



HAL
open science

Resilience, Reliability, and Coordination in Autonomous Multi-Agent Systems

Rafael Cardoso, Brian Logan, Felipe Meneguzzi, Nir Oren, Bruno Yun

► **To cite this version:**

Rafael Cardoso, Brian Logan, Felipe Meneguzzi, Nir Oren, Bruno Yun. Resilience, Reliability, and Coordination in Autonomous Multi-Agent Systems. *AI Communications*, 2022, 35 (4), pp.339-356. 10.3233/AIC-220136 . hal-04318342

HAL Id: hal-04318342

<https://hal.science/hal-04318342>

Submitted on 1 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Resilience, Reliability, and Coordination in Autonomous Multi-Agent Systems

R. C. Cardoso B. Logan F. Meneguzzi N. Oren
B. Yun

{rafael.cardoso,brian.logan,felipe.meneguzz,n.oren,bruno.yun}@abdn.ac.uk

Abstract

Multi-agent systems is an evolving discipline that encompasses many different branches of research. The long-standing Agents at Aberdeen (A^3) group undertakes research across several areas of multi-agent systems, focusing in particular on aspects related to resilience, reliability, and coordination. In this article we introduce the group and highlight past research successes in those themes, building a picture of the strengths within the group. We close the paper outlining the future direction of the group and identify key open challenges and our vision towards solving them.

1 Introduction

Agents research at Aberdeen has a long history, with work ranging from fundamental theoretical research on topics such as trust [1, 2], norms [3, 4] and argumentation [5], to applications of such technologies [6, 7, 8]. The *Agents at Aberdeen* group¹ (A^3) has changed over time, with several members joining the group in the last two years. There are currently 5 academics working on research related to agents, and whose interests cover multiple areas.² The focus of the group revolves around ensuring that agents not only work well with each other, but also behave predictably when working as part of a human-agent team and that such agents can cope with unexpected changes in their environment. The first strand of research aims at enhancing the *reliability* of a Multi-Agent System (MAS), and improving its *resilience* in the face of a dynamic and changing environment. The second strand of research revolves around *coordination* among agents, as well as within human-agent teams.

Resilience determines the ability to detect and deal (recover) with problems that may arise during execution (failures, unforeseen circumstances, etc.). Autonomous systems require resilience due to constant changes in the environment. There are many techniques for increasing resilience, such as runtime monitoring to detect failures [9, 10], the use of norms and rules to design compliant systems [11, 12], and automated planning to attempt to recover from failures [13, 14]. Due to the increasing ubiquity of autonomous systems, there has been significant interest in making such systems more resilient. For example, the Trustworthy Autonomous Systems Hub [15], is an initiative funded by UKRI which includes an element focused on resilience.

A reliable system is a system that performs as expected. Reliability is critical when decisions made autonomously by a system are mission or safety critical, or may affect humans. A recent Dagstuhl seminar [16] brought together researchers working on software engineering, verification, ethics, and machine learning to discuss reliability in MAS, and highlighted the topic's importance. One of the outcomes of the seminar was a special issue [17] that posed two challenges: defining what is a good decision, and how to check that the system will choose to make good decisions. There are many ways of improving the reliability of a system and to try to tackle these challenges. For autonomous systems, the techniques often include verification and validation, ranging from formal verification to simulation-based and physical-based testing [18]. Reliability

¹<https://agentsataberdeen.github.io/> (Accessed: 29/07/2022)

²In this paper, we focus on the work of the current members of the group rather than taking a historical perspective.

is somewhat similar to resilience since both deal with increasing the trustworthiness of a system, and the two concepts are therefore closely interconnected.

Coordination in a MAS concerns the coordination of individual and joint goals/actions of multiple agents. Despite this simple concept, coordination can be achieved using a multitude of different techniques from MAS. For example: agents can engage in argumentation to deliberate about how and when to coordinate and how to avoid conflicts arising from such coordination [19]; a trust system can be used to determine potential agents for coordination [20]; norms can be specified to guide agent coordination in order to avoid conflicts or improve efficiency [4]; multi-agent planning can solve scheduling problems in coordination as well as generate decentralised plans and/or centralised joint plans [14]; and using goal and plan recognition to reduce the need for explicit communication during coordination [21].

In this paper, we describe recent work done by A^3 in resilience, reliability, and coordination in the context of multi-agent systems. Figure 1 provides a general overview of the different research topics that the members of A^3 have been working on recently. While this word cloud illustrates some common topics of research among the group, it does not cover the entirety of the A^3 's research portfolio.

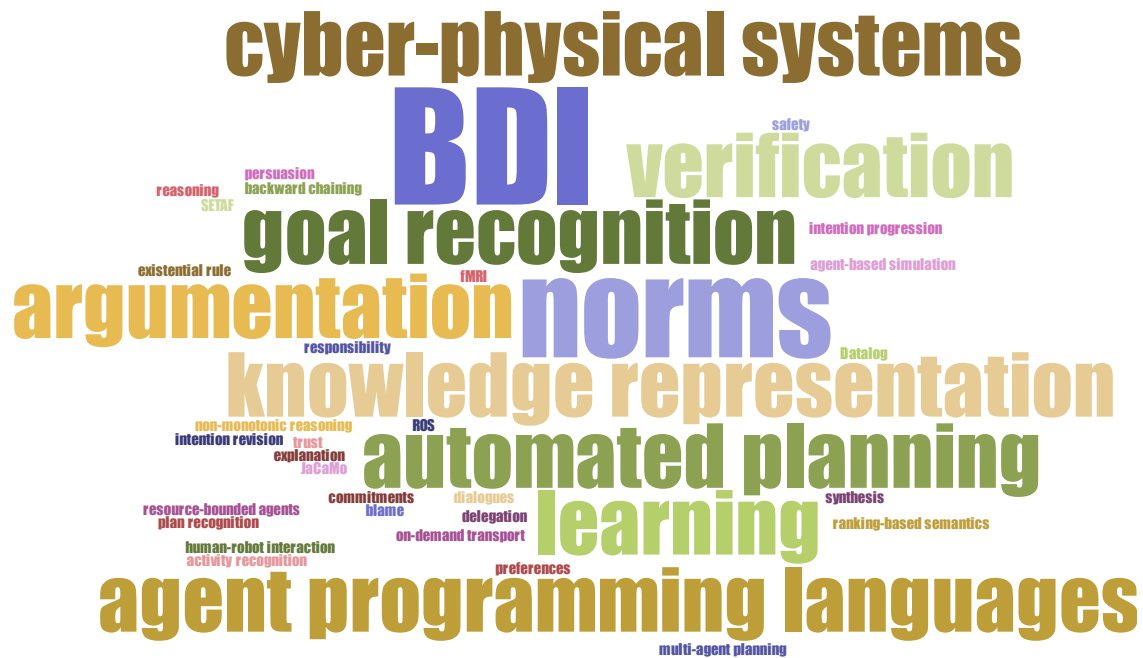


Figure 1: Word cloud containing the most common keywords in recent publications by the members of A^3 (2019–2022).

This paper is organised as follows. In Section 2 we list our combined efforts towards resilience and reliability in autonomous systems. Section 3 briefly contextualises coordination in MAS and describes the research that we have done in various topics in this area. In Section 4 we discuss open research challenges in MAS including, but not limited to resilience, reliability, and coordination. Finally, Section 5 summarises the contributions of the group so far and highlights future work.

2 Resilient & Reliable Autonomy

We use the word *resilience* to express the ability of an autonomous agent to withstand several types of failure while remaining functional (to some degree) and ensuring a minimum level of performance. The notion of reliability is more general than resilience as the agent must have consistently good performance in

all situations. Reliability is also a prerequisite for a system to be *trusted* by other agents/users. While trust has many definitions [22] and is influenced by factors such as reputation, risk, experience, and transparency among others; the members of the A^3 have examined the transparency aspect by providing insightful explanations [7] and visualisations [23], examined how trust can be used to underpin decision-making [6], and how trust can be computed [2]. In what follows we describe our current work to achieve resilient and reliable autonomy in MAS.

2.1 Formal Verification of Autonomous Decision-Making

One way of providing stakeholders with assurance in the reliability of a system is verification. Formal verification uses formal methods to prove specific properties (often related to requirements) of system behaviour. Examples of formal methods include *model checking* [24], where the desired property is specified in an appropriate logic and a model checking algorithm is used to determine whether the property holds for a model of the system; and runtime verification (RV) [25], which uses monitors to collect runtime events and then checks formal properties against these events. Many autonomous systems and robotic applications, particularly in safety-critical domains [26], require some form of formal verification to assure stakeholders that unsafe behaviours are not possible (or at least very unlikely) to occur, and/or that system objectives will be achieved. This is evidenced by, for example, the overview of verification challenges for inspection robots reported in [18] which also describes the common issues encountered in verifying remote inspection tasks during the authors' experience in three research hubs within the UK's "Robots for a Safer World" programme and which included involvement by members of A^3 . Verification research in the A^3 group has focused in particular on the verification of *resource-bounded* agents and multi-agent systems, where the behaviour of the agents depends on the resources they have available (e.g., time, money, battery power). Initial work considered the verification of the time, space and communication requirements of agent and multi-agent programs, e.g., [27, 28, 29], and the strategic ability of coalitions of agents under bounded resources [30, 31, 32, 33], including the first work to consider systems in which the amount of resources agents are able to produce is not known in advance [34, 35, 36, 37, 38]. More recently, we have investigated the problem of determining the minimal amount of resources required to achieve a goal [39] rather than whether a given amount of resource is sufficient to achieve a goal.

Another challenge in verification of agents is modelling the environment due to its dynamic nature, which often results in an incomplete abstraction of the real environment. Therefore, a static verification of the system using this abstraction of the environment does not provide sufficient assurances about the real reliability of the system. In [40], Ferrando, Dennis, and Cardoso et al. introduce a domain specific language to model the environment in such a way that a model checker can statically verify the environment. This language also allows the compilation of runtime monitors that can verify the same constraints at runtime, effectively checking if the environment abstraction holds during execution.

Members of the A^3 group have also developed several tools and applications related to formal verification. For example, in [32, 41, 33], an extension to the MCMAS [42] model checker is described that allows verification of properties of resource-bounded multi-agent systems. More recently, Ferrando and Cardoso et al. developed ROSMonitoring [9], a tool for performing RV of applications designed in the Robot Operating System (ROS)³. ROS is a middleware for developing software for robots. Its main advantages are modularity, interoperability, and range of tools and library packages containing drivers and state-of-the-art algorithms for robotic tasks such as navigation. ROSMonitoring allows the user to specify formal properties in a variety of formal logics, automatically synthesising the monitor based on a configuration file to be ready to deploy in a ROS application. The monitor then observes events at runtime (or from a log file) and validates it against its formal properties. Ferrando and Cardoso also developed the RVPlan [10] tool to automatically synthesise runtime monitors from assumptions in classical planning. Plans found by a classical planner can often become outdated when executing in dynamic environments and may be prone to failure. RVPlan converts preconditions of planning operators and instantiated actions into properties that can be verified at runtime in order to detect plan failures. Finally, in [43], Cardoso et al. apply heterogeneous verification

³<https://www.ros.org/> (Accessed: 03/05/2022)

techniques, such as RV, model checking, and theorem proving, to verify different nodes in a ROS 3D-based simulation of an autonomous Mars Curiosity rover.

2.2 Automated Planning and Synthesis

Formal verification is a post-design (or in the case of runtime verification post-development) approach, in that it checks properties of an existing agent design, program or system. If the desired property does not hold, the system design or implementation must be revised by a developer and re-verified in an iterative process. An alternative approach is to derive a (correct) behaviour, plan, program or policy from the desired property together with a model of the environment and the information available to an agent. Such approaches differ from learning-based model-free techniques in three key ways [44]. First, they require no data in order to decide valid actions in any given environment. Second, they provide formal guarantees about the resulting policies that learning-based approaches cannot. Finally, while no data is required, the (offline or online) computational cost of deriving a plan or program may be high. Similarly, planning differs from preprogrammed policies (e.g., using agent programming languages) in that it relieves the burden on the designer in accounting for every possible contingency imposed by the environment.

Work in the A^3 group has focused on two main approaches to the automated derivation of behaviour: *AI behaviour composition* [45] where provably correct behaviours or programs are synthesised from a formal specification of a task; and *automated planning*, a class of AI algorithms concerned with automatically generating plans or policies of action using a model of the environment and the information available to an agent.

In terms of AI behaviour composition, A^3 members initiated a new area of research in the synthesis of provably correct programs for the control of autonomous cyber-physical systems, such as smart manufacturing systems consisting of multiple interacting robots [46, 47, 48, 49, 50], and whether a controller can be synthesised given sufficient resources [51]. In the unbounded setting, only the types of available manufacturing resources are given, and we want to know whether it is possible to manufacture a product using only resources of those types, and, if so, how many resources of each type are needed. This work has led to the development of synthesis tools capable of generating industry-standard control software for manufacturing assembly cells [52, 53, 50].

Members of the A^3 group have also generated a number of contributions in automated planning modelling techniques and formalism extensions. First, Maguaguagno, Meneguzzi, and others [54, 55] developed a series of tools for developing, testing, and visualising planning domain descriptions for classical planning. These tools have been recognised not only as useful pedagogical tools, but also as informative mechanisms for visualising heuristic informativeness based on the design of each planning domain. Second, Maguaguagno and Meneguzzi [56] introduced the notion of *semantic attachments* to Hierarchical Task Network (HTN) planning languages, formalising the connection between the syntax of predicates reasoned about by the planning system, and dynamics external to the planner. Such connection is critical for planning in symbolic-geometric domains of the type required by robotics applications [57].

While most research in the past 20 years on automated planning has focused on algorithms for state-space search using formalisms for domain-independent planning, recent efforts [58] have revived interest in HTN planning, leading to new algorithms and heuristics. By contrast, members of our group have leveraged techniques from BDI agent reasoning cycles to substantially optimise HTN planning, leading to the development of HyperTensioN [59]. This planner won the 2020 International Planning Competition for Hierarchical Planning against contenders from multiple research groups. Meneguzzi and others have also made substantial contributions to planning using reinforcement learning and imitation learning achieving a high degree of sample efficiency [60].

Finally, we have consistently made contributions to applications of planning formalisms and algorithms. Specifically, Meneguzzi and others have applied planning techniques to a number of domains, both within computer science, and in transdisciplinary research. First, Meneguzzi et al. automatically generate and validate agent commitment protocols [61, 62, 63] using an HTN planning formalism. Second, Meneguzzi et al. carried out a body of research on applying a combination of automated planning and machine learning as a tool for neuroscientific research [64, 65].

2.3 Intention Progression

Autonomous agents often have to achieve multiple goals in parallel. Even if the behaviour or plan to achieve each goal is (provably) correct, e.g., has been verified or derived from a specification of the goal or task, the problem of *acting* (plan execution) remains [66]. The *Intention Progression Problem* (IPP) is the problem of what a Belief-Desire-Intention (BDI) agent should do next [67], that is, what means (i.e., plan) to use to achieve a given (sub)goal; and which of the currently adopted plans (i.e., intentions), to progress at the current moment. An important capability of an intelligent agent is the ability to progress multiple intentions in parallel, by interleaving the steps in each intention to provide the best outcome for the agent. This problem is both central to agent reasoning and complex in its nature. For example, ‘best outcome’ may have different definitions depending on the application, while goals and plans may conflict given the resources available. A key challenge is the interleaving of steps in plans in different intentions to avoid conflicts, i.e., when the execution of a step in one plan makes the execution of a step in another concurrently executing plan impossible.

Members of the A^3 group have contributed both to state of the art approaches to the IPP, and to the development of the field through the Intention Progression Competition⁴. For example, approaches based on Monte Carlo Tree Search to intention progression at both the plan [68] and action [69] level have been proposed, while in [70] it was shown how similar approaches can be used to recover from execution failures, by exploiting positive interactions between intentions. A positive interaction occurs when the execution of an action in one intention assists the execution of plans or actions in other intentions, for example, by (re)establishing their preconditions. More recent work has focused on intention progression under uncertainty [71], e.g., when it is not known whether a particular plan is applicable or an action is executable, and how offline analysis of possible conflicts between an agent’s plans can significantly reduce the computation required at runtime [72], increasing the responsiveness of agents to unanticipated changes in the environment.

2.4 Argumentation-based Explanations

As agents are becoming more complex and ubiquitous in our everyday lives, it has become crucial to be able to explain their behaviour and allow them to have profound interactions with humans. While there are myriads of explanation techniques using graphs, statistics, and natural language descriptions, argumentation has a unique advantage in transparently explaining the procedure and the results of reasoning through an “arguing process”. The usual process of arguing can be viewed as (1) identifying and modelling the information at hand and (2) generating an explanation for the topic, usually through some fictitious proponent and opponent debate game. Of course, the structure of the arguments and the protocol of the debate game will vary depending on the desired expressiveness of the explanations and the nature of the topic.

For instance, Yun, Vesic and Oren [73] proposed a dialectical framework based on argumentation to compute and explain pure strategy Nash equilibria. Roughly speaking, this framework extends the framework of [74] (allowing attacks on attacks) and models a two-agent dialogue, where the proponent’s goal is to show that an argument is a Nash equilibrium while the opponent seeks to demonstrate that the proponent’s argument is not a Nash equilibrium by proposing alternatives. Another example is the work of Dennis and Oren [75], where they proposed a dialogue-based approach for explaining the behaviour of a BDI agent. This dialogue considers two participants, who may have different views regarding the beliefs, plans and external events, and make utterances which incrementally reveal their traces to each other, allowing them to identify divergences or to conclude that their traces agree. Moreover, as part of the EPSRC funded “Scrutable Autonomous Systems” project, the A^3 group made use of their expertise to design a tool based on argumentation to explain automated planning [7], allowing for humans and agents to coordinate actions. Much of this project examined efficient proof dialogues for providing such explanation [76], as well approaches to knowledge representation which can naturally lead to argumentative structures [77].

The members of the A^3 group also researched how to convey explanations through various forms of graphs. This is critical as Vesic, Yun and Teovanovic have shown that interacting with the graphical representation

⁴<http://intentionprogression.org/> (Accessed: 03/05/2022).

of arguments can increase the compliance with desirable and rational behaviours [78]. As a result, the group has developed several models and concrete tools to generate argumentation graphs automatically from logical knowledge bases [79, 80, 81, 82].

2.5 Computational Trust and Normative Reasoning

In an open MAS, agents must identify others to coordinate with, either to undertake joint actions, or to delegate tasks to. A core problem here is that some agents within the system may be malicious, or simply incompetent, and such agents should ideally be avoided, or that measures be put into place mitigating the effect of such agents in order to improve resilience and reliability of the system.

A trust and reputation system allows an agent to combine their direct experiences of others (*trust*) and obtain information from third parties about others (*reputation*), so as to compute a reliability ordering about these other agents. This reliability ordering can then be used to decide who to interact with [6, 20], or to put other measures into place so as to minimise the harm these other agents can cause. For example, we combined a trust rating with principal agent theory to determine how much utility transfer should take place between a delegator and a delegatee, and how much monitoring the delegator should undertake of the delegatee’s behaviour [6, 83].

The Eigentrust framework [84] is a simple and popular approach to computing trust among agents, originally used in the context of peer-to-peer file sharing. Eigentrust assumes that trust is transitive (if A trusts B and B trusts C then A should trust C), and that an agent’s ability to provide reputational information is correlated to its ability to perform a task. It then represents the (direct) trust relationship between agents in a matrix. Multiplying a vector by this matrix yields a new vector representing direct trust, and — under the assumptions described above — further multiplications of the resultant vector by the matrix yield a new vector describing first hand, second hand, and so on reputational information. However, Eigentrust does not discriminate between lack of information about an agent, and lack of trust in the agent. We proposed the MaxTrust algorithm [2] to take this difference into account, and it has been shown to obtain more accurate trust ratings in some situations.

Another element of work on trust taking place in the A^3 group considers the case where a task can be delegated multiple times. For example, a company could outsource some task to another, which further outsources the task until it is eventually executed by the last element of a *delegation chain*. The question we ask is how trust should be updated by, and for, all agents in the delegation chain. Beginning with an empirical evaluation [85], we have described more principled approaches building on ideas from game theory and multi-arm bandits [20] to deal with trust in such sub-delegative domains. Our current research focuses on extending this work by integrating reputational information, introducing parallelism, and also considering a model of others in trust update decisions. A related strand of work considers responsibility and blame for failures in team plans [86, 87], and allows the degree of responsibility of each agent for the failure, and the degree of blame attached to each agent to be determined.

While trust allows for the filtering of interaction partners, another aspect of resilience and robustness involves putting appropriate sanctions into place to punish misbehaving agents, and to compensate those who suffered. The A^3 group has a long history of research into normative systems and electronic contracting to perform such tasks. For example, [88] introduces a contracting language from which obligations, permissions and prohibitions with regards to a target can be described (c.f., commitments [89]), while [90] details how an agent can reason in the context of norms encoded as numerical constraints, and [91] studies the problem of checking whether agents can bring about a state satisfying a property without incurring a specified number of sanctions. Norms are inherently non-monotonic [92], and our recent work has examined how argumentation can be used to undertake practical reasoning in the presence of both norms and goals [93].

A key problem in normative MAS is detecting norm violations. For large-scale MAS, monitoring for violations may involve significant computational costs. Members of the A^3 group have proposed decentralised approaches to norm monitoring, in which agents are incentivised to monitor the actions of other agents for norm violations [94]. In [95], this work is extended to settings in which there is no crisp definition of a norm violation; rather, it is a matter of judgement whether an agent’s behaviour conforms to generally accepted standards of behaviour, and agents may legitimately disagree about borderline cases. Other norm-related

work has considered what (and where) monitors need to be placed to check for compliance with obligations and prohibitions [96], the strongest norm that can be enforced given the available monitoring infrastructure [97], and which norms can be enforced in practical settings [98]. When such monitors are unavailable, plan recognition can be used to detect deviation from norms [99].

Norms may not always be explicit, and an agent entering a new system may need to learn the system's norms. We have described a plan recognition based approach to doing so [100], which effectively filters out norms inconsistent with executed plans. A refinement to this work [4] used a Bayesian approach to identify the most likely norms.

3 Coordination

There is a large body of research dealing with coordination among agents including work in game theory [101] and multi-agent planning [102, 103]. At A^3 , the focus of research into coordination has revolved around four elements, namely (a) argumentation as a mechanism for reasoning about how and why coordination should take place; (b) how to generate decentralised plans and joint plans using multi-agent planning; (c) goal recognition as a means to understand others' objectives so as to allow coordination to take place; and (d) a framework for evaluating the effectiveness of coordination. We now describe each of these topics in more detail and summarise our contributions.

3.1 Making Decisions in a Multi-Agent Setting

The problem of making a decision when multiple agents are involved is very complex, especially when many parameters must be taken into account for the decision. There can be one or several global costs functions that need to be optimised, agents can have conflicting viewpoints about the situation at hand and can also have different objectives. Within A^3 , argumentation [104] and dialogue [105] are used as the underpinning paradigm allowing agents to exchange information, reason in the presence of conflict, and reach and understand the reasons for a (joint) decision. For example, as part of the EU funded "EcoBioCAP" project, the problem of designing innovative packaging solutions was investigated. Here, experts taking the role of agents exchange information to optimise multiple criteria regarding food packaging such as the packaging's cost, environmental impact, and bio-safety. Yun et al. [19, 106] then proposed an argumentation-based tool to drive this information exchange, take account of conflicting agent preferences, and help the experts understand the impact of arguments advanced by them and others on the final outcome. In other works, Yun et al. [107] have studied how to best support decision making in more constrained situations, i.e., when only given agents' arguments and no preferences are elicited. They proposed an approach that makes use of ranking-based semantics [108, 109] to provide agents with viewpoints (non-conflicting sets of arguments) that are the most relevant with respect to several desirable postulates.

It should also be noted that within a MAS, concepts such as trust, goals, obligations and permissions, as well as the provenance of information also affect decision-making. Research within A^3 has examined how an argumentation semantics can also be used to reason alongside such concepts [110, 93, 111, 112, 8]. As part of such decision making, strength of evidence (due to different levels of trust in sources or simple uncertainty) must be considered, and working with different types of argument strength is another element of the group's research [113, 114, 115, 116].

While there are increasing evidence that argumentation-based approaches are able to model human reasoning to a certain extent [117, 118], the generation of arguments [119] and the expensive computation of some argumentation extensions [120], such as the preferred-semantics, are often limitations that prevent their wide-adoption in settings where fast decision-making is required. To accommodate such settings, Yun et al. have also considered rule-based approaches that can be applied directly to knowledge bases in the Ontology-Based Data Access (OBDA) setting [121, 122]. In this setting, a set of well-designed rules (an *ontology*), is used to access and combine the knowledge of each agent, allowing for heterogeneous agents with different vocabularies (e.g., agents working in different application domain) to seamlessly communicate. For instance, in [122], a multi-criteria decision-making system is proposed to model each agent criteria evaluation

over the alternatives using expressive languages. This avenue of research was further investigated by Yun et al. [121], who combined inconsistency measures with *Shapley* techniques to assist decision-making. The approach was successfully applied on the aforementioned packaging use-case, allowing for the data from more than 20 professionals working in the food industry to be aggregated and used in the decision-making process.

3.2 Multi-Agent Planning

Multi-Agent Planning (MAP) had many different interpretations over the years, but in general the overall process can be interpreted around two main aspects: *a)* the planning process itself is either centralised (performed by a single agent) or distributed (performed by multiple agents), and *b)* the solution is for a single agent or for multiple agents. Planning done by a single agent (centralised planner) will encounter the search state-space explosion problem as the number of agents to plan for increases. Distributing/decentralising the planning process can lead to faster computation times depending on the number of agents to plan for and how tightly coupled the actions of these agents are. Tightly coupled actions require more coordination before/after planning, but in loosely coupled scenarios using multiple agents to perform planning can result in substantial improvements in planning time [14]. Another advantage of MAP is that it can utilise privacy-preserving algorithms [123] to maintain various levels of privacy during planning so that private planning information (such as actions, goals, etc.) are not shared amongst agents.

Cardoso and Bordini developed the Decentralised Online Multi-Agent Planning (DOMAP) framework [14] that combines HTN planning with the JaCaMo MAS development platform [124], resulting in BDI agents that can plan at runtime using HTN planning, and then coordinate their actions during execution using JaCaMo’s organisational dimension. The most notable contributions of the framework is that it can be used to bridge the gap between planning and execution often found in the literature, it provides fair goal allocation prior to planning using a contract net protocol mechanism, and it performs decentralised HTN planning using an off-the-shelf HTN planner. Results show that DOMAP outperforms the best planners from the 2015 International Competition of Distributed and Multi-Agent Planners (CoDMAP) [125] in terms of planning time, execution time, and parallelism (variance of the plan size of each individual agent, used to indicate the spread of actions and to identify how well execution loads are balanced) for the most difficult problems (large number of agents). As shown by DOMAP, MAS dimensions (agent, environment, and organisation) and agent programming can help to bridge the gap between planning and execution, and future work in this topic can tackle some of the most difficult problems in MAP such as dealing with conflicting actions, coordinating concurrent and joint actions, and enforcing privacy constraints.

3.3 Goal and Plan Recognition

Goal Recognition is a task related to automated planning (Section 2.2), where an agent employs abductive reasoning to infer the most likely goal pursued by another agent [126]. The evidence for this kind of reasoning often consists of a sequence of observations of the observed agent’s plan. Here, instead of deducing a plan from an initial state towards a goal using some kind of domain theory, the deduction is about filtering the correct goal out of a set of goal hypotheses. Research on goal recognition lies within the context of *Plan, Activity, and Intent Recognition* [127] and employs distinct inference techniques to recognise the ultimate goals of agents under observation. Goal recognition is also related to the problem of *Plan Recognition* [128], which consists of trying to infer the actual plan adopted by the observed agent. The task of goal recognition has a number of potential and actual applications, including assisting the handicapped [129], activities of daily living [130], workplace safety [131], among others [132, 133].

In the context of the A^3 group, goal and plan recognition provide a key mechanism for non-verbal communication, as agents that can recognise the goals and plans adopted by others can more effectively coordinate their activities, minimising the need for explicit communication. Members of the A^3 group have made significant contributions in addressing both of these problems at the theoretical and practical levels.

On the theoretical side, Meneguzzi, Oren and Pereira developed the current generation of the state-of-the-art in goal and plan recognition algorithms. This has started with the application of planning landmarks to

perform accurate and efficient goal recognition [134, 135, 21, 136]. Building upon this work, the group developed approaches for online goal recognition in continuous domains [137], as well as in incomplete [138], and learned domains [139]. The latest iteration of novel goal recognition techniques employs linear programming and operator-counting heuristics to perform goal and plan recognition under noisy and low observability conditions [140]. More recently, Meneguzzi and others developed a series of techniques for hybrid, neuro-symbolic reasoning towards goal recognition, to both automatically derive symbolic representations from unstructured data [130, 141, 142, 143, 144, 145], and enhancing the recognition process with learned preferences of the agents under observation [146]. Finally, Meneguzzi and others have bridged the gap between reinforcement learning and goal recognition techniques by providing the first formalisation and efficient algorithms for goal recognition as reinforcement learning [147].

On the practical side, Meneguzzi and others have developed a number of applications of plan and activity recognition, including activities of daily living [148] and scene recognition [149]. Goal recognition is an area that is at least as broad as the applications of automated planning itself [44]. These areas include network security [150], crowd safety [151], among others.

If the goals of other agents can be recognised (or are known), then agents are able to coordinate more effectively. The multi-agent intention progression problem for BDI agents extends the IPP presented in Section 2.3. In the multi-agent setting, how an agent progresses its intentions has implications for both the achievement of its own goals and the achievement of the goals of other agents, e.g., if the agent selects a plan that consumes a resource necessary for another agent to achieve its goal. In [152], the MCTS-based approach to intention progression discussed in Section 2.3 is extended to support *intention aware* scheduling, where each agent attempts to anticipate the possible actions of other agents in the environment when progressing their own intentions. Intention-aware scheduling was shown to be more effective than the approaches in [69] in cooperative, selfish and competitive environments. More recently, this approach to intention-aware scheduling has been extended to the case in which the plans used by other agents are unknown, and agents use an abstraction of their own program called a partially-ordered goal-plan tree (pGPT) to schedule their intentions and predict the actions of other agents [153].

3.4 Multi-Agent Programming

Agent programming languages focus on rational agents, which are computational programs that usually make use of a reasoning cycle in order to make their decisions. The most common reasoning model in existing agent programming languages follows the BDI model, as shown in recent literature surveys [154, 155, 156, 157].

The Multi-Agent Programming Contest (MAPC)⁵ is an annual competition aimed at advancing the state-of-the-art in terms of tools, programming languages, and methodologies for the engineering of MAS, in particular to try to bring forward the advantages of agent programming languages. Teams of agents compete against each other in a synchronous simulation, with every step lasting only a few seconds in which agents have to submit their actions for the next step. During this limited period, agents are encouraged to coordinate and agree on their actions in order to work together as a team and achieve the goals of the scenario efficiently.

Members of A^3 have participated and obtained 1st place in MAPC 2016 [158], 2nd place in MAPC 2018 [159], 1st place in MAPC 2019 [160] and 2nd place in MAPC 2020/21 [161]. In all cases the team was programmed using JaCaMo [124]. The main strategies used by our teams in the past included the use of an automated planner to perform optimal planning for the individual movement of agents in the grid [162], a formally verified map merging protocol [163] that agents followed to coordinate and build a global map, and the organisational layer in JaCaMo to specify roles and coordination schemes that were used to guide the high-level coordination of joint actions.

⁵<https://multiagentcontest.org/> (Accessed: 03/05/2022)

4 Open Challenges and Future Directions

Recent developments in AI, and particularly in machine learning, have the potential to significantly impact the development and runtime behaviour of agent systems. However, while there has been remarkable advances in machine learning, many tasks remain resistant to learning-based approaches, for reasons of explainability, the need for commonsense or causal reasoning, among many others. We believe that future agent systems must exploit the advantages of both symbolic and sub-symbolic techniques to ensure reliability and resilience in challenging domains, such as long-term autonomy, human-robot teamwork, and hybrid intelligence. For example, such systems must be capable of robust autonomous decision-making taking into account socio-technical concepts (including accountability, responsibility and trust). To be trusted by users and approved by regulators, we must be able to offer guarantees that the behaviour of such systems is, at the very least, safe. This will require the development of new approaches to the verification and synthesis of MAS that combine symbolic and sub-symbolic approaches. Verification of such hybrid systems is extremely challenging, and will require significant advances across a range of areas, including logics of strategic ability, automata-based learning, stochastic search and quantitative games on graphs, among others.

We believe that argumentation and dialogue can also address several challenges related to agent reasoning, coordination and explanation. One strand of active research involves investigating how closely complex argumentation systems (e.g., those that incorporate weights or structured arguments) mimic human reasoning [164, 165], and how such systems can be made to comply with desirable properties. Such work can allow an agent to infer a human’s preferences [5], or identify how strong information needs to be to convince a user of some conclusion [115]. At a more applied level, we intend to investigate argument-based reasoning in the context of socio-technical concepts such as trust, responsibility and accountability [166].

Despite recent trends in deep learning [167], key applications of AI in our daily lives will involve collaboration between AI systems and humans [168]. Such collaboration imposes several challenges to AI systems, one of which involves coordinating the activities of AI-driven algorithms and their human partners. To accomplish such coordination, AI systems must be capable of reasoning about a number of aspects of cognition and behaviour. This includes inferring the intentions of their human interlocutors (Section 3.3), reaching consensus among mixed teams of agents and humans (Section 3.1), and planning for joint courses of actions (Section 3.2). Our group envisions addressing these kinds of challenges using a mixture of classic AI models (for robustness and explainability), as well as machine learning models to deal with inherently noisy environmental data. Recent work in this direction includes eliciting (human) preferences from the arguments they advance [5], and identifying what strength an argument requires to be acceptable to a human reasoner [115].

Most of our work in the past has been in symbolic AI (representation of AI problems through the use of symbols, simulating the human reasoning process) with some applications of machine learning [146]. However, there are still many open challenges and potential advantages of combining symbolic AI and machine learning techniques. For example, symbolic AI has been shown to be effective in scenarios where reactivity is necessary [161], while machine learning has demonstrated excellent results in scenarios where learning is essential [169]. Therefore, scenarios that require both reactive and learning behaviours, such as self-driving cars or autonomous vehicles, would greatly benefit from such integration. Moreover, with recent advances in neuro-symbolic AI [170], we can now align ontologies and concepts with neural networks. A new research avenue will be to integrate argumentation with such technologies to boost the reliance of neural network systems by verifying that they comply with basic rules and possibly discover hidden patterns in data. This will also allow A^3 to secure a new way to obtain arguments, possibly bringing argumentation closer to machine learning.

Alongside the purely architectural questions of how symbolic and sub-symbolic approaches can be effectively combined, work in this direction gives rise to a range of new research problems centred around the notion of *bounded adaptation* [171]. How should the split between predefined or canonical behaviours and learned behaviours (e.g., refinements, or implementations of very high-level actions, etc.) be characterised? What development methodologies and verification approaches can be used to specify and certify the behaviour of agents that integrate significant AI capabilities into their decision making? This can be seen as establishing a new strand of research exploring hybrids of programming-based, learning-based, and model-based approaches to developing AI capabilities [172]. We believe that the agent programming paradigm forms

an ideal framework in which to explore the resulting complex mix of scientific and engineering questions.

Since several members of A^3 have only joined the group recently, most of the research topics discussed in this paper have been investigated separately. Combining the different elements of A^3 's research may also yield interesting insights. For example, argument-based explanation could enhance formal verification of systems by providing easy to understand reasons as to why certain properties are violated. Similarly, plan recognition techniques may be used to ascribe responsibility with regards to certain undesirable outcomes in a system, while multi-agent planning techniques can be combined with other approaches to achieve robust system behaviour.

5 Conclusion

We have discussed recent work in A^3 related to resilience, reliability, and coordination in MAS. We focused on a variety of topics for each theme. Our work in verification, in particular model checking and runtime verification, has been demonstrated to improve reliability of cyber-physical systems through the use of various tools and applications. Still related to reliability, we have described our contributions for automated generation of programs in both program synthesis and automated planning, allowing the generation of provably correct programs. In regards to resilience, we have identified the intention progression problem and proposed a set of state of the art solutions, presented a range of argumentation-based techniques for providing explanations about agents' behaviour, and discussed our past work in trust, reputation, and normative systems. Moving on to coordination, we have reported a series of works based on using argumentation to aid in making group decisions, a decentralised planner for generating multi-agent plans, goal and plan recognition techniques as a mechanism for non-verbal communication, and a series of applications in multi-agent oriented programming scenarios where coordination is essential.

There are still many open challenges that are directly related to resilience, reliability, and coordination in MAS that we will continue to investigate by extending some of the work reported here. Nevertheless, we also plan to use our past experience in these research themes to combine and apply them with socio-technical concepts, human-agent scenarios, and hybrid symbolic and statistical agents.

Acknowledgements

The research reported in this paper was funded and supported by various grants over the years: Robotics and AI in Nuclear (RAIN) Hub (EP/R026084/1); Future AI and Robotics for Space (FAIR-SPACE) Hub (EP/R026092/1); Offshore Robotics for Certification of Assets (ORCA) Hub (EP/R026173/1); the Royal Academy of Engineering under the Chair in Emerging Technologies scheme; Trustworthy Autonomous Systems "Verifiability Node" (EP/V026801); Scrutable Autonomous Systems (EP/J012084/1); Supporting Security Policy with Effective Digital Intervention (EP/P011829/1); The International Technology Alliance in Network and Information Sciences.

References

- [1] A. Etuk, T.J. Norman, N. Oren and M. Sensoy, Strategies for Truth Discovery under Resource Constraints, in: *AAMAS*, ACM, 2015, pp. 1807–1808.
- [2] J. Afanador, N. Oren, M.S. Baptista and M. Araujo, From Eigentrust to a Trust-Measuring Algorithm in the Max-Plus Algebra, in: *ECAI*, Frontiers in Artificial Intelligence and Applications, Vol. 325, IOS Press, 2020, pp. 3–10.
- [3] F. Meneguzzi, O. Rodrigues, N. Oren, W.W. Vasconcelos and M. Luck, BDI reasoning with normative considerations, *Eng. Appl. Artif. Intell.* **43** (2015), 127–146.

- [4] S. Cranefield, F. Meneguzzi, N. Oren and B.T.R. Savarimuthu, A Bayesian Approach to Norm Identification, in: *ECAI*, Frontiers in Artificial Intelligence and Applications, Vol. 285, IOS Press, 2016, pp. 622–629.
- [5] Q. Mahesar, N. Oren and W.W. Vasconcelos, Preference Elicitation in Assumption-Based Argumentation, in: *PRIMA*, Lecture Notes in Computer Science, Vol. 12568, Springer, 2020, pp. 199–214.
- [6] C. Burnett, T.J. Norman, K.P. Sycara and N. Oren, Supporting Trust Assessment and Decision Making in Coalitions, *IEEE Intell. Syst.* **29**(4) (2014), 18–24.
- [7] N. Oren, K. van Deemter and W.W. Vasconcelos, Argument-Based Plan Explanation, in: *Knowledge Engineering Tools and Techniques for AI Planning*, Springer, 2020, pp. 173–188.
- [8] A. Toniolo, T. Norman, A. Etuk, F. Cerutti, R.W. Ouyang, M. Srivastava, N. Oren, T. Dropps, J.A. Allen and P. Sullivan, Supporting reasoning with different types of evidence in intelligence analysis (2015).
- [9] A. Ferrando, R.C. Cardoso, M. Fisher, D. Ancona, L. Franceschini and V. Mascardi, ROSMonitoring: A Runtime Verification Framework for ROS, in: *Towards Autonomous Robotic Systems*, Springer International Publishing, Cham, 2020, pp. 387–399. ISBN ISBN 978-3-030-63486-5. doi:10.1007/978-3-030-63486-5_40.
- [10] A. Ferrando. and R. Cardoso., RVPLAN: Runtime Verification of Assumptions in Automated Planning, in: *Proceedings of the 14th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART*,, SciTePress, 2022, pp. 67–77, INSTICC. ISBN ISBN 978-989-758-547-0. doi:10.5220/0010776500003116.
- [11] N. Oren, W. Vasconcelos, F. Meneguzzi and M. Luck, Acting on Norm Constrained Plans, in: *Computational Logic in Multi-Agent Systems, 11th International Workshop*, LNCS, 2011, pp. 347–363. <http://www.meneguzzi.eu/felipe/pubs/clima-identifying-2011.pdf>.
- [12] J.P. Aires, D. Pinheiro, V.S.d. Lima and F. Meneguzzi, Norm conflict identification in contracts, *Artificial Intelligence and Law* **25**(4) (2017), 397–428. doi:10.1007/s10506-017-9205-x.
- [13] F. Meneguzzi and M. Luck, Leveraging new plans in AgentSpeak(PL), in: *Proceedings of the Sixth Workshop on Declarative Agent Languages*, M. Baldoni, T.C. Son, M.B. van Riemsdijk and M. Winikoff, eds, 2008, pp. 63–78.
- [14] R.C. Cardoso and R.H. Bordini, Decentralised Planning for Multi-Agent Programming Platforms, in: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2019, pp. 799–807–. ISBN ISBN 9781450363099.
- [15] M. Naiseh, C. Bentley, S. Ramchurn, E. Williams, E. Awad and C. Alix, Methods, Tools and Techniques for Trustworthy Autonomous Systems (TAS) Design and Development, in: *Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '22 Companion, Association for Computing Machinery, Sophia Antipolis, France, 2022, pp. 66–69–. ISBN ISBN 9781450390316. doi:10.1145/3531706.3536459.
- [16] J. Dix, B. Logan and M. Winikoff, Engineering Reliable Multiagent Systems (Dagstuhl Seminar 19112), *Dagstuhl Reports* **9**(3) (2019), 52–63. doi:10.4230/DagRep.9.3.52. <http://drops.dagstuhl.de/opus/volltexte/2019/11291>.
- [17] J. Dix, B. Logan and M. Winikoff, Preface to the Special Issue on engineering reliable multi-agent systems, *Auton. Agents Multi Agent Syst.* **35**(2) (2021), 37. doi:10.1007/s10458-021-09520-y.

- [18] M. Fisher, R.C. Cardoso, E.C. Collins, C. Dadswell, L.A. Dennis, C. Dixon, M. Farrell, A. Ferrando, X. Huang, M. Jump, G. Kourtis, A. Lisitsa, M. Luckcuck, S. Luo, V. Page, F. Papacchini and M. Webster, An Overview of Verification and Validation Challenges for Inspection Robots, *Robotics* **10**(2) (2021). doi:10.3390/robotics10020067. <https://www.mdpi.com/2218-6581/10/2/67>.
- [19] B. Yun, P. Bisquert, P. Buche, M. Croitoru, V. Guillard and R. Thomopoulos, Choice of environment-friendly food packagings through argumentation systems and preferences, *Ecol. Informatics* **48** (2018), 24–36. doi:10.1016/j.ecoinf.2018.07.006.
- [20] J. Afanador, M.S. Baptista and N. Oren, Algorithms for recursive delegation, *AI Communications* **32**(4) (2019), 303–317.
- [21] R.F. Pereira, N. Oren and F. Meneguzzi, Landmark-based approaches for goal recognition as planning, *Artificial Intelligence* **279** (2020), 103217. doi:10.1016/j.artint.2019.103217. <https://linkinghub.elsevier.com/retrieve/pii/S0004370219300013>.
- [22] I. Pinyol and J. Sabater-Mir, Computational trust and reputation models for open multi-agent systems: a review, *Artificial Intelligence Review* **40**(1) (2013), 1–25.
- [23] M. Croitoru, N. Oren, S. Miles and M. Luck, Graphically explaining norms, in: *9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010), Toronto, Canada, May 10-14, 2010, Volume 1-3*, W. van der Hoek, G.A. Kaminka, Y. Lespérance, M. Luck and S. Sen, eds, IFAAMAS, 2010, pp. 1405–1406. <https://dl.acm.org/citation.cfm?id=1838404>.
- [24] E.M. Clarke, O. Grumberg, D. Kroening, D.A. Peled and H. Veith, *Model checking, 2nd Edition*, MIT Press, 2018. ISBN ISBN 978-0-262-03883-6. <https://mitpress.mit.edu/books/model-checking-second-edition>.
- [25] E. Bartocci and Y. Falcone (eds), *Lectures on Runtime Verification - Introductory and Advanced Topics*, Lecture Notes in Computer Science, Vol. 10457, Springer, 2018. ISBN ISBN 978-3-319-75631-8. doi:10.1007/978-3-319-75632-5.
- [26] M. Luckcuck, M. Farrell, L.A. Dennis, C. Dixon and M. Fisher, Formal Specification and Verification of Autonomous Robotic Systems: A Survey, *ACM Computing Surveys* **52**(5) (2019), 1–41. doi:10.1145/3342355.
- [27] N. Alechina, P. Bertoli, C. Ghidini, M. Jago, B. Logan and L. Serafini, Verifying space and time requirements for resource-bounded agents, in: *Proceedings of the Fourth Workshop on Model Checking and Artificial Intelligence (MoChArt-2006)*, S. Edelkamp and A. Lomuscio, eds, 2006, pp. 16–30.
- [28] A. Albore, N. Alechina, P. Bertoli, C. Ghidini, B. Logan and L. Serafini, Model-checking memory requirements of resource-bounded reasoners, in: *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI 2006)*, AAAI Press, 2006, pp. 213–218. ISBN ISBN 978-1-57735-279-2.
- [29] N. Alechina, B. Logan, H.N. Nguyen and A. Rakib, Verifying time, memory and communication bounds in systems of reasoning agents, in: *Proceedings of the Seventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Vol. 2, L. Padgham, D. Parkes, J. Müller and S. Parsons, eds, IFAAMAS, Estoril, Portugal, 2008, pp. 736–743, IFAAMAS. ISBN ISBN 978-0-981-7381-0-9.
- [30] N. Alechina, B. Logan, H.N. Nguyen and A. Rakib, A Logic for Coalitions with Bounded Resources, in: *Proceedings of the Twenty First International Joint Conference on Artificial Intelligence (IJCAI 2009)*, Vol. 2, C. Boutilier, ed., AAAI Press, Pasadena CA, USA, 2009, pp. 659–664, IJCAI/AAAI.

- [31] N. Alechina, B. Logan, H.N. Nguyen and A. Rakib, Resource-bounded alternating-time temporal logic, in: *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, W. van der Hoek, G. Kaminka, Y. Lespérance, M. Luck and S. Sen, eds, IFAAMAS, Toronto, Canada, 2010, pp. 481–488, IFAAMAS.
- [32] N. Alechina, B. Logan, L. Mostarda, H.N. Nguyen and F. Raimondi, Symbolic Model-Checking for Resource-Bounded ATL, in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, R. Bordini, E. Elkind, G. Weiss and P. Yolu, eds, IFAAMAS, Istanbul, Turkey, 2015, pp. 1809–1810, IFAAMAS.
- [33] N. Alechina, B. Logan, H.N. Nguyen and F. Raimondi, Model-checking for Resource-Bounded ATL with production and consumption of resources, *Journal of Computer and System Sciences* **88** (2017), 126–144. doi:doi.org/10.1016/j.jcss.2017.03.008.
- [34] N. Alechina, B. Logan, H.N. Nguyen and F. Raimondi, Decidable Model-Checking for a Resource Logic with Production of Resources, in: *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI-2014)*, T. Schaub, G. Friedrich and B. O’Sullivan, eds, IOS Press, Prague, Czech Republic, 2014, pp. 9–14, ECCAI.
- [35] N. Alechina, N. Bulling, B. Logan and H.N. Nguyen, On the Boundary of (Un)Decidability: Decidable Model-Checking for a Fragment of Resource Agent Logic, in: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, Q. Yang, ed., AAAI Press, Buenos Aires, Argentina, 2015, pp. 1494–1501, IJCAI.
- [36] N. Alechina, M. Dastani and B. Logan, Verifying Existence of Resource-Bounded Coalition Uniform Strategies, in: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, S. Kambhampati, ed., IJCAI/AAAI Press, 2016, pp. 24–30.
- [37] N. Alechina, N. Bulling, B. Logan and H.N. Nguyen, The virtues of idleness: A decidable fragment of resource agent logic, *Artificial Intelligence* **245** (2017), 56–85. doi:http://dx.doi.org/10.1016/j.artint.2016.12.005.
- [38] N. Alechina, N. Bulling, S. Demri and B. Logan, On the complexity of resource-bounded logics, *Theoretical Computer Science* **750** (2018), 69–100. <https://doi.org/10.1016/j.tcs.2018.01.019>.
- [39] N. Alechina, S. Demri and B. Logan, Parameterised Resource-Bounded ATL, in: *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*, AAAI Press, New York, USA, 2020, AAAI.
- [40] A. Ferrando, L.A. Dennis, R.C. Cardoso, M. Fisher, D. Ancona and V. Mascardi, Toward a Holistic Approach to Verification and Validation of Autonomous Cognitive Systems, *ACM Trans. Softw. Eng. Methodol.* **30**(4) (2021). doi:10.1145/3447246.
- [41] N. Alechina, B. Logan, H.N. Nguyen and F. Raimondi, Symbolic Model-Checking for One-Resource RB+-ATL, in: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, Q. Yang, ed., AAAI Press, Buenos Aires, Argentina, 2015, pp. 1069–1075, IJCAI.
- [42] A. Lomuscio, H. Qu and F. Raimondi, MCMAS: An Open-Source Model Checker for the Verification of Multi-Agent Systems, *International Journal on Software Tools for Technology Transfer* **19**(1) (2017), 9–30.
- [43] R.C. Cardoso, M. Farrell, M. Luckcuck, A. Ferrando and M. Fisher, Heterogeneous Verification of an Autonomous Curiosity Rover, in: *Proc. 12th International NASA Formal Methods Symposium (NFM)*, LNCS, Vol. 12229, Springer, 2020, pp. 353–360.
- [44] H. Geffner and B. Bonet, *A Concise Introduction to Models and Methods for Automated Planning*, 1st edn, Morgan & Claypool, 2013.

- [45] G. De Giacomo, F. Patrizi and S. Sardiña, Automatic behavior composition synthesis, *Artificial Intelligence* **196** (2013), 106–142.
- [46] P. Felli, B. Logan and S. Sardina, Parallel Behavior Composition for Manufacturing, in: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, S. Kambhampati, ed., IJCAI/AAAI Press, 2016, pp. 271–278.
- [47] L. de Silva, P. Felli, J.C. Chaplin, B. Logan, D. Sanderson and S. Ratchev, Realisability of Production Recipes, in: *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI-2016)*, G.A. Kaminka, M. Fox, P. Bouquet, H. E., D. F., D. V. and van Harmalen F., eds, IOS Press, The Hague, The Netherlands, 2016, pp. 1449–1457, ECCAI. doi:10.3233/978-1-61499-672-9-1449.
- [48] P. Felli, L. de Silva, B. Logan and S. Ratchev, Process Plan Controllers for Non-Deterministic Manufacturing Systems, in: *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI 2017)*, C. Sierra, ed., IJCAI, Melbourne, Australia, 2017, pp. 1023–1030. doi:10.24963/ijcai.2017/142.
- [49] G. De Giacomo, M. Vardi, P. Felli, N. Alechina and B. Logan, Synthesis of Orchestrations of Transducers for Manufacturing, in: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, AAAI Press, New Orleans, USA, 2018, pp. 6161–6168, AAAI. <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17250/16186>.
- [50] L. de Silva, P. Felli, D. Sanderson, J.C. Chaplin, B. Logan and S. Ratchev, Synthesising process controllers from formal models of transformable assembly systems, *Robotics and Computer Integrated Manufacturing* **58** (2019), 130–144.
- [51] N. Alechina, T. Brázdil, G. De Giacomo, P. Felli, B. Logan and M.Y. Vardi, Unbounded Orchestrations of Transducers for Manufacturing, in: *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, AAAI Press, Honolulu, USA, 2019, AAAI.
- [52] L. de Silva, P. Felli, J.C. Chaplin, B. Logan, D. Sanderson and S. Ratchev, Synthesising Industry-Standard Manufacturing Process Controllers (Demonstration), in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)*, S. Das, E. Durfee, K. Larson and M. Winikoff, eds, IFAAMAS, Sao Paulo, Brazil, 2017, pp. 1811–1813, IFAAMAS.
- [53] P. Felli, L. de Silva, , B. Logan and S. Ratchev, Composite Capabilities for Cloud Manufacturing (Demonstration), in: *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, M. Dastani, G. Sukthankar, E. Andre and S. Koenig, eds, IFAAMAS, Stockholm, Sweden, 2018, pp. 1809–1811, IFAAMAS.
- [54] M.C. Magnaguagno, R.F. Pereira, M.D. Móre and F. Meneguzzi, Develop, Visualize and Test Classical Planning Descriptions in your Browser (Demo), in: *Proceedings of the 29th International Conference on Automated Planning and Scheduling (ICAPS)*, AAAI Press.
- [55] M.C. Magnaguagno, R.F. Pereira, M.D. Móre and F. Meneguzzi, *Web Planner: A Tool to Develop, Visualize, and Test Classical Planning Domains*, in: *Knowledge Engineering Tools and Techniques for AI Planning*, M. Vallati and D. Kitchin, eds, Springer International Publishing, Cham, 2020, pp. 209–227. ISBN ISBN 978-3-030-38561-3.
- [56] M.C. Magnaguagno and F. Meneguzzi, HTN Planning with Semantic Attachments, in: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, AAAI Press, 2020.
- [57] L. de Silva and F. Meneguzzi, On the Design of Symbolic-Geometric Online Planning Systems, in: *2015 Workshop on Hybrid Reasoning (HR 2015)*, pp. 1–8. <http://www.meneguzzi.eu/felipe/pubs/hr-symbolic-geometric-2015.pdf>.

- [58] D. Höller, P. Bercher, G. Behnke and S. Biundo, A Generic Method to Guide HTN Progression Search with Classical Heuristics, in: *Proceedings of the Twenty-Eighth International Conference on Automated Planning and Scheduling, ICAPS 2018, Delft, The Netherlands, June 24-29, 2018*, M. de Weerdt, S. Koenig, G. Röger and M.T.J. Spaan, eds, AAAI Press, pp. 114–122. <https://aaai.org/ocs/index.php/ICAPS/ICAPS18/paper/view/17764>.
- [59] M.C. Magnaguagno, F.R. Meneguzzi and L. de Silva, HyperTensioN: A three-stage compiler for planning, in: *Proceedings of the 30th International Conference on Automated Planning and Scheduling (ICAPS), 2020, França.*, 2020.
- [60] N. Gavenski, J. Monteiro, R. Granada, F. Meneguzzi and R. Barros, Imitating Unknown Policies via Exploration, in: *Proceedings of the Thirty First British Machine Vision Conference*, 2020.
- [61] P.R. Telang, F. Meneguzzi and M.P. Singh, Hierarchical Planning about Goals and Commitments, in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, 2013. <https://dl.acm.org/doi/abs/10.5555/2484920.2485059>.
- [62] F. Meneguzzi, P.R. Telang and M.P. Singh, A First-Order Formalization of Commitments and Goals, in: *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, 2013.
- [63] F. Meneguzzi, P. Telang and N. Yorke-Smith, Towards Planning Uncertain Commitment Protocols, in: *Proceedings of the Thirteenth International Conference on Autonomous Agents and Multiagent Systems*, pp. 1681–1682. <http://www.meneguzzi.eu/felipe/pubs/aamas-commitments-probabilities-2015.pdf>.
- [64] R.F. Pereira, A.S. Heinsfeld, A.R. Franco, A. Buchweitz and F. Meneguzzi, Detecting task-based fMRI compliance using plan abandonment techniques **5** (2016), 21–22.
- [65] K.B. Esper and F. Meneguzzi, Automated Design of fMRI Paradigms, in: *Proceedings of the Thirty-First International Conference on Automated Planning and Scheduling (ICAPS)*, AAAI Press, 2021, pp. 445–449.
- [66] M. Ghallab, D.S. Nau and P. Traverso, *Automated Planning and Acting*, Cambridge University Press, 2016. ISBN ISBN 978-1-107-03727-4.
- [67] B. Logan, J. Thangarajah and N. Yorke-Smith, Progressing Intention Progression: A Call for a Goal-Plan Tree Contest, in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)*, S. Das, E. Durfee, K. Larson and M. Winikoff, eds, IFAAMAS, Sao Paulo, Brazil, 2017, pp. 768–772, IFAAMAS.
- [68] Y. Yao, B. Logan and J. Thangarajah, SP-MCTS-based Intention Scheduling for BDI Agents, in: *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI-2014)*, T. Schaub, G. Friedrich and B. O’Sullivan, eds, IOS Press, Prague, Czech Republic, 2014, pp. 1133–1134, ECCAI.
- [69] Y. Yao and B. Logan, Action-Level Intention Selection for BDI Agents, in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, J. Thangarajah, K. Tuyls, C. Jonker and S. Marsella, eds, IFAAMAS, Singapore, 2016, pp. 1227–1235, IFAAMAS.
- [70] Y. Yao, B. Logan and J. Thangarajah, Robust Execution of BDI Agent Programs by Exploiting Synergies Between Intentions, in: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, D. Schuurmans and M.P. Wellman, eds, AAAI Press, Phoenix, USA, 2016, pp. 2558–2564, AAAI. <http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12148>.
- [71] Y. Yao, N. Alechina, B. Logan and J. Thangarajah, Intention Progression under Uncertainty, in: *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI 2020)*, C. Bessiere, ed., IJCAI, Yokohama, Japan, 2020.

- [72] Y. Yao, N. Alechina, B. Logan and J. Thangarajah, Intention Progression using Quantitative Summary Information, in: *Proceedings of the 20th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2021)*, U. Endriss, A. Nowé, F. Dignum and A. Lomuscio, eds, IFAAMAS, London, 2021, pp. 1416–1424, IFAAMAS.
- [73] B. Yun, S. Vesic and N. Oren, Representing pure Nash equilibria in argumentation, *Argument & Computation Preprint*(Preprint) (2021), 1–14, Publisher: IOS Press.
- [74] S. Modgil, Labellings and Games for Extended Argumentation Frameworks, in: *Proc. IJCAI-09*, 2009, pp. 873–878.
- [75] L.A. Dennis and N. Oren, Explaining BDI Agent Behaviour through Dialogue, in: *AAMAS '21: 20th International Conference on Autonomous Agents and Multiagent Systems, Virtual Event, United Kingdom, May 3-7, 2021*, F. Dignum, A. Lomuscio, U. Endriss and A. Nowé, eds, ACM, 2021, pp. 429–437. <https://dl.acm.org/doi/10.5555/3463952.3464007>.
- [76] Z. Shams and N. Oren, A two-phase dialogue game for skeptical preferred semantics, in: *European Conference on Logics in Artificial Intelligence*, Springer, 2016, pp. 570–576.
- [77] M.W.A. Caminada, S. Modgil and N. Oren, Preferences and unrestricted rebut, *Computational Models of Argument* (2014).
- [78] S. Vesic, B. Yun and P. Teovanovic, Graphical Representation Enhances Human Compliance with Principles for Graded Argumentation Semantics, 2021, International Conference on Autonomous Agents and Multi-Agent Systems 2022, AAMAS 2022 ; Conference date: 09-05-2022 Through 13-05-2022. <https://aamas2022-conference.auckland.ac.nz/>.
- [79] B. Yun, M. Croitoru, S. Vesic and P. Bisquert, DAGGER: Datalog+/- Argumentation Graph Generator., in: *Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS 2018, Stockholm, Sweden, July 10-15, 2018*, 2018, pp. 1841–1843. <http://dl.acm.org/citation.cfm?id=3237996>.
- [80] B. Yun, M. Croitoru and S. Vesic, NAKED: N-Ary graphs from Knowledge bases Expressed in Datalog+/-, in: *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS 2019, Montreal, Canada, May 13-17, 2019*, 2019.
- [81] B. Yun, S. Vesic and M. Croitoru, Sets of Attacking Arguments for Inconsistent Datalog Knowledge Bases, in: *Computational Models of Argument - Proceedings of COMMA 2020, Perugia, Italy, September 4-11, 2020*, H. Prakken, S. Bistarelli, F. Santini and C. Taticchi, eds, Frontiers in Artificial Intelligence and Applications, Vol. 326, IOS Press, 2020, pp. 419–430. doi:10.3233/FAIA200526.
- [82] M.W.A. Caminada, R. Kutlák, N. Oren and W.W. Vasconcelos, Scrutable plan enactment via argumentation and natural language generation, in: *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14, Paris, France, May 5-9, 2014*, A.L.C. Bazzan, M.N. Huhns, A. Lomuscio and P. Scerri, eds, IFAAMAS/ACM, 2014, pp. 1625–1626. <http://dl.acm.org/citation.cfm?id=2616095>.
- [83] F. Cerutti, C. Burnett and N. Oren, Representative agents and the cold start problem in contract negotiation, in: *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, Springer, 2015, pp. 42–58.
- [84] S.D. Kamvar, M.T. Schlosser and H. Garcia-Molina, The Eigentrust Algorithm for Reputation Management in P2P Networks, in: *Proceedings of the 12th International Conference on World Wide Web, WWW '03*, Association for Computing Machinery, New York, NY, USA, 2003, pp. 640–651-. ISBN 1581136803. doi:10.1145/775152.775242.

- [85] C. Burnett and N. Oren, Sub-Delegation and Trust, in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3, AAMAS '12*, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2012, pp. 1359–1360–. ISBN ISBN 0981738133.
- [86] N. Alechina, J.Y. Halpern and B. Logan, Causality, Responsibility and Blame in Team Plans, in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)*, S. Das, E. Durfee, K. Larson and M. Winikoff, eds, IFAAMAS, Sao Paulo, Brazil, 2017, pp. 1091–1099, IFAAMAS.
- [87] V. Yazdanpanah, M. Dastani, W. Jamroga, N. Alechina and B. Logan, Strategic Responsibility Under Imperfect Information, in: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, N. Agmon, M.E. Taylor, E. Elkind and M. Veloso, eds, IFAAMAS, Montreal, Canada, 2019, pp. 592–600, IFAAMAS. ISBN ISBN 978-1-4503-6309-9. <http://dl.acm.org/citation.cfm?id=3306127>.
- [88] N. Oren, S. Panagiotidi, J. Vázquez-Salceda, S. Modgil, M. Luck and S. Miles, Towards a formalisation of electronic contracting environments, in: *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, Springer, 2008, pp. 156–171.
- [89] M.P. Singh, Commitments Among Autonomous Agents in Information-Rich Environments, in: *MAA-MAW*, Lecture Notes in Computer Science, Vol. 1237, Springer, 1997, pp. 141–155.
- [90] N. Oren, W. Vasconcelos, F. Meneguzzi and M. Luck, Acting on norm constrained plans, in: *International Workshop on Computational Logic in Multi-Agent Systems*, Springer, 2011, pp. 347–363.
- [91] N. Alechina, M. Dastani and B. Logan, Reasoning about Normative Update, in: *Proceedings of the Twenty Third International Joint Conference on Artificial Intelligence (IJCAI)*, F. Rossi, ed., AAAI Press, Beijing, China, 2013, pp. 20–26, IJCAI.
- [92] J.F. Horty, Deontic Logic as Founded on Nonmonotonic Logic, *Ann. Math. Artif. Intell.* **9**(1–2) (1993), 69–91.
- [93] Z. Shams, M.D. Vos, N. Oren and J. Padget, Argumentation-based reasoning about plans, maintenance goals, and norms, *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* **14**(3) (2020), 1–39.
- [94] N. Alechina, J. Halpern, I. Kash and B. Logan, Incentive-Compatible Mechanisms for Norm Monitoring in Open Multi-Agent Systems, *Journal of Artificial Intelligence Research* **62** (2018), 433–458.
- [95] N. Alechina, J. Halpern, I. Kash and B. Logan, Incentivising Monitoring in Open Normative Systems, in: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI 2017)*, S.P. Singh and S. Markovitch, eds, AAAI Press, San Francisco, USA, 2017, pp. 305–311, AAAI. <http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14911>.
- [96] F. Meneguzzi, R. Fraga Pereira and N. Oren, Sensor Placement for Plan Monitoring Using Genetic Programming, in: *International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2018, pp. 544–551.
- [97] N. Alechina, M. Dastani and B. Logan, Norm Approximation for Imperfect Monitors, in: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*, A. Lomuscio, P. Scerri, A. Bazzan and M. Huhns, eds, IFAAMAS, Paris, France, 2014, pp. 117–124.
- [98] N. Alechina, N. Bulling, M. Dastani and B. Logan, Practical Run-Time Norm Enforcement with Bounded Lookahead, in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, R. Bordini, E. Elkind, G. Weiss and P. Yolum, eds, IFAAMAS, Istanbul, Turkey, 2015, pp. 443–451, IFAAMAS.

- [99] R.F. Pereira, N. Oren and F. Meneguzzi, Using sub-optimal plan detection to identify commitment abandonment in discrete environments, *ACM Transactions on Intelligent Systems and Technology (TIST)* **11**(2) (2020), 1–26.
- [100] N. Oren and F. Meneguzzi, Norm Identification through Plan Recognition, in: *15th International Workshop on Coordination, Organizations, Institutions, and Norms*, 2013, pp. 161–175. <http://www.meneguzzi.eu/felipe/pubs/coin-norm-detect-2013.pdf>.
- [101] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, Cambridge, UK, 2009. ISBN ISBN 978-0-521-89943-7.
- [102] A. Torreño, E. Onaindia, A. Komenda and M. Štolba, Cooperative Multi-Agent Planning: A Survey, *ACM Comput. Surv.* **50**(6) (2017). doi:10.1145/3128584.
- [103] D. Borrajo and S. Fernandez, Efficient approaches for multi-agent planning, *Knowledge and Information Systems* **58** (2019), 425–479. doi:10.1007/s10115-018-1202-1.
- [104] P.M. Dung, On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games, *Artif. Intell.* **77**(2) (1995), 321–358. doi:10.1016/0004-3702(94)00041-X.
- [105] D.N. Walton and E.C.W. Krabbe, *Commitment in Dialogue*, State University of New York Press, 1995.
- [106] B. Yun, P. Bisquert, P. Buche and M. Croitoru, Arguing About End-of-Life of Packagings: Preferences to the Rescue, in: *Metadata and Semantics Research - 10th International Conference, MTSR 2016, Göttingen, Germany, November 22-25, 2016, Proceedings*, E. Garoufallou, I.S. Coll, A. Stellato and J. Greenberg, eds, Communications in Computer and Information Science, Vol. 672, 2016, pp. 119–131. doi:10.1007/978-3-319-49157-8_10.
- [107] B. Yun, S. Vesic, M. Croitoru and P. Bisquert, Viewpoints Using Ranking-Based Argumentation Semantics, in: *Computational Models of Argument - Proceedings of COMMA 2018, Warsaw, Poland, 12-14 September 2018*, S. Modgil, K. Budzynska and J. Lawrence, eds, Frontiers in Artificial Intelligence and Applications, Vol. 305, IOS Press, 2018, pp. 381–392. doi:10.3233/978-1-61499-906-5-381.
- [108] E. Bonzon, J. Delobelle, S. Konieczny and N. Maudet, A Comparative Study of Ranking-Based Semantics for Abstract Argumentation., in: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, 2016, pp. 914–920. <http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12465>.
- [109] L. Amgoud and J. Ben-Naim, Ranking-Based Semantics for Argumentation Frameworks., in: *Scalable Uncertainty Management - 7th International Conference, SUM 2013, Washington, DC, USA, September 16-18, 2013. Proceedings*, 2013, pp. 134–147.
- [110] Z. Shams, M. De Vos, N. Oren and J. Padget, Normative practical reasoning via argumentation and dialogue, in: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016, pp. 1244–1250.
- [111] B. Liao, N. Oren, L. van der Torre and S. Villata, Prioritized norms in formal argumentation, *Journal of Logic and Computation* **29**(2) (2019), 215–240.
- [112] G. Ogunniye, A. Toniolo and N. Oren, Meta-Argumentation Frameworks for Multi-party Dialogues, in: *International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2018, pp. 585–593.
- [113] R. Riveret and N. Oren, On probabilistic argumentation and subargument-completeness, *Journal of Logic and Computation* **32**(1) (2022), 115–128.

- [114] H. Li, N. Oren and T.J. Norman, Probabilistic argumentation frameworks, in: *International Workshop on Theorie and Applications of Formal Argumentation*, Springer, 2011, pp. 1–16.
- [115] N. Oren, B. Yun, S. Vesic and M. da Silva Baptista, Inverse Problems for Gradual Semantics, in: *Proceedings of the 2022 International Joint Conference on Artificial Intelligence*, 2022.
- [116] J. Wu, H. Li, N. Oren and T.J. Norman, Gödel Fuzzy Argumentation Frameworks., in: *Proceedings of COMMA 2016*, 2016.
- [117] I. Rahwan, M.I. Madakkatel, J. Bonnefon, R.N. Awan and S. Abdallah, Behavioral Experiments for Assessing the Abstract Argumentation Semantics of Reinstatement, *Cogn. Sci.* **34**(8) (2010), 1483–1502. doi:10.1111/j.1551-6709.2010.01123.x.
- [118] M. Cramer and L. van der Torre, SCF2 - an Argumentation Semantics for Rational Human Judgments on Argument Acceptability, in: *Proceedings of the 8th Workshop on Dynamics of Knowledge and Belief (DKB-2019) and the 7th Workshop KI & Kognition (KIK-2019) co-located with 44nd German Conference on Artificial Intelligence (KI 2019), Kassel, Germany, September 23, 2019*, C. Beierle, M. Ragni, F. Stolzenburg and M. Thimm, eds, CEUR Workshop Proceedings, Vol. 2445, CEUR-WS.org, 2019, pp. 24–35. http://ceur-ws.org/Vol-2445/paper_3.pdf.
- [119] B. Yun, S. Vesic, M. Croitoru, P. Bisquert and R. Thomopoulos, A Structural Benchmark for Logical Argumentation Frameworks, in: *Advances in Intelligent Data Analysis XVI - 16th International Symposium, IDA 2017, London, UK, October 26-28, 2017, Proceedings*, 2017, pp. 334–346.
- [120] P.E. Dunne and M.J. Wooldridge, Complexity of Abstract Argumentation, in: *Argumentation in Artificial Intelligence*, G.R. Simari and I. Rahwan, eds, Springer, 2009, pp. 85–104. doi:10.1007/978-0-387-98197-0_5.
- [121] B. Yun, S. Vesic, M. Croitoru and P. Bisquert, Inconsistency Measures for Repair Semantics in OBDA., in: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, 2018, pp. 1977–1983. doi:10.24963/ijcai.2018/273.
- [122] N. Karanikolas, M. Croitoru, P. Bisquert, C. Kaklamanis, R. Thomopoulos and B. Yun, Multi-criteria Decision Making with Existential Rules Using Repair Techniques., in: *Artificial Intelligence XXXV - 38th SGA International Conference on Artificial Intelligence, AI 2018, Cambridge, UK, December 11-13, 2018, Proceedings*, 2018, pp. 177–183.
- [123] S. Maliah, G. Shani and R. Stern, Collaborative Privacy Preserving Multi-Agent Planning, *Autonomous Agents and Multi-Agent Systems* **31**(3) (2017), 493–530-. doi:10.1007/s10458-016-9333-9.
- [124] O. Boissier, R.H. Bordini, J. Hubner and A. Ricci, *Multi-Agent Oriented Programming: Programming Multi-Agent Systems Using JaCaMo*, Intelligent Robotics and Autonomous Agents series, MIT Press, 2020. ISBN ISBN 9780262360661. https://books.google.com.br/books?id=GM_tDwAAQBAJ.
- [125] A. Komenda, M. Stolba and D.L. Kovacs, The International Competition of Distributed and Multiagent Planners (CoDMAP), *AI Magazine* **37**(3) (2016), 109–115. doi:10.1609/aimag.v37i3.2658. <https://ojs.aaai.org/index.php/aimagazine/article/view/2658>.
- [126] F. Meneguzzi and R. F. Pereira, A Survey on Goal Recognition as Planning, in: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, International Joint Conferences on Artificial Intelligence Organization, 2021.
- [127] G. Sukthankar, R.P. Goldman, C. Geib, D.V. Pynadath and H.H. Bui, *Plan, Activity, and Intent Recognition: Theory and Practice*, Elsevier, 2014. ISBN ISBN 978-0123985323. <http://store.elsevier.com/Plan-Activity-and-Intent-Recognition/isbn-9780123985323/>.

- [128] R. Mirsky, S. Keren and C. Geib, Introduction to Symbolic Plan and Goal Recognition **16**(1) (2021), 1–190.
- [129] C.W. Geib, Problems with intent recognition for elder care, in: *Proceedings of the AAAI-02 Workshop "Automation as Caregiver"*, 2002. <https://www.aaai.org/Papers/Workshops/2002/WS-02-02/WS02-02-003.pdf>.
- [130] R. Granada, R.F. Pereira, J. Monteiro, R. Barros, D. Ruiz and F. Meneguzzi, Hybrid Activity and Plan Recognition for Video Streams, in: *The AAAI 2017 Workshop on Plan, Activity, and Intent Recognition*, 2017.
- [131] R. Inam, K. Raizer, A.Y. Hata, R.S. Souza, E. Fersman, E. Cao and S. Wang, Risk assessment for human-robot collaboration in an automated warehouse scenario. (2018).
- [132] R. Singh, T. Miller, J. Newn, E. Velloso, F. Vetere and L. Sonenberg, Combining gaze and AI planning for online human intention recognition **284** (2020), 103275. doi:10.1016/j.artint.2020.103275.
- [133] C. Wayllace, S. Ha, Y. Han, J. Hu, S. Monadjemi, W. Yeoh and A. Ottley, DRAGON-V: Detection and Recognition of Airplane Goals with Navigational Visualization, in: *AAAI*, 2020.
- [134] R.F. Pereira and F. Meneguzzi, Landmark-based Plan Recognition, in: *European Conference on Artificial Intelligence*, 2016, pp. 1706–1707. doi:10.3233/978-1-61499-672-9-1706. <http://ebooks.iospress.nl/publication/44992>.
- [135] R.F. Pereira, N. Oren and F. Meneguzzi, Landmark-Based Heuristics for Goal Recognition, in: *Proceedings of the 32st AAAI Conference on Artificial Intelligence*, 2017. <https://ojs.aaai.org/index.php/AAAI/article/view/11021>.
- [136] K.M.P. Gusmão, R.F. Pereira and F. Meneguzzi, The More the Merrier?! Evaluating the Effect of Landmark Extraction Algorithms on Landmark-Based Goal Recognition, in: *The AAAI 2020 Workshop on Plan, Activity, and Intent Recognition (PAIR@AAAI)*, 2020.
- [137] M. Vered, R.F. Pereira, M.C. Magnaguagno, G.A. Kaminka and F. Meneguzzi, Towards Online Goal Recognition Combining Goal Mirroring and Landmarks., in: *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2018. <https://dblp.org/rec/conf/atal/VeredPMKM18>.
- [138] R.F. Pereira, A.G. Pereira and F. Meneguzzi, Landmark-Enhanced Heuristics for Goal Recognition in Incomplete Domain Models, in: *Proceedings of the 29th International Conference on Automated Planning and Scheduling*, 2019.
- [139] R.F. Pereira, M. Vered, F. Meneguzzi and M. Ramírez, Online Probabilistic Goal Recognition over Nominal Models., in: *International Joint Conference on Artificial Intelligence*, International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 5547–5553. ISBN ISBN 978-0-9992411-4-1. doi:10.24963/ijcai.2019/770. <https://www.ijcai.org/proceedings/2019/770>.
- [140] L.R. de A. Santos, F. Meneguzzi, R.F. Pereira and A. Pereira, An LP-Based Approach for Goal Recognition as Planning, in: *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, AAAI Press, 2021.
- [141] L. Amado, J.P. Aires, R.F. Pereira, M.C. Magnaguagno, R. Granada, G.P. Licks and F. Meneguzzi, LatRec: Recognizing Goals in Latent Space (Demo), in: *Proceedings of the 29th International Conference on Automated Planning and Scheduling (ICAPS)*, AAAI Press, 2019.
- [142] R. Granada, R.F. Pereira, J. Monteiro, L. Amado, R. Barros, D. Ruiz and F. Meneguzzi, HAPRec: Hybrid Activity and Plan Recognizer, in: *The AAAI 2020 Workshop on Plan, Activity, and Intent Recognition (PAIR@AAAI): Demo Track*, 2020.

- [143] R. Granada, J. Monteiro, N. Gavenski and F. Meneguzzi, Object-Based Goal Recognition Using Real-World Data, in: *Proceedings of the 19th Mexican International Conference on Artificial Intelligence (MICA)*, 2020.
- [144] L.R. Amado and F. Meneguzzi, LatRec: Recognizing Goals in Latent Space, in: *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, AAAI Press, 2020.
- [145] L. Amado, J.P. Aires, R.F. Pereira, M.C. Magnaguagno, R. Granada, G.P. Licks, M. Marcon and F. Meneguzzi, LatRec+: Learning-based Goal Recognition in Latent Space (Demo), in: *The AAAI 2020 Workshop on Plan, Activity, and Intent Recognition (PAIR@AAAI): Demo Track*, 2020.
- [146] L.R. Amado, R.F. Pereira and F. Meneguzzi, Combining LSTMs and Symbolic Approaches for Robust Plan Recognition, in: *Proceedings of the Twentieth International Conference on Autonomous Agents and Multiagent Systems*, 2021.
- [147] L.R. Amado, R. Mirsky and F. Meneguzzi, Goal Recognition as Reinforcement Learning, in: *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, AAAI Press, 2022.
- [148] J. Monteiro, R. Granada, R.C. Barros and F. Meneguzzi, Deep neural networks for kitchen activity recognition, in: *2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2017, pp. 2048–2055. ISBN ISBN 978-1-5090-6182-2. doi:10.1109/IJCNN.2017.7966102. <http://ieeexplore.ieee.org/document/7966102/>.
- [149] R.L. Granada, J.P. Aires, J. Monteiro, F.R. Meneguzzi and R.C. Barros, Improving Action Recognition using Temporal Regions, *J. Inf. Data Manag.* **9**(2) (2018), 108–123. <https://sol.sbc.org.br/journals/index.php/jidm/article/view/2047>.
- [150] J. Hoffmann, Simulated Penetration Testing: From "Dijkstra" to "Turing Test++", in: *Proceedings of the International Conference on Automated Planning and Scheduling*, Vol. 25, 2015, pp. 364–372.
- [151] S. Keren, A. Gal and E. Karpas, Goal recognition design-survey, in: *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 2021, pp. 4847–4853.
- [152] M. Dann, J. Thangarajah, Y. Yao and B. Logan, Intention-Aware Multiagent Scheduling, in: *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)*, B. An, N. Yorke-Smith, A. El Fallah Seghrouchni and G. Sukthankar, eds, IFAAMAS, Auckland, 2020, IFAAMAS.
- [153] M. Dann, Y. Yao, B. Logan and J. Thangarajah, Multi-Agent Intention Progression with Black Box Agents, in: *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI 2021)*, Z.-H. Zhou, ed., IJCAI, Montreal, Canada, 2021.
- [154] B. Logan, An agent programming manifesto, *International Journal of Agent-Oriented Software Engineering* **6**(2) (2018), 187–210.
- [155] L. de Silva, F. Meneguzzi and B. Logan, BDI Agent Architectures: A Survey, in: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 2020, pp. 4914–4921.
- [156] R.H. Bordini, A.E.F. Seghrouchni, K.V. Hindriks, B. Logan and A. Ricci, Agent programming in the cognitive era, *Auton. Agents Multi Agent Syst.* **34**(2) (2020), 37. doi:10.1007/s10458-020-09453-y.
- [157] R.C. Cardoso and A. Ferrando, A Review of Agent-Based Programming for Multi-Agent Systems, *Computers* **10**(2) (2021), 16. doi:10.3390/computers10020016.

- [158] R.C. Cardoso, R.F. Pereira, G. Krzisch, M.C. Magnaguagno, T. Baségio and F. Meneguzzi, Team PUCRS: a Decentralised Multi-Agent Solution for the Agents in the City Scenario, *International Journal of Agent-Oriented Software Engineering* **6**(1) (2018), 3–34. doi:10.1504/IJAOSE.2018.10010601. <https://www.inderscienceonline.com/doi/abs/10.1504/IJAOSE.2018.089595>.
- [159] T. Krausburg, R.C. Cardoso, J. Damasio, V. Peres, G.P. Farias, D.C. Engelmann, J.F. Hübner and R.H. Bordini, SMART–JaCaMo: An Organisation-Based Team for the Multi-Agent Programming Contest, in: *The Multi-Agent Programming Contest 2018*, T. Ahlbrecht, J. Dix and N. Fiekas, eds, Springer International Publishing, Cham, 2019, pp. 72–100. ISBN ISBN 978-3-030-37959-9.
- [160] R.C. Cardoso, A. Ferrando and F. Papacchini, LFC: Combining Autonomous Agents and Automated Planning in the Multi-Agent Programming Contest, in: *The Multi-Agent Programming Contest 2019*, Springer International Publishing, Cham, 2020, pp. 31–58. ISBN ISBN 978-3-030-59299-8. doi:10.1007/978-3-030-59299-8_2.
- [161] R.C. Cardoso, A. Ferrando, F. Papacchini, M. Luckcuck, S. Linker and T.R. Payne, MLFC: From 10 to 50 Planners in the Multi-Agent Programming Contest, in: *The Multi-Agent Programming Contest 2021*, T. Ahlbrecht, J. Dix, N. Fiekas and T. Krausburg, eds, Springer International Publishing, Cham, 2021, pp. 82–107. ISBN ISBN 978-3-030-88549-6. doi:10.1007/978-3-030-88549-6_4.
- [162] R.C. Cardoso, A. Ferrando and F. Papacchini, Automated Planning and BDI Agents: A Case Study, in: *Advances in Practical Applications of Agents, Multi-Agent Systems, and Social Good. The PAAMS Collection*, F. Dignum, J.M. Corchado and F. De La Prieta, eds, Springer International Publishing, Cham, 2021, pp. 52–63. ISBN ISBN 978-3-030-85739-4. doi:10.1007/978-3-030-85739-4_5.
- [163] M. Luckcuck and R.C. Cardoso, Formal Verification of a Map Merging Protocol in the Multi-agent Programming Contest, in: *Engineering Multi-Agent Systems*, N. Alechina, M. Baldoni and B. Logan, eds, Springer International Publishing, Cham, 2022, pp. 198–217. ISBN ISBN 978-3-030-97457-2. doi:10.1007/978-3-030-97457-2_12.
- [164] I. Rahwan, I. Madakkatel, J.-F. Bonnefon, N. Awan and S. Abdallah, Behavioral Experiments for Assessing the Abstract Argumentation Semantics of Reinstatement, *Cognitive science* **34** (2010), 1483–502. doi:10.1111/j.1551-6709.2010.01123.x.
- [165] S. Vesic, B. Yun and P. Teovanovic, Graphical Representation Enhances Human Compliance with Principles for Graded Argumentation Semantics, 2021, International Conference on Autonomous Agents and Multi-Agent Systems 2022, AAMAS 2022 ; Conference date: 09-05-2022 Through 13-05-2022. <https://aamas2022-conference.auckland.ac.nz/>.
- [166] S. Cranefield, N. Oren and W.W. Vasconcelos, Accountability for practical reasoning agents, in: *International conference on agreement technologies*, Springer, 2018, pp. 33–48.
- [167] H. Trivedi, The Business of Artificial Intelligence in Radiology Has Little to Do With Radiologists, *Journal of the American College of Radiology* **19**(4) (2022), 564–566.
- [168] R. Brooks, A Human in the Loop: AI won’t Surpass Human Intelligence Anytime Soon, *IEEE Spectrum* **58**(10) (2021), 48–49. doi:10.1109/MSPEC.2021.9563963.
- [169] J. Monteiro, J.P. Aires, R. Granada, R.C. Barros and F. Meneguzzi, Virtual guide dog: An application to support visually-impaired people through deep convolutional neural networks, in: *2017 International Joint Conference on Neural Networks, IJCNN 2017, Anchorage, AK, USA, May 14-19, 2017*, IEEE, 2017, pp. 2267–2274. doi:10.1109/IJCNN.2017.7966130.
- [170] M. de Sousa Ribeiro and J. Leite, Aligning Artificial Neural Networks and Ontologies towards Explainable AI, in: *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium*

on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021, AAAI Press, 2021, pp. 4932–4940. <https://ojs.aaai.org/index.php/AAAI/article/view/16626>.

- [171] R.H. Bordini, A. El Fallah Seghrouchni, K. Hindriks, B. Logan and A. Ricci, Agent programming in the cognitive era, *Autonomous Agents and Multi-Agent Systems* **34**(37) (2020). doi:<https://doi.org/10.1007/s10458-020-09453-y>.
- [172] H. Geffner, The Model-Based Approach to Autonomous Behavior: A Personal View, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-2010)*, AAAI Press, 2010, pp. 1709–1712.