E. DIDAY". Dissimilarité en analyse des données symboliques". JFA 90, Lannion 25-26 Avril 1990, p 47-75

/DERENTY 92/ : V. DERENTY. "Réalisation d'une interface Homme-Machine pour l'exploitation de deux mécanismes d'apprentissage symbolique automatique : CLASCA, EVALSCA". Rapport de projet DESS - ICHM. Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Novembre1991 / Mars 1992

/DIDAY 89/ : E. DIDAY. "Introduction à l'analyse des données symboliques : objets symboliques modaux et implicites". Actes des 2èmes Journées symboliques-numériques pour l'apprentissage de connaissances à partir d'observation. Université de Paris-Sud, LRI. Edités par E. DIDAY et Y. KODRATOFF, 1989, p 1-30

/DIDAY 89/ : E. DIDAY. "Introduction à l'analyse des données symboliques". Rapport de recherche n° 1074, INRIA, Août 1989, p 1-37

/DIENG 90/ : R. DIENG. "Méthodes et outils d'acquisition des connaissances" - ERGO IA 90 - 19-21 Septembre 1990, P245-271

/DUPAS 90/ : R. DUPAS. "Apport des méthodes d'apprentissage symbolique automatique pour l'aide à la maintenance industrielle". Thèse de Doctorat, Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Janvier 1990

/DUVAL 91/ : B. DUVAL. "Abduction pour l'apprentissage à partir d'explication" JFA 91, 6èmes Journées Françaises de l'Apprentissage, Sète 1991, p 25-45

/EL KOURSI 90/ : M. EL KOURSI, A. STUPARU. "Etude comparative des architectures

microprogrammées utilisées dans les applications de sécurité". Rapport INRETS n° 134, Décembre 1990, 79p

/EL KOURSI 92/ : M. EL KOURSI, B. LE TRUNG, H. HADJ-MABROUK, B. HOURIEZ. "Base de connaissances pour l'aide à l'analyse de sécurité des systèmes de transports terrestres automatisés". 8ème colloque de fiabilité et de maintenabilité, Grenoble 6-8 Octobre 1992

/FARRENY 85/ : H. FARRENY. "Les systèmes experts : principes et exemples". Editions CEPADUS 1985 - 363p

/FARRENY 87/ : H. FARRENY, M. GHALLAB. "Eléments d'Intelligence Artificielle". Paris - HERMES 1987, 363p

/FAYOLLE 90/ : B. FAYOLLE. "Spécification et validation d'automatismes à haut niveau de sûreté de fonctionnement appliquées aux transports terrestres automatisés". Thèse de Doctorat, Université de Lille, 24 Avril 1990

/FISHER 87/ : D.H. FISHER. "Knowledge Acquisition via Incremental Conceptual Clustering. Machine Learning, vol 2, 1987, p 139-172

/GABER 91/ : K. GABER. "Contribution à la réalisation et à l'évaluation d'un système d'apprentissage par classification : CLASCA". Rapport de DEA, Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Juillet 1992

/GABILLARD 79/ : R. GABILLARD, M. FICHEUR. "Sur les rôles respectifs de l'homme et du matériel dans la sécurité d'un système de transport". Annales des Ponts et Chaussées, 2ème trimestre 1979

/GAINES 87/ : B. GAINES. "An overview of knowledge Acquisition and Transfer". International Journal Man-Machine Studies 265, p 453-472, 1987

/GALLOUIN 88/ : J.F. GALLOUIN. "Transfert de connaissances - systèmes experts : techniques et méthodes" Editions EYROLLES, 1988, 159p

/GANASCIA 85/ : J.G. GANASCIA. "La conception des systèmes experts". La recherche en Intelligence Artificielle n°170, Octobre 1985

/GANASCIA 85/ : J.G. GANASCIA. "L'apprentissage dans les systèmes experts". ARTEFACT, Microsystèmes, Novembre 1985, p 146-151

/GANASCIA 87/ : J.G. GANASCIA. "AGAPE : de l'appariement structurel à l'apprentissage". Intellectica : Apprentissage et Machine (1), n° 2/3, 1987

/GANASCIA 87/ : J.G. GANASCIA. "CHARADE : apprentissage de bases de connaissances". Journées Symboliques numériques pour l'apprentissage de connaissances à partir de données. Université Paris Dauphine 8-9 Décembre 1987

/GANASCIA 87/ : J.G. GANASCIA. "AGAPE et CHARADE : deux mécanismes d'apprentissage symbolique appliqués à la construction de bases de connaissances". Thèse d'état - Université Paris-sud, 27 Mai 1987

/GANASCIA 88/ : J.G. GANASCIA. "CHARADE : Une sémantique cognitive pour les heuristiques d'apprentissage". 8èmes Journées internationales : les systèmes experts et leurs applications. Avignon, Mai 1988, p 567-585

/GANASCIA 88/ : J.G. GANASCIA, N. HELFT. "Evaluation des systèmes d'apprentissage", 3èmes Journées Françaises de l'apprentissage. Cassis, 5-6 Mai 1988, p 3-20

/GANASCIA 89/ : J.G. GANASCIA, J.F. PUGET. "Modèle comportemental d'apprentissage : représentation de systèmes existants". Actes des 4èmes Journées Françaises de l'Apprentissage. St Malo, mai 1989

/GANASCIA 90/ : J.G. GANASCIA. "L'âme Machine - les enjeux de l'Intelligence Artificielle". Editions du SEUIL, Janvier 1990, 280p

/GANASCIA 90/ : J.G. GANASCIA, J.F. PUGET, N. HELFT. "Comportement des systèmes d'apprentissage : vers une modélisation générale". Actes des 3èmes Journées Nationales PRC-GDR Intelligence Artificielle, 5-7 Mars 1990, Edition HERMES, Chapitre 5, p 237-269

/GANASCIA 91/ : J.G. GANASCIA. "L'apprentissage symbolique". Encyclopédie de la communication - PUF, sous la direction de Lucien SFEZ, 1991

/GENNARI 89/ : J.H. GENNARI, P. LANGLEY, D. FISHER. "Models of Incremental Concept Formation". Artificial Intelligence, vol 40, 1989, p 11-61

/GRUNDSTEIN 88/ : M. GRUNDSTEIN, P. de BONNIERES, S. PARA. "Les systèmes à base de connaissances, systèmes experts pour l'entreprise". AFNOR 1988, 261p

/HADJ-MABROUK 89/ : H. HADJ-MABROUK, R. DUPAS, P. MILLOT. "Outil d'aide à la maintenance préventive par apprentissage". FORUM Intelligence Artificielle et Systèmes Experts, Paris 13-15 Décembre 1989

/HADJ-MABROUK 91a/ : H. HADJ-MABROUK, M. EL KOURSI, B. HOURIEZ, P. MILLOT. "Approche méthodologique d'aide à la certification des systèmes de transport automatisés basée sur l'apprentissage". Journées Acquisition de Connaissances (JAC), Sète 17 Mai 1991

/HADJ-MABROUK 91b/ : H. HADJ-MABROUK, B. HOURIEZ, P. MILLOT. "Etude de faisabilité d'un système à base de connaissances pour l'aide à la certification des systèmes de transport automatisés". Convention INRETS-CRESTA/LAIH du 15/01/91. Rapport de fin de Contrat, Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Juin 1991

/HADJ-MABROUK 92a/ : H. HADJ-MABROUK, B. HOURIEZ, M. EL KOURSI, B. LE TRUNG. "Méthodologie d'analyse et d'évaluation de la sécurité basée sur les techniques d'intelligence artificielle". Revue européenne de diagnostic et sûreté de fonctionnement. Editions Hermes 1992, volume 2 - n° 1/1992, p 5-35

/HADJ-MABROUK 92b/ : H. HADJ-MABROUK, B. HOURIEZ. "Acquisition de connaissances et apprentissage automatique pour l'élaboration d'une base de connaissances". Actes des 1ères Journées Francophones d'Apprentissage et d'Explicitation des Connaissances (JFAEC). AFIA-AFCET, PRC GRECO IA, Dourdan, 15-17 Avril 1992, p 29-46

/HADJ-MABROUK 92c/ : H. HADJ-MABROUK, L. MEJRI, B. HOURIEZ. "Complémentarité des deux mécanismes d'apprentissage : CLASCA et CHARADE pour le développement d'un système à base de connaissances". 3èmes Journées symbolique-numérique. SFC-AFCET-AFIA, Paris 14-15 Mai 1992, p 157-170

/HADJ-MABROUK 92d/ : H. HADJ-MABROUK, L. MEJRI, M. EL KOURSI, B. HOURIEZ. "Système expert à apprentissage pour l'aide à l'analyse de sécurité. Application à la certification des systèmes de transport automatisés". 12èmes Journées Internationales d'Avignon 92. Conférence Intelligence Artificielle, Défense et Sécurité civile. Avignon 2-3 Juin 1992, volume 2, p 97-112

/HADJ-MABROUK 92e/ : H. HADJ-MABROUK, L. MEJRI, M. EL KOURSI, B. HOURIEZ. "Méthodologie d'aide à l'évaluation de la sécurité des systèmes de transport automatisés, basée sur l'apprentissage automatique". Forum 1992 de la SCGM (Sté Canadienne de Génie Mécanique), "TRANSPORT 1992+" Université Concordia (session : Sécurité dans les transports), Montréal, Québec, Canada 1-4 Juin 1992, volume III, p 957-964

/HADJ-MABROUK 92f/ : H. HADJ-MABROUK, B. HOURIEZ, P. MILLOT. "CLASCA et EVALSCA deux mécanismes d'apprentissage Symbolique-Numérique pour le développement d'un système à base de connaissances d'aide à la certification des systèmes de transport automatisés". Convention INRETS-CRESTA/VALUVAL-LAIH du 22/01/91. Rapport de fin de Contrat, Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Juin 1992

/HART 88/ : A. HART. "Acquisition du savoir pour les systèmes experts". Masson Paris 1988, 142p

/HOLSAPPLE 87/ : C.W. HOLSAPPLE, A.B. WHINSTON. "GURU, l'utilisation des systèmes experts dans l'entreprise". Les éditions d'organisation, Paris 1987, 363 p

/IGL Technology 89/ : IGL Technology . "SADT : un langage pour communiquer". Editeur Eyrolles, Paris 1998, 336p

/KODRATOFF 85/ : Y. KODRATOFF. "Quand l'ordinateur apprend". La recherche n° 170, Volume 16, Octobre 1985, p 1252-1262

/KODRATOFF 86/ : Y. KODRATOFF. "Leçons d'apprentissage symbolique automatique". Cepadues - Edition, 1986

/KODRATOFF 87/ : Y. KODRATOFF. "Problématique et applications de l'apprentissage symbolique automatique". Séminaire en Intelligence Artificielle et circulation aérienne. Ecole Nationale de l'Aviation Civile, Avril 1987, p 127-140

/KODRATOFF 89/ : Y. KODRATOFF. "Apprentissage : science des explications ou sciences des nombres ?" Ann. Télécom., 44, n° 5-6, 1989, p 251-264

/KODRATOFF 89/ : Y. KODRATOFF. "Présentation des travaux de recherche de l'équipe LRI : Laboratoire de Recherche en Informatique de l'Université de Paris-Sud : responsable KODRATOFF 4èmes JFA 89, Saint-Malo 22-24 Mai 1989, p 25-27

/KODRATOFF 91/ : Y. KODRATOFF, E. DIDAY. "Induction symbolique et numérique à partir de données". CEPADUES-EDITIONS 1991. 460p

/LAPRIE 85/ : J.C. LAPRIE. "Sûreté de fonctionnement des systèmes informatiques et tolérance aux fautes: concepts de base". Technique et Science Informatique, Volume 4 n°5, 1985, p 419-429

/LAPRIE 87/ : J.C. LAPRIE. "Evaluation de la sûreté de fonctionnement des systèmes informatiques : Matériel et logiciel". Derniers développements en automatique, informatique, robotique, micro-électronique. Le séminaire du LAAS-CNRS, Cepadues-Editions, Octobre 1987, p 141-157

/LAPRIE 88/ : J.C. LAPRIE. "Sûreté de fonctionnement et tolérance aux fautes : concepts de base". LAAS-CNRS, les techniques de l'Ingénieur, Collection Automatique et Informatique, Novembre 1988

/LAPRIE 89/ : J.C. LAPRIE, B. COURTOIS, M.C. GAUDEL, D. POWELL. "Sûreté de fonctionnement des systèmes informatiques : Matériel et logiciel". Bordas, Dunod, Paris 1989, 184 p

/LAURIERE 84/ : J.L. LAURIERE. "Les systèmes experts : caractéristique, état de l'art et perspective". Colloque international d'intelligence artificielle. Marseille 24-27 Octobre 1984

/LAURIERE 87/ : J.L. LAURIERE. "L'apprentissage", chapitre IX, Edition EYROLLES 1987

/LEBOWITZ 87/ : M. LEBOWITZ. "Experiments with Incremental Concept Formation : UNIMEM". Machine Learning, vol 2, 1987, p 103-138

/LERMAN 90/ : I.C. LERMAN, J. NICOLAS, M. OUALI, P. PETER. "Classification conceptuelle : une approche centrée sur la similarité"

/LESCORT 85/ : A. LESCORT. "Intelligence Artificielle et Systèmes Experts". Edition Cedic/Nathan 1985, 141 p

/LE TRUNG 84/ : B. LE TRUNG. "Procédures d'agrément des automatismes des nouveaux systèmes de

métro". RTS : Recherche Transports Sécurité - Juillet 1984

/LE TRUNG 89/: B. LE TRUNG. "Approval procedure for automatic equipement of unmanned Metro system". IFAC - CCCT 1989, Septembre 19-21, 1989 Paris

/LE TRUNG 92/: B. LE TRUNG, M. EL KOURSI, B. HOURIEZ, H. HADJ-MABROUK. " Knowledge base for safety analysis in unmanned metro system". COMPRAIL 92, 18-20 August 1992, WASHINGTON DC, USA

/LIEVENS 76/ : C. LIEVENS. "Sécurité des systèmes". E.N.S.A.E., Toulouse, 1976, 352p

/LIQUIERE 89/ : M. LIQUIERE. "INNE : un algorithme d'apprentissage de structures dans le cas d'exemples bruités". 4émes Journées Françaises de l'apprentissage (JFA 89). Saint-Malo 22-24 Mai 1989, p 93-105

/LOUNIS 91/ : H. LOUNIS, G. BISSON. "Etude des performances des algorithmes d'apprentissage : apport de bases artificielles". 6èmes JFA 91, Sète 1991, p 145-164

/MAHE 87/ : H. MAHE, P. VESOUL. "Acquisition des connaissances et adaptation à l'utilisateur : outils et méthodes". 7èmes journées internationales d'Avignon. Les systèmes experts et leurs applications. Avignon 13-15 Mai 1987, Volume 1, p 625-646

/MANAGO 88/ : M. MANAGO. "Intégration de techniques numériques et symboliques en apprentissage automatique". Thèse de Docteur en Science, Université de Paris-Sud, Centre d'Orsay, 20 Décembre 1988

/MANAGO 89/ : M. MANAGO, N. CONRUYT. "KATE un système d'apprentissage avec objets". 4èmes Journées Françaises de l'apprentissage : JFA 89. Saint-Malo 22-24 Mai 1989, p 121-135

/MARCOVICI 82/ : C. MARCOVICI, B. LE TRUNG. "Méthodologie sécurité appliquée au métro de Lille". Proceedings of the 3rd International Conference on Reliability & Maintainability, 1982

/MEJRI 91/ : L.MEJRI. "Réalisation d'une maquette de faisabilité d'un système d'apprentissage de descriptions de classes de scénarios d'accident pour l'aide à la certification des systèmes de transport automatisés". Rapport de DEA, Laboratoire d'Automatique Industrielle et Humaine, Université de Valenciennes, Juin 1991

/MEJRI 92/ : L. MEJRI, H. HADJ MABROUK, M. EL KOURSI, B. HOURIEZ. "Learning based assistance tool for the generation of scenarios for automated transport systems certification". 11th European Annual Conference on Human decision making and manual control. Valenciennes, France, 17-19 november 1992

/MICHALSKI 80/ : R.S. MICHALSKI, R.L. Chilausky. "Knowledge acquisition by encoding expert rules versus computer induction from examples : A case study involving soyben pathology". Internatinal Journal of Man-Machine Studies, 12, P 63-87, 1980

/MICHALSKI 83/ : R.S. MICHALSKI, R.E. STEPP. "Learning from observation : conceptual clustering". Machine learning : An Artificial Intelligence Approach. R.S. MICHALSKI, J.G. Carbonell, T.M. Mitchell. Vol. 1, Morgan Kaufmann, P 331-363, 1983

/MICHALSKI 84/ : R.S. MICHALSKI. "Inductive Learning as Rule-guide Transformation of Symbolic descriptions : a Theory and Implementation". Automatic Program Construction Techniques. A.W. BIERMANN, G. GUIHO, Y. KODRATOFF, Macmillan Publishing company, p. 517-552, 1984

/MICHALSKI 89/ : R.S. MICHALSKI. "Introduction to Machine Learning". Machine Learning : An Artificial Intelligence Approach, Volume III, Y. KODRATOFF & R.S. MICHALSKI. Morgan KAUFMANN 1989

/MILLOT 88/ : P. MILLOT ."Supervision des procédés automatisés et ergonomie". Editions Hermes, Paris, décembre 1988

/MILLOT 90/ : P. MILLOT ."Coopération Homme-Machine : exemple de la téléopération". Journées du G.R. Automatique, Strasbourg 17-19 octobre 1990

/MITCHELL 82/ : T.M. MITCHELL. "Generalization as search". Artificial Intelligence n°18, 1982

/MITCHELL 83/ : T.M. MITCHELL. "Learning and problem solving". Proceeding IJCAI-83, Karlsruhe, p. 1139-1151, 1983

/NASSIET 87/ : D. NASSIET. "Contribution à la méthodologie de développement des systèmes experts : application au domaine du diagnostic technique". Thèse de Docteur Ingénieur, Université de Valenciennes, 20 Novembre 1987

/PARODI 90/ : A. PARODI, S. KHOVAS. "Une approche de raisonnement et d'apprentissage sur un modèle de représentation explicite des connaissances". 5èmes Journées Françaises de l'apprentissage JFA 90, Lannion 25-26 Avril 1990, p 233-249

/PIERREVAL 90/ : H. PIERREVAL. "Les méthodes d'analyse et de conception des systèmes de production". Edition Hermès, Paris 1990

/POMORSKI 91/ : D. POMORSKI. "Apprentissage automatique symbolique-numérique : construction et évaluation d'un ensemble de règles à partir des données". Thèse de Doctorat, Université de Lille, Décembre 1991

/QUILAN 83/ : J.R. QUILAN. "Learning Efficient Classification Procedures and their Applications to Chess and Games". Machine Learning : An Artificial Intelligence Approach, R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds), Morgan Kaufmann, 1983, PP 463-482

/QUILAN 86/: J.R. QUILAN. "Induction of Decision Trees". Machine Learning 1, p. 81-106, 1986

/RIALLE 90/: V. RIALLE. "Contribution à la définition du profil de cogniticien : caractéristiques fonctionnelles, savoir et savoir-faire dans l'élaboration d'une base de connaissances". Applica 90, 2ème congrès Européen. Lille 24-26 Septembre 1990

/ROGER 91/ : B. ROGER. "Calcul incrémental de caractérisations disjonctives de concepts". 6èmes JFA 91, Sète 1991, p 88-112

/ROOK 89/: F.W. ROOK, J.W. CROGHAN. "The Knowledge Acquisition Activity Matrix : A systems Engineering Conceptual Framework". IEEE Transactions on systems, Mand, and cybernetics. vol 19, n°3, May, June 1989, p 586-597

/ROYER 90/ : J.C. ROYER, A. MERLE, C de SAINTE MARIE. "Application de l'apprentissage au problème de réglage en contrôle non destructif". JAC 90, Lannion 27 Avril 1990, p 5-18

/SEBAG 90/ : M. SEBAG. "Une approche symbolique-numérique pour la discrimination à partir d'exemples et de règles : l'apprentissage multi-couches". Thèse de Docteur en Science, Université Paris IX Dauphine, U.E.R. Mathématique de la décision. 1990

/SEBAG 90/ : M. SEBAG, M. SCHOENAUER. "Apprentissage de règles par découverte". 5èmes Journées Françaises de l'Apprentissage : JFA 90, Lannion 25-26 Avril 1990, p 213-231

/SEROUSSI 89/ : B. SEROUSSI, J.J. BOISVIEUX. "Apprentissage et raisonnement par analogie dans le système ALEX". 4èmes JFA, Saint-Malo 22-24 Mai 1989, p 169-181

/STAROSWIECKI 91/ : M. STAROSWIECKI, M. BARBOUCHA. "Apprentissage automatique par filtrage de tableaux de contingence". Induction symbolique et numérique à partir de données, volume I. CEPADUES-EDITIONS 1991, p 405-434

/THAYSE 91/ : A. THAYSE & co-auteurs. "De l'apprentissage artificiel aux frontières de l'IA". Approche logique de l'Intelligence Artificielle, Volume 4, chapitre 1, Edition DUNOD 1991

/THIBAULT 87/ : H.B. THIBAULT, J.F. DHALLUIN. "Sur la réglementation de la sécurité des transports nouveaux par objectifs quantifiés". Recherche Transports Sécurité, Mars 1987

/TONG 91/ : X. TONG. "Modèles de connaissances et leur acquisition". JAC, Sète 1991

/VILLEMEUR 88/ : A. VILLEMEUR. "Sûreté de fonctionnement des systèmes industriels - fiabilité, facteurs humains, informatisation". Edition Eyrolles Paris 1988, 795 p

/VISSER 88/ : W. VISSER. "Méthodes pour l'accès à l'expertise. L'utilisation concurrente de différentes méthodes de recueil de données pour l'étude de l'activité de programmation". Psychologie Française 33.3, Novembre 1988, p 127-132

/VISSER 90/ : W. VISSER. "Acquisition de connaissances : l'approche de la psychologie cognitive illustrée par le recueil d'expertise en conception". Journée Acquisition de Connaissances (JAC), Lannion 27, Avril 1990

/VOGEL 88/ : C. VOGEL. "Génie cognitif". Collection sciences cognitives, Edition Masson, Mai 1988

/VOGEL 89/ : C. VOGEL. "Comment qualifier un système à base de connaissances ?". Génie logiciel et système expert - n° 6, Septembre 1989

/VRAIN 87/ : C. VRAIN. "Un outil pour la généralisation utilisant systématiquement les théorèmes : le système OGUST". Thèse de Docteur de troisième cycle en Informatique. Université de Paris-Sud, Centre d'Orsay, 25 Février 1987