Report from Dagstuhl Seminar 14332

Formal Methods for Coordinating Multi-Agent Systems

Edited by

Thomas Ågotnes[1] and Nils Bulling[2]

1 **University of Bergen, NO,** thomas.agotnes@infomedia.uib.no
2 **TU Clausthal, DE,** bulling@in.tu-clausthal.de

Abstract

This report documents the programme and outcomes of the Dagstuhl Serminar 14332 "Formal Methods for Coordinating Multi-Agent Systems", that took place from 10 to 14 August, 2014. This seminar brought together researchers from the following subfields of multi-agent systems: logic, game theory, and agreement technologies. It is set up at the intersection of these active fields of research and aimed at fostering collaborations between them. A key objective of the seminar has been to shed light on formal methods for coordinating multi-agent systems, in particular, how to combine research and tools from the different areas to obtain new techniques for coordinating the behavior of agents. The *coordination problem* is a key problem in multi-agent systems: how can we coordinate the individual behaviour of the agents such that the global behaviour of the system as a whole satisfies our needs? Dagstuhl was an excellent venue to bring together leading researchers from logics, game theory, and agreement technologies to learn about their research activities, to discuss as well as to work on timely problems, and to establish new collaborations between researchers. The outcome of the working groups and discussions provides promising avenues and open questions for future research in the field.

Seminar August 10–14, 2014 – http://www.dagstuhl.de/14332
1998 ACM Subject Classification I.2.11 Distributed Artificial Intelligence – Multiagent systems, Intelligent agents
Keywords and phrases multi-agent systems, coordination, formal methods, game theory, logic, agreement technologies
Digital Object Identifier 10.4230/DagRep.4.8.21
Edited in cooperation with Marija Slavkovic

1 Executive Summary

Thomas Ågotnes
Nils Bulling

Formal methods form an active and broad field of research in multi-agent systems, ranging from bottom-up to top-down approaches. Properties of individual agents, e. g., aspects related to decision making and knowledge representation, are rather low-level, while the specification and verification of multi-agent systems are higher-level. In particular, *logic-based approaches* have been successfully used for the modeling of intelligent agents and for reasoning about them: epistemic logics allow to talk about knowledge; temporal logics to reason about the evolution of actions; and strategic logics have been proposed to reason about abilities of agents and coalitions. Alternating-time temporal logics and STIT logics are prominent members of the latter type, and more expressive logics like Strategy Logic have recently been proposed. What they all have in common is their descriptive flavor. Typically, they are not used to actively change the state of the agent system but to talk and to reason about the system. Multi-agent logics are particularly relevant for the coordination problem. The latter is concerned with global properties of a system. Since the global behaviour of a system emerges from the individual behaviour of agents, it is not obvious what the global properties are. By specifying global properties using multi-agent logics, verification techniques can be employed to verify what of these properties are met by the system; thus, to find out what the global properties are. Interaction between rational decision makers in general, and coordination problems in particular, have been studied in game theory for decades. However, game theory is not concerned with *computational* or *logical* aspects of coordination: how we can represent and reason about coordination in computers. In contrast, many *agreement technologies* are used in an interactive way, e. g., for arriving at agreements about joint actions or coalition structures. Techniques like norms and social laws coordinate the agent's behavior and often require less interaction of agents with their peers. Agents have to decide whether to comply with the rules or not. A difficult problem is the synthesis of appropriate norms and social laws. Related issues important for appropriate control techniques include the detection of norm violations and sanctioning mechanisms.

The seminar aimed at opening up new directions of research into the coordination problem, by bringing together researchers working in different areas of multi-agent systems as well as related fields, and in particular, to combine insights from research in the following fields:

- formal methods and verification, and multi-agent logics in particular,
- game theory in multi-agent systems, and
- agreement technologies.

The seminar took place between 10 and 14 August, 2014. This medium-size, four day seminar was highly international: the 27 participants came from 12 different countries. The scientific program consisted of presentations, discussions and working groups. We scheduled presentations of three different types: overview, medium, and short. The aim of the four one hour overview talks was to give a broad introduction of the main fields relevant to the seminar – to provide a common ground. They covered Argumentation Theory, Normative Systems, Judgement Aggregation, and Computational Social Choice. Then, we scheduled ten medium (20 minutes long) and ten short (15 minutes long) presentations. We encouraged the speakers to give rather informal, non conference-style talks focussing on high-level ideas in order to provide input for the discussion groups.

From the discussions, two working groups emerged which focused on one of the following topics (cf. Sections 4.1 and 4.2):

- Concepts: conceptual definition and classification – what is coordination, coordination problems, and solutions?
- Formalisation of coordination

We organized three meetings for the working groups and two joint discussion sessions for presenting and discussing the results of the working groups.

In addition to the scientific program, we enjoyed a hike which was followed a Barbecue, and the unique atmosphere of Dagstuhl.

2 Table of Contents

3 Overview of Talks

3.1 Strategic Voting and Strategic Candidacy

Markus Brill (Duke University, US)

Joint work of Brill, Markus; Conitzer, Vincent
Main reference M. Brill, V. Conitzer, "Strategic Voting and Strategic Candidacy," in Proc. of the 29th AAAI Conf. on Artificial Intelligence (AAAI'15), to appear; pre-print available from author's webpage.
URL http://www.cs.duke.edu/~brill/papers/candidacy.pdf

Models of strategic candidacy analyze the incentives of candidates to run in an election. Most work on this topic assumes that strategizing only takes place among candidates, whereas voters vote truthfully. In this paper, we extend the analysis to also include strategic behavior on the part of the voters. (We also study cases where only candidates or only voters are strategic.) We consider two settings in which strategic voting is well-defined and has a natural interpretation: majority-consistent voting with single-peaked preferences and voting by successive elimination. In the former setting, we analyze the type of strategic behavior required in order to guarantee desirable voting outcomes. In the latter setting, we determine the complexity of computing the set of potential outcomes if both candidates and voters act strategically.

3.2 Many questions on the semantics of AT(E)L(*)-type logics and some answers

Jan M. Broersen (Utrecht University, NL)

I propose to analyse the semantics of AT(E)L(*) type logics using standard tree-based semantics with epistemic indistinguishability relations. The aim is to come to a unified view on (1) memory, (2) uncertainty and (3) uniformity of strategies in the semantics for these logics. The new setting should also enable one to develop a theory of strategic ability that includes strategies with observations.

3.3 Agents with Perfect and Truly Perfect Recall

Nils Bulling (TU Clausthal, DE)

Joint work of Bulling, Nils; Jamroga, Wojciech; Popovici, Matei
Main reference N. Bulling, W. Jamroga, M. Popovici, "ATL* with truly perfect recall: Expressivity and validities," in Proc. of the 21st European Conf. on Artificial Intelligence (ECAI'14), Frontiers in Artificial Intelligence and Applications, Vol. 263, pp. 177–182, IOS Press, 2014.
URL http://dx.doi.org/10.3233/978-1-61499-419-0-177

In alternating-time temporal logic ATL* [1], agents with perfect recall assign choices to sequences of states, i.e., to possible finite histories of the game. However, when a nested strategic modality is interpreted, the new strategy does not take into account the previous sequence of events. It is as if agents collect their observations in the nested game again from scratch, thus effectively forgetting what they observed before. Intuitively, it does not fit the assumption of agents having perfect recall of the past.

In this talk I shall review ATL and its semantic variants. Then, I present a new semantics for ATL* where the past is not forgotten in nested games [3, 4]. I give a formal treatment and show that the standard semantics of ATL* coincides with the new semantics in case of agents with perfect information. On the other hand, both semantics differ if agents have imperfect information about the state of the game. I compare the expressivity of the logics and their sets of validities; the latter characterizes general properties of the underlying classes of games (cf. [2]).

References

1 R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002.

2 Nils Bulling and Wojciech Jamroga. Comparing variants of strategic ability: how uncertainty and memory influence general properties of games. *Journal of Autonomous Agents and Multi-Agent Systems*, 28(3):474–518, 2014.

3 Nils Bulling, Wojciech Jamroga, and Matei Popovici. Agents with truly perfect recall in alternating-time temporal logic (extended abstract). In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2014)*, pages 1561–1562, Paris, France, May 2014. ACM Press.

4 Nils Bulling, Wojciech Jamroga, and Matei Popovici. ATL* with truly perfect recall: Expressivity and validities. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI 2014)*, pages 177–182, Prague, Czech Republic, August 2014.

3.4 Reasoning About Norms under Uncertainty in Dynamic Environments

Natalia Criado (John Moores University – Liverpool, GB)

One of the main goals of the agent community is to provide a trustworthy technology that allows humans to delegate some specific tasks to software agents. Frequently, laws and social norms regulate these tasks. As a consequence, agents need mechanisms for reasoning about these norms. Up until now the existing proposals on normative agents assume that agents interact within a deterministic environment that is certainly perceived. In practice, agents interact by means of sensors and actuators under uncertainty with non-deterministic and dynamic environments. In this talk, I presented my work on normative agents that are able to deal with uncertainty in dynamic environments

3.5 Norm-based Coordination in Multi-Agent Systems

Mehdi Dastani (Utrecht University, NL)

Norms have been widely proposed as a means of coordinating the behaviours of agents. This presentation discusses some application areas where norms can be used for coordination purposes. A programming view on norms will be provided and it will be explained how programmed norms can be enforced and regimented by means of sanctions. In order to study the behaviours of norm programs, a logical analysis will be provided that allows reasoning about the behaviours of multi-agent systems under norm enforcement. We also discuss the enforcement of norms by multiple sanctions.

References

1 Mehdi Dastani and Davide Grossi and John-Jules Meyer. A Logic for Normative Multi-Agent Programs. Journal of Logic and Computation, 2011.
2 Max Knobbout and Mehdi Dastani. Reasoning under Compliance Assumptions in Normative Multiagent Systems. Eleventh International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2012.
3 Nils Bulling and Mehdi Dastani and Max Knobbout. Monitoring Norm Violations in Multi-Agent Systems, Twelfth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2013.
4 Natasha Alechina and Mehdi Dastani ad Brian Logan. Norm Approximation for Imperfect Monitors. Thirteenth International Conference Autonomous Agents and Multi-Agent Systems (AAMAS), 2014.

3.6 Verifying Agents that Plan

Louise A. Dennis (University of Liverpool, GB)

Joint work of Dennis, Louise A.; Fisher, Michael

We are interested in the issue of "graceful degradation" of BDI based autonomous systems. If such as system gracefully degrades then it can adapt to the failure of some of its components, for instance by modifying its goals, adapting its plans, or requesting assistance from other agents. A key part of such a capability is likely to be access to planning sub-systems.

This raises some interesting issues for the verification of such systems. This will be a largely speculative short talk, exploring some of those issues in a preliminary way.

3.7 Automata techniques for temporal epistemic logics

Catalin Dima (University Paris-Est – Créteil, FR)

Joint work of Bozianu, Rodica; Dima, Catalin; Filiot, Emmanuel; Maubert, Bastien; Pinchinat, Sophie
Main reference R. Bozianu, C. Dima, E. Filiot, "Safraless Synthesis for Epistemic Temporal Specifications," in Proc. of the 26th Int'l Conf. on Computer Aided Verification (CAV'14), LNCS, Vol. 8559, pp. 441-456, Springer, 2014.
URL http://dx.doi.org/10.1007/978-3-319-08867-9_29
Main reference C. Dima, B. Maubert, S. Pinchinat, "The Expressive Power of Epistemic mu-Calculus," arXiv:1407.5166v1 [cs.LO], 2014.
URL http://arxiv.org/abs/1407.5166v1

The classical duality between automata and logics lies at the foundations of many results on decidability and expressivity of modal and temporal logics, with applications in model-checking and synthesis. It's therefore natural to investigate the possibility of extending this duality to the case of temporal epistemic logics. We present two recent results in which automata techniques and their relation with temporal epistemic logics play an essential role: a non-expressivity result concerning ATL with imperfect information and the epistemic mu-calculus, and an implemented technique for controller synthesis from temporal epistemic goals.

3.8 Sharing information in teams: what, when and with whom?

Maaike Harbers (TU Delft, NL)

A multi-agent system with a joint goal, a team, can increase its performance by sharing information between team members. For instance, sharing information helps to anticipate each other's actions, plan own activities efficiently, and help each other when possible and needed. However, sharing more information is not always better. Agents may have a limited capacity for sending and processing information, and the exchange of certain information can lead to privacy loss. In this talk, I will discuss a formal model of teamwork that captures positive and negative effects of information sharing in teams. The model allows for the exploration of different information sharing policies, thus benefiting the design of effective teams.

3.9 Exploiting Speculative Computation with Defeasible Reasoning in Multi-Agent System

Ho-Pun Lam (NICTA – Brisbane, AU)

Joint work of Lam, Ho-Pun; Governatori, Guido; Satoh, Ken; Hosobe, Hiroshi
Main reference H.-P. Lam, G. Governatori, K. Satoh, H. Hosobe, "Distributed Defeasible Speculative Reasoning in Ambient Environment," in Proc. of the 13th Int'l Workshop on Computational Logic in Multi-Agent Systems (CLIMA'12), LNCS, Vol. 7486, pp. 43–60, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-32897-8_5

Speculative Computation is an effective means for solving problems with incomplete information in an open and distributed environment, such as peer- to-peer environment. It allows such a system to compute tentative (and possibly final) solutions using default knowledge about the current environment, or the agent's perception, even if the communications between peers are delayed or broken. However, previous work in speculative reasoning assumed that agents are hierarchically structured, which may not be the case in reality. We propose a more general multi-agents system with no centralized control. Agents in the framework have equivalent functionalities and can collaborate with each other to achieve their common goals. We characterize the framework using the argumentation semantics of defeasible logic, which provides support of speculative reasoning in the presence of conflicting information, and provide an operational model for the framework.

3.10 Formal Argumentation and Its Roles in Multi-Agent Systems

Beishui Liao (Zhejiang University, CN)

Formal argumentation is an increasingly active research area in artificial intelligence. In this talk, after discussing some properties of various kinds of reasoning in multi-agent systems and the limitations of first-order logic and traditional nonmonotonic formalisms, I briefly introduce formal argumentation, and some research directions of formal argumentation in multi-agent systems.

3.11 Fair allocation of group tasks according to social norms

Brian Logan (University of Nottingham, GB)

Joint work of Alechina, Natasha; van der Hoek, Wiebe; Logan, Brian
Main reference N. Alechina, W. van der Hoek, B. Logan, "Fair Allocation of Group Tasks According to Social Norms," in Proc. of the 15th Int'l Workshop on Computational Logic in Multi-Agent Systems (CLIMA'14), LNCS, Vol. 8624, pp. 19–34, Springer, 2014.
URL http://dx.doi.org/10.1007/978-3-319-09764-0_2

This talk considers the problem of decomposing a group norm into a set of individual obligations for the agents comprising the group, such that if the individual obligations are fulfilled, the group obligation is fulfilled. The group norms we consider may be non-repeating or repeating (e. g., a group obligation that should be discharged each week). We assume that the assignment of tasks in a group norm to agents is subject to additional social or organisational norms that specify permissible ways in which tasks can be assigned. An important type of social norms are 'fairness constraints', that seek to distribute individual responsibility for discharging the group norm in a 'fair' or 'equitable' way, e. g., an agent may be required to perform a particular task no more than once a week. I briefly present our initial attempts to formalise both group norms and social norms/fairness constraints, and highlight some open problems.

3.12 Towards future road networks: considering traffic system's fairness trap

Marin Lujak (University Rey Juan Carlos, ES)

Joint work of Lujak, Marin; Giordani, Stefano; Ossowski, Sascha
Main reference M. Lujak, S. Giordani, S. Ossowski, "Fair route guidance: bridging system and user optimization," in Proc. of the 17th Int'l IEEE Conf. on Intelligent Transportation Systems (ITSC'14), IEEE, to appear.

In this talk we study the problem of the assignment of road paths to vehicles. Due to the assumption that a low percentage of vehicles follow the routes proposed by route guidance systems (RGS) and the increase of the use of the same, the conventional RGS might shortly result obsolete. Assuming a complete road network information at the disposal of RGSs, their proposed paths are related with user optimization which in general can be arbitrarily more costly than the system optimum. However, the user optimum is fair for the drivers on the same Origin-Destination (O-D) but it doesn't guarantee fairness for different O-D pairs. Contrary, the system optimum can produce unfair assignments both for the vehicles of the same as of different O-D pairs. This is the reason why, in this talk, we present and discuss an optimization model which bridges this gap between the user and system optimum, and propose a new mathematical programming formulation based on Nash Welfare optimization which results in a good egalitarian and utilitarian welfare for all O-D pairs. Furthermore, we discuss issues related with the proposed distributed approach based on multi-agent system principles and paradigms.

3.13 Practical Reasoning, Norms and Argument

Nir Oren (University of Aberdeen, GB)

Joint work of Oren, Nir; van Riemsdijk, Birna

This talk describes a model of normative system based on a transition system wherein agents act based on goals and directed norms. These impose a partial preference relation over paths through the system, from which an equilibrium can be computed. I then map the preference relations for an agent to argument schemes, from which an argumentation framework for the system can be derived; evaluating this system aims to again identify system equilibria, and the arguments themselves can also be used to explain the functioning of the system.

3.14 An action language approach to normative specification, analysis and revision

Julian Padget (University of Bath, GB)

Joint work of Li, Tingting; Balke, Tina; De Vos, Marina; Padget, Julian; Satoh, Ken
Main reference T. Li, T. Balke, M. De Vos, J. Padget, K. Satoh, "Legal conflict detection in interacting legal systems," in Kevin D. Ashley, editor, "Legal Knowledge and Information Systems," Frontiers in Artificial Intelligence and Applications, Vol. 259, pp. 107–116. IOS Press, 2013.
URL http://dx.doi.org/10.3233/978-1-61499-359-9-107

InstAL (Institutional Action Language) is both a declarative domain-specific language for the specification of collections of interacting normative systems and a framework for a set of associated tools. The computational model is realized by translating the specification language to AnsProlog, a logic programming language under the answer set semantics (ASP), and is underpinned by a set-theoretic formal model and a formalized translation process. Among its notable features are: (i) non-inertial fluents, which allow the creation of institutional fluents that hold for as long as some condition over existing institutional facts is true, (ii) interacting institutions, which allow for events in one institution to affect another and support the modularization of institutional specifications, and (iii) durations, which allow a fluent to hold at some number of time steps in the future. A recent substantive extension is the means to carry out conflict detection and resolution for coordinated institutions. This enables the specification of normative positions through use cases comprising a (partial) sequence of events and a (partial) state description and the consequent synthesis using inductive logic programming of a minimal self-consistent rule set.

3.15 Anonymity in ATL – Strategic Homogeneity

Truls Pedersen (University of Bergen, NO)

We discuss anonymous alternating-time temporal logic (ATL) which is ATL interpreted over symmetric concurrent game structures (CGSs). By symmetry in the model, we mean that all permutations of any action profile leads to bisimilar states. Symmetry in the model permit representation for which the model checking problem is no longer exponential in the number of agents. We also give a complete axiomatization of anonymous ATL over symmetric models. As the symmetry assumption is very strong, we indicate how we can reintroduce heterogeneity in the model by roles without seriously affecting the complexity of the model checking problem.

3.16 Modal Logic for Mixed Strategies in Games

Joshua Sack (University of Amsterdam, NL)

Joint work of Sack, Joshua; van der Hoek, Wiebe
Main reference J. Sack, W. van der Hoek, "A Modal Logic for Mixed Strategies," Studia Logica, 102(2):339–360, 2014.
URL http://dx.doi.org/10.1007/s11225-014-9548-1

Mixed strategies are useful for reasoning about equilibria in games. Although it is true that every finite strategic form game has a Nash equilibrium, this is not true if one restricts oneself to pure strategies, as can be exemplified by the Matching Pennies Game. This talk introduces Modal Logic for Mixed Strategies, a modal logic that can reason about mixed strategies and mixed Nash equilibria in games. A sound and strongly complete proof system for it is given that makes use of a number of non-standard infinitary rules.

3.17 Abstract Formal Basis for Digital Crowds

Marija Slavkovik (University of Bergen, NO)

Joint work of Slavkovik, Marija; Dennis, Louise; Fisher, Michael
Main reference M. Slavkovik, L. A. Dennis, M. Fisher, "An Abstract Formal Basis for Digital Crowds," arXiv:1408.1592v1 [cs.LO], 2014.
URL http://arxiv.org/abs/1408.1592v1

Crowdsourcing, together with its related approaches, has become very popular in recent years. All crowdsourcing processes involve the participation of a digital crowd, a large number of people that access a single Internet platform or shared service. In this paper we explore the possibility of applying formal methods, typically used for the verification of software and hardware systems, in analysing the behaviour of a digital crowd. More precisely, we provide a formal description language for specifying digital crowds. We represent digital crowds in which the agents do not directly communicate with each other. We further show how this specification can provide the basis for sophisticated formal methods, in particular formal verification.

3.18 Judgment Aggregation – an overview

Marija Slavkovik (University of Bergen, NO)

Judgment aggregation is a social choice theory concerned with aggregation of collection of judgments, or truth-value assignments, made for a set of logically related issues. This talk was a general overview of the field. It presumed that the participants are entirely unfamiliar with judgment aggregation and somewhat familiar with social choice theory. The talk focused on answering the following questions: what is judgment aggregation, how does it relate to the better known voting theory, what are the points of interest from a MAS and AI perspective, and what are the interesting open problems in this area.

3.19 Resource-sensitive interactions

Nicolas Troquard (National Research Council – Povo (Trento), IT)

I insist on the importance of considering resources in multi-agent interactions. I present some motivations with examples in socio-technical systems. I introduce a resource-sensitive logic of agency in a more technical part.

3.20 Pre-vote negotiations and voting games

Paolo Turrini (Imperial College London, GB)

Joint work of Grandi, Umberto; Grossi, Davide; Turrini, Paolo
Main reference U. Grandi, D. Grossi, P. Turrini, "Pre-vote negotiations and binary voting with constraints," arXiv:1404.5433v1 [cs.GT], 2014.
URL http://arxiv.org/abs/1404.5433v1

I have talked about voting games on possibly interconnected issues, where voters might hold a principled opinion about a subset of the issues at stake while willing to strike deals on the remaining ones, and can influence one another before casting their ballots in order to obtain an individually more favourable outcome. We analyse voters' rational behaviour in a two-phase game, allowing players to undergo a negotiation phase before their vote, and showing under what conditions undesirable equilibria can be removed as an effect of the pre-vote phase. What I presented is joint work with Davide Grossi (Liverpool) and Umberto Grandi (Toulouse).

3.21 Computational Reasoning for Socially Adaptive Electronic Partners

Birna Van Riemsdijk (TU Delft, NL)

Main reference M. B. van Riemsdijk, L. Dennis, M. Fisher, K. V. Hindriks, "Agent reasoning for norm compliance: a semantic approach," in Proc. of the 12th Int'l Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS'13), pp. 499–506, IFAAMAS, 2013; pre-print available from author's webpage.
URL http://dl.acm.org/citation.cfm?id=2485000
URL http://ii.tudelft.nl/~birna/publications/2013/riemsdijk13aamas.pdf

Technology is becoming an integral part of our daily lives. To ensure that this process unfolds with sufficient respect for important values such as privacy and freedom, I develop software that adapts to norms and values of people. In this talk I introduce a new computational reasoning framework for Socially Adaptive Electronic Partners (SAEPs) that support people in their daily lives without people having to adapt their way of living to the software. The computational reasoning techniques are aimed at determining when and to what extent norm-compliance can be guaranteed, and deciding what to do if in exceptional situations a norm cannot be complied with. The reasoning framework is based on executable temporal logic, integrating the agent's execution semantics with adopted norms for guaranteeing compliance.

3.22 Collective Intention Revision from a Database Perspective

Marc Van Zee (University of Luxembourg, LU)

Joint work of Van Zee, Marc; Dastani, Mehdi; Van der Torre, Leon; Shoham, Yoav
Main reference M. van Zee, M. Dastani, Y. Shoham, L. van der Torre, "Collective Intention Revision from a Database Perspective," in Proc. of the Collective Intentionality Conference 2014; pre-print available from author's webpage.
URL http://icr.uni.lu/marc/publications/ci2014.pdf

Icard et al. recently formalized Shoham's "database perspective" with a logical model to capture action, belief and intention. We extend this model to a multi-agent setting by introducing a collective intention base that captures dependencies between intentions of different agents. We provide AGM-like postulates for multi-agent revision of beliefs, individual intentions, and collective intentions, and conjecture a representation theorem relating our postulates to the formal model.

3.23 Adaptation of social control and trust mechanisms

Laurent Vercouter (INSA – St-Etienne-du-Rouvray, FR)

Multiagent systems offer an interesting approach to deal with large scale distributed systems as it proposes to focus on local design and development of agents on on their interaction rather than trying to consider the system as a whole. However, it is not an easy task to define control mechanisms ensuring that the global behavior comply to what is expected. Agreement technologies and more specifically multiagent trust and reputation models brings solution for controlling such networks when it is not possible to get a global and complete observation of the agents' behaviors. Following the early proposition of Castelfranchi & Falcone, social control mechanisms have been developed these last years in the multiagent community as a way to provide to artificial agent societies the means to control itself using trust and reputation concepts in order to allow each agent to observe, evaluate and sanction its neighborhood. This talk will present the general issues involved in the development of social control mechanisms. We will present some specific concrete examples of their applications to decentralized systems (sensor networks, peer-to-peer systems, social networks). We will consider especially situations where uncertainty (in the agents' identity or in the definition of expectations) prevents the use of classical control mechanisms and for which adapted solutions have been developped. Finally, we will present our view of open issues and perspectives for this topic.

4 Working Groups

There were four working group discussion sessions of about 75 minutes each, including an initial meeting where the organizers proposed general topics, centered around coordination, for the discussions groups. They were then concretized by the seminar participants. After some general discussion among all seminar participants the organizers separated the seminar into two groups of about equal size. The outcomes of the working groups were presented to the whole seminar at two wrap-up sessions, in the middle and at the very end of the seminar.

For the first meeting of the plenary discussion group, the seminar participants identified interesting sub-topics of the general topic. These were then, after some discussion, divided into two general topics:

- Concepts: conceptual definition and classification – what is coordination, coordination problems, and solutions?
- Formalisation of coordination

The following two subsections are reports that sums up the outcomes of the two working groups, edited by the respective working group coordinator(s).

4.1 Working group: Concepts – What is coordination?

Joint work of working group members; edited by: Louise Dennis (group coordinator)

4.1.1 Introduction

The working group began by considering the question "What is a Coordination Problem" and then worked on understanding the space of such problems and their solutions.

4.1.2 What is a coordination problem?

▶ **Definition 1** (Coordination Problem). Given a purpose/requirement and n agents (where $n > 1$), where one of those agents alone may not achieve the purpose then the complete solution to the coordination problem is a system which guarantees the achievement of the purpose.

Partial solutions

- achieve compromises,
- increase the chance of achieving goals

4.1.3 Purpose

We assume that a coordination is designed by some person for the purpose of achieving some thing.

Requirements Ultimately the system designed wants to realise the requirements. They may be informal – in fact they will probably be informal.

Motivations Motivations are a formal set of expressions that explicitly state a purpose of the coordination. If the requirements were formal then the motivations can equal the requirements. Motivations may be exogenous to the system (i. e., nothing in the system has any explicit awareness of them) or they may be endogenous to the system (i. e., at least one entity in the system is aware of the motivations).
We note that motivations can be engineered, particularly if they are to be endogenous to the system. i. e., finding a suitable formal expression for the requirements may be a part of the solution.

Soft Constraints Additionally there may be a number of properties which are desirable in the system but which are not required for a solution. "Optimal" solutions (and compromises) will optimise these soft constraints.

Goals

Note: Both motivations and soft constraints are referred to as *goals*. A *goal* is therefore an explicit formal property that we wish the system to have.

We categorise goals into three types.

A Type A goals are discrete yes/no properties. The system either has them or it does not.
B Type B goals are metrics with a threshold. The system is required to exceed the threshold in order for the goal to be achieved. Optimal systems will more than exceed the threshold
C Type C goals are metrics with no threshold. The system is asked to optimise these but there is no minimum requirement.

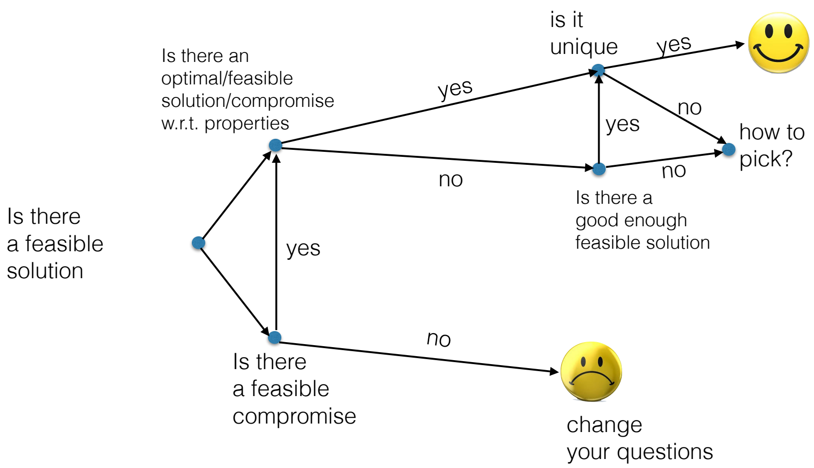

Figure 1 Idealised Coordination Design Flow.

Solutions and Compromises

This terminology let us further refine our definitions of solutions and compromises. We note that:

- A solution achieves all type A and type B goals.
- Given some evaluation function over goals of type B and type C, and optimal solution achieves the maximal value for that function.
- A compromise fails to satisfies all type A and type B goals.
- Givens some evaluation function over goals of type B, type C and over the number-/desirability of goals of type A and type B, an optimal compromise maximises the value of that function.

Coordination Measures

We will refer to goals of type B and type C as *coordination measures*. Typical coordination measures include: Fairness, How much free riding is tolerated, Envy-freeness, Efficiency, Social welfare, Robustness, Scalability, Security, Safety, Computation time.

Coordination Design Flow

Ideally we should have a clear formal framework for assessing coordination problems and answering questions about whether solutions exist and are optimal. This is unlikely to be the case in general but it seems reasonable to suppose that classes of problems could be identified which were amenable to such formal analysis. This would give a design flow something like that shown in Figure 1.

4.1.4 Patterns and Mechanisms

We identify a number of features of existing coordinated multi-agent systems. We separate these into coordination patterns and coordination mechanisms.

Coordination Patterns

Orchestration There is some global entity in the system that computes solutions and directs actions.

Synchronisation/Choreography Groups of entities with the system come together to coordinate their actions.

Cooperation Entities in the system may have different endogenous goals but these facilitate each other – e. g., if one entity wishes a table to be in room A and another wishes it removed from room B then they may cooperate to move the table from room B to room A

Collaboration Entities in the system share the same goals.

Competition Entities in the system compete for resources in order to maximise resource usage/utility/or something.

Coordination Mechanisms

There are a wide range of techniques that have been used to implement coordination in multi-agent systems. We identified the following: Incentives, Negotiation, Argumentation, Norms, Social Choice, Auctions, Operations Research Methods, Roles.

4.1.5 Features

In some cases a designer may have no initial system to work with and will be able to design both the entities within the system as well as the coordination mechanisms. In most cases, however, a coordination problem will include some pre-existing entities and environments which the problem needs to include. These may have certain features.

Features of Entities

When considering the individual entities or agents within a system we identified the following features each of which exists on a spectrum: Compatability of Goals, Reliability, Altruism, Malice.

Features of the Final System

We believe the final system representing a solution to a coordination problem can be placed in a two-dimensional space governed by how much centralisation there is, and how strict the control of the individual entities is. We illustrate that in figure 2.

We have mapped out some common coordination designs in this space. Regimentation is a coordination in which agents are strictly controlled by some centralised authority. Normative systems exist somewhere in the middle of the Strict-Lax axis since agents may ignore norms but will incur sanctions if they do so. These systems may be centralised with some organisation or institution making judgements about violations and imposing sanctions or de-centralised in which individual agents make judgements about the behaviour of other agents and then apply sanctions individually.

While it is difficult to imagine systems in which there is strict de-centralised control, there are systems in which the centralisation may be both static and not strictly-speaking an entity in the system. This would include situations where the built environment passively controls the agent behaviour – e. g., by physically restricting access to certain locations.

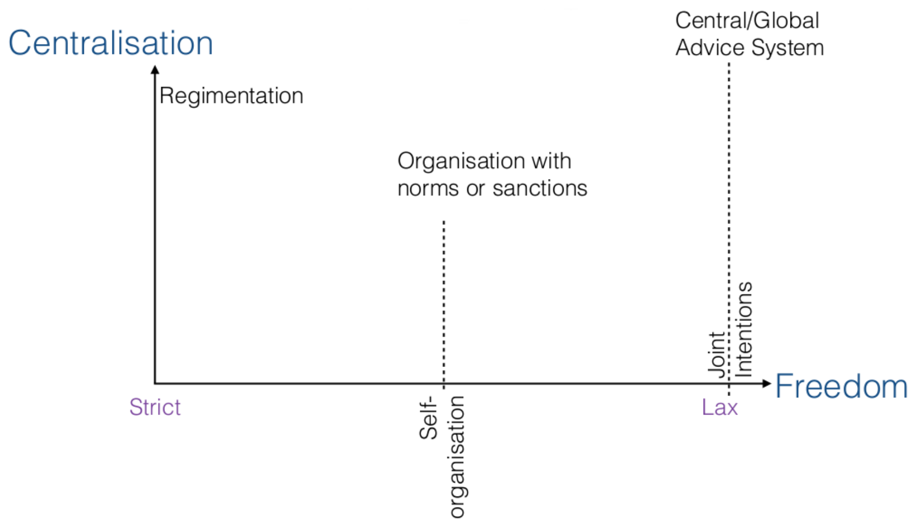

Figure 2 Features of a Coordination.

4.1.6 Examples of Coordination Problems

We conidered two high level examples of significant coordination problems and discuss how they relate to the engineering concepts outlined above.

Traffic Control

Driverless cars have a high profile and strong commercial support eager to bring them to market. The presence of autonomous vehicles, as well as new opportunities to network drivers and traffic systems allows us to consider more sophisticated systems for controlling traffic in place of the current systems which rely on norms, regulation, and minimally networked signalling.

Smart Homes and Smart Grids

Our other example considered the interaction of two coordinations. A *smart grid* manages the flow of electricity around a large network. A *smart home* forms a node in a grid and consists of a number of devices which aim to make the life of the home owner easier and safer.

Comparing and Contrasting the Two Systems

We note

- Humans in the traffic control example accept a far greater loss of autonomy than they do in the smart home. It makes more sense to talk about using norms to regulate human behaviour in one than in the other.
- Although legacy coordination systems exist in both cases, it will be harder to get rid of them entirely in the traffic control system where people have strong beliefs that they are required for safety.
- In the traffic control example there are a large number of different activities that require coordinations to achieve.

- In the traffic control example all entities have similar capabilities (they are all vehicles) while in the smart home the entities all have very different capabilities.
- In the traffic control example we must account for a much higher degree of self-interest (or even malice) from the agents than in the home where all computational entities exist to serve the home owner.

4.1.7 Open Research Questions

Given the framework we have outlined above, we can identify some open research questions.

- What are the techiques for answering the following question?
  Given X pre-existing components for a system, and Y requirements (or more realistically motivations – a formal expression of the requirements) is there a feasible coordination that achieves the requirements and, if so, is it optimal? What happens if we restrict ourselves only to certain patterns or mechanisms for building the solution.
- What are coordination patterns and how do they differ from mechanisms?
- How does monitoring fit in with the scheme?
- We note that coordinations do not exist in isolation. There may be external entities that influence the coordination and which the coordination influences in turn. In some cases these external entities may be additional coordinations. Coordinating a set of coordinations is also a design problem.
  Examples of these are, for instance, the interactions between parliaments and peoples. The people choose (design) the parliament which in turn regulates the behaviour of the people.

4.2 Working group: Formalization of Coordination

Joint work of working group members; written and edited by: Jan Broersen, Marc van Zee, Joshua Sack, Paolo Turrini

4.2.1 Definitions of coordination

The working group came up with several views on the nature of coordination, resulting in different definitions. There was a separation between two camps, roughly described by the following definitions of coordination.

> **Definition 1**: Coordination is the merging of plans of individual agents into a single group plan, independent of any goals on the individual or group level.
> **Definition 2**: Coordination is the aggregation / merging of goals of individual agents into a group goal and the ensuing process of coming to, and implementing, a joint plan to reach the aggregated group goal.

To the surprise of the coordinator of this group (and co-author of this report), support for the first definition was strong. This led some to claim that under that definition just any joint behaviour would count as a coordination. That there is this dichotomy is an interesting phenomenon. We see this in many areas. See the discussion on 'capability theorists' versus 'utility theorists' in economics. Or the dichotomy between 'mentalists' and 'physicalists' in theories of action (e. g. in stit theory).

The working group arrived at the conclusion that the notion of 'protocol' is crucial, because the result of a coordination is best thought of as a protocol. We came, roughly, to the following definition.

Definition 3: A protocol is a restriction on the maximal game tree containing all combinations of free choices of the group of coordinating agents.

Note that the definition of 'protocol' does not go into the issue of the 'character' of the restriction. The restriction may be either such that particular actions are made impossible (regimentation) or they may lead to changes in the utility or goal structures. Again, the difference between these options is a subject of heated debate.

An issue that came up several times is that it is important to distinguish between coordination 'upfront' and coordination 'on the fly'. The definitions are roughly as follows.

Definition 4: Coordination 'upfront' separates the deliberation phase and the execution phase of a coordination: first there is the group deliberation phase and then there is the execution phase where the interaction game is played with the protocol in place.

Definition 5: Coordination 'on the fly' does not separate the deliberation phase and the execution phase of a coordination: during execution, agents make new observations, get new gaols, stumble into impossibilities, etc, etc.

When looking at concrete examples of coordination, it is clear that 'coordination on the fly' is the more realistic conceptualisation. But, as always, to come to a reasonable formalisation, it makes sense to simplify reality and focus on idealised scenarios where the execution phase and deliberation phase are treated separately.

4.2.2 A concrete example

In order to understand and illustrate the many ways in which a multi-agent system is able to coordinate we will consider a football team (for the American reader: a soccer team) that is engaged in playing a football match. We will explain the coordination process within the football team as a layered process consisting of the following three levels:

3. Normative powers (obligations, permissions, counts as)
2. Collective mental attitudes (collective beliefs, intentions, goals)
1. Individual mental attitudes (individual beliefs, intentions, goals),

where the individual agents each contain their own doxastic and motivational attitudes, i. e. their beliefs, goals, intentions, preferences, etc. These attitudes may depend on each other in different ways, and changing one of them may cause a change in others. For instance, if an agent believes that it is impossible to reach his goal to receive a ball that has been played by another team member, the agent may decide to pursue another goal such as defending.

We can view coordination in a multiagent system as a coordination process of these different mental attitudes, giving rise to higher-level group attitudes such as collective beliefs, collective intentions, and collective goals. There is still much debate whether these higher-level attitudes can in some way be reduced to the individual attitudes[1].

Going up another layer in the coordination hierarchy, we can say that agents do not just form collective intentions and beliefs, they also decide to follow the rules of the game. For instance, they agree as a group that the man in blak counts as a referee, and that playing the ball to a player of the own team, when there is no more than one player of the other team between him and the enemy goal, counts as offside, which is penalized with a free kick. Thus, speaking in John Searle's terms, a group of agent is able to designate "status functions" to objects, giving them a certain "deontic power". Note that these deontic powers are only

[1] There is a good Stanford encyclopedia article about this called "Collective Intentionality".

there because the group has chosen to accept them. Another example of a concept that we as humans have given a special status is money. It is interesting to note that most of the money does not even exist. It is merely a concept that exist because we want it, but there is no physical representation of it.

How does this affect choice of action of an agent? We can say that there are two kinds of reasons for actions, namely desire-dependent reasons for actions and desire independent reasons for action. The desire dependent reasons for action come from within the agent. For instance, a football player aims to score a goal because of his desire to be celebrated by his team. Desire independent reasons come from the deontic powers, and explain why we sometimes do things we don't want. For instance, a defender might not even like to play left behind and would rather play as a striker, but since he counts as a good defender, he has to play there anyway.

4.2.3 Towards formalization

Most contributors within the working group thought of 'Alternating Time Temporal Logic [1] / Coalition Logic [10] / Strategic *stit* Logic [3]' as the typical formal setting in which to study coordination. However, a drawback of the ontological commitments made by these logic is maybe that they do not easily connect to logics of programs or logics of action that depart from an event-based perspective (Dynamic Logic [6], Situation Calculus [8], etc.). To accommodate this criticism, we may look at the following alternatives:

- dynamic logic for strategies and concurrent processes [2, 11, 12]
- strategy logic (making strategies explicit in the object language) [7]
- interpreted systems (the aim of interpreted systems is to ground abstract logical models as runs of concrete systems) [4, 5]
- protocols as sets of strategies [9]

We can extract several important ingredients necessary to model coordination from the following example. Consider two agents, Ann and Bob, who are choosing a dress for a party. They can only choose between a black suit and a white suit. They both prefer to wear a suit of the same color, but they have no way to know what the other has chosen. This situation is common knowledge among the two.

Necessary ingredients

- There must be agents that can perform different actions independently and concurrently. Independently means that their choices of action cannot interfere in the choice of actions of other agents. Concurrently means that agents don't know, before making their choice, what choices the others have made. Note that concurrently is not the same as simultaneously, which means that actions must happen at the same time.
- There must be desired outcomes (i. e., A& B choose the same suit) and undesired ones. In particular, each agent must have an action that is not a sure success, but only a success conditional to an action of another agent. All actions by Ann and Bob are not a sure success (success being choosing the same suit).

Unneccesary ingredients. These ingredients are not necessary to model coordination and we can abstract away from them.

- It is not necessary to have a strict linear order on the outcomes, nor a numerical utility function, but only a dichotomous (success failure) relation.

- It is not necessary to have paths, time lines etc. but only a dichotomous (now, later) one. Clearly time adds notions such as interleaving, but every extensive structure can be translated into a normal form one (plus information sets), therefore we can do with the latter.
- It is not necessary to have more than two agents. Each problem involving more agents can be simulated.
- It is not necessary to have more than two actions. Also because each problem involving more agents can be simulated.
- It is not necessary to have probabilistic beliefs. They don't add anything to knowledge relations in modelling the problem.

The big claim is that each coordination problem can be ultimately described with the ingredients used above and no more.

4.2.4 Conclusion

The above insights are the result of 4 days of discussion and later consideration and rationalisations by the four authors. We think they form a sound basis for further study.

The discussions during the workshop tended to focus on rather basic differences of a philosophical nature, which hampered progress. This was somewhat unexpected. The joint goal of the workshop contributors is to come to formalisation and conceptualisations of coordination in Multi–Agent Systems. To reach this goal it would be good if for the next edition of the workshop we find a format (indeed: a coordination) that prevents discussions on preliminaries that are too long.

References

1 R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, 2002.
2 J. F. A. K van Benthem. In praise of strategies. Unpublished manuscript, 2007.
3 Jan Broersen. CTL.STIT: enhancing ATL to express important multi-agent system verification properties. In *Proceedings of the ninth international joint conference on Autonomous agents and multiagent systems (AAMAS '10)*, New York, NY, USA, 2010. ACM.
4 R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge.* The MIT Press, 1995.
5 J. Y. Halpern and R. Fagin. Modelling knowledge and action in distributed systems. *Distributed Computing*, 3(3):159–177, 1989.
6 D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic.* The MIT Press, 2000.
7 Thomas A. Henzinger Krishnendu Chatterjee and Nir Piterman. Strategy logic. In *Proceedings of the 18th International Conference on Concurrency Theory (CONCUR)*, Lecture Notes in Computer Science 4703, pages 59–73. Springer, 2007.
8 J. McCarthy and P. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie, editors, *Machine Intelligence*, volume 4, pages 463–502. Edinburgh University Press, 1969.
9 Rohit Parikh and Ramaswamy Ramanujam. A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12(4):453–467, September 2003.
10 Marc Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.
11 D. Peleg. Communication in concurrent dynamic logic. *Journal of Computer and System Sciences*, 35:23–58, 1987.
12 Jan van Eijck. PDL as a multi-agent strategy logic. In *Proceedings TARK 2013*, 2013.

5 Open Problems

The seminar aimed at bringing together researchers from (at least) three disciplines – logic, game theory and agreement technologies – to discuss and solve (formal) problems related to coordination. As expected not all problems could have been solved and new open problems have emerged. Some are listed below.

- What are the techniques for answering the following question?
  Given X pre-existing components for a system, and Y requirements (or more realistically motivations – a formal expression of the requirements) is there a feasible coordination that achieves the requirements and, if so, is it optimal? What happens if we restrict ourselves only to certain patterns or mechanisms for building the solution.
- What are coordination patterns and how do they differ from mechanisms?
- How does monitoring fit in with the scheme?
- We note that coordinations do not exist in isolation. There may be external entities that influence the coordination and which the coordination influences in turn. In some cases these external entities may be additional coordinations. Coordinating a set of coordinations is also a design problem.
- What are interesting coordination properties that it would be useful to express in a formal logic?
- How suitable are different existing logical formalisms (e. g., coalition logic or PDL) for expressing interesting coordination properties?
- How should a general and flexible model of coordination look like, if there is one at all?
- Should coordination always be conscious or can it just emerge incidentally?

Participants

- Thomas Ågotnes
University of Bergen, NO
- Markus Brill
Duke University, US
- Jan M. Broersen
Utrecht University, NL
- Nils Bulling
TU Clausthal, DE
- Natalia Criado
John Moores University – Liverpool, GB
- Mehdi Dastani
Utrecht University, NL
- Louise A. Dennis
University of Liverpool, GB
- Catalin Dima
University Paris-Est – Créteil, FR
- Valentin Goranko
Technical Univ. of Denmark – Lyngby, DK
- Maaike Harbers
TU Delft, NL
- Andreas Herzig
Paul Sabatier University – Toulouse, FR
- Max Knobbout
Utrecht University, NL
- Ho-Pun Lam
NICTA – Brisbane, AU
- Beishui Liao
Zhejiang University, CN
- Brian Logan
University of Nottingham, GB
- Marin Lujak
University Rey Juan Carlos, ES
- Nir Oren
University of Aberdeen, GB
- Julian Padget
University of Bath, GB
- Truls Pedersen
University of Bergen, NO
- Joshua Sack
University of Amsterdam, NL
- Marija Slavkovik
University of Bergen, NO
- Bas J. G. Testerink
Utrecht University, NL
- Nicolas Troquard
National Research Council – Trento, IT
- Paolo Turrini
Imperial College London, GB
- Birna Van Riemsdijk
TU Delft, NL
- Marc Van Zee
University of Luxembourg, LU
- Laurent Vercouter
INSA – St-Etienne-du-Rouvray, FR