Modelling Rational Agents in Multi-Agent Systems (Extended Abstract)

Nils Bulling
Department of Informatics, Clausthal University of Technology, Germany
bulling@in.tu-clausthal.de

ABSTRACT

This extended abstract gives an overview about my current and future research as well as a summary of my PhD thesis. The thesis is about rational agents in multi-agent systems where the main focus is on formal methods that allow for modelling and reasoning about such systems and its comprised agents. Several aspects of rational agency are treated, for instance, rational agents' behaviors, coalition formation processes, communication among rational agents, and acting with limited resources. The main questions which are tried to be answered are of the following nature: How do rational agents behave under various restrictions and settings? Complexity issues are considered as well, mainly with respect to model checking.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent Systems*; I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Modal logic, Temporal logic*

General Terms

Theory, Logical foundations

Keywords

multi-agent systems, argumentation, coalition formation, game theory, temporal logic

1. MOTIVATION AND OVERVIEW

A vast amount of software systems can be considered as multi-agent systems (MAS) [10, 11], even if the notion of an agent is not explicitly used. Such systems (e.g. online shops, distributed systems, web services, and computer games) often require human or rational decision making to provide a good and up-to-date service. Computer games are excellent examples of software programs with an increasing demand for such rational decision techniques; the same is valid for a variety of commercial applications. Arguably, basic needs of most of those programs include *knowledge representation techniques* and interfaces that allow to *query* or *infer data* from the information stored. The logic-based approaches for rational agents we present here provide means for both points mentioned above. We discuss rationality in MAS from a more theoretical point of view: The focus is on *modelling*, *specifying*, and *verifying* the behavior of rational agents, issues important to guarantee that software is reliable and to ensure that it does what it is supposed to do.

Apart from these cases we consider formal logics as tools to *speak about* and to better understand the complex interactions taking place in MAS, again, focussing on rationality. The main questions we try to answer are of the following form: *How to model and to reason about rational agents?*; *How do rational agents cooperate and communicate?*; or *How do rational agents act under incomplete information and limited resources?*

In addition to the main part, in which we take on a *model theoretic point of view*, we also discuss how rationality aspects can be analyzed and implemented in the more practical setting of *agent programming languages*. Here we especially focus on *communicative acts* and how to interpret them. We also sketch how these tools might be used for reasoning *within* agents; clearly, agent programming languages prepare an appropriate ground for that.

Finally, throughout the thesis we are interested in the computational aspect of the presented formal frameworks, in particular in the analysis of the model checking complexity.

2. RATIONALITY ASPECTS IN MAS

In the past it has been shown that modal logics are applicable to a great many of heterogeneous systems. *Epistemic logics*, for instance, are used to model and to reason about knowledge of agents; *temporal logics* allow to verify temporal properties of systems. *Strategic logics* have attracted quite some interest in recent research. They describe what agents can enforce and what power coalitions have. Among these logics *Alternating-Time Temporal Logic* (ATL for short) [1] is one of the most influential; it combines temporal concepts with basic game theoretic ones. ATL is very flexible regarding extensions by other modal concepts, e.g. by epistemic logic, which often result in powerful and interesting logics applicable to various areas of MAS.

In this thesis we analyze how ATL can be extended in such a way that it is suitable for the modelling of various *rationality aspects* in MAS as pure ATL does not allow to speak about sensible strategies *per se* rather about *all* possible behaviors of agents regardless whether they make sense or not.

One of the main questions addressed is the following: *How do agents behave if they act according to a given* plausibility *or* rationality *assumption?* Apart from the epistemological gain about agents' rational behavior the answer provides, there also is a more practical aspect: In many games, from a game theoretic point of view, the number of all possible outcomes is infinite, although only some of them "make sense"; hence, a notion of rationality (like subgame-perfect Nash equilibrium) allows to discard the "less sensible" ones, and to determine what should happen had the game been played by ideal players. For this purpose we extend ATL with a notion of *plausibility* [5] and refer to the logic as ATLP. This extension of ATL enables us (1) to *express* various rationality assumptions of intelligent agents; (2) to *specify* sets of rational strategy profiles *in the object language*; and (3) to reason about agents' play if only those strategy profiles were allowed. For example, we may assume the agents to play only Nash equilibria, Pareto-optimal profiles or undominated strategies, and ask about the resulting behaviour (and outcomes) under such an assumption. The logic also gives rise to generalized versions of classical solution concepts through characterizing patterns of payoffs by suitably parameterized formulae of ATLP. We investigate the complexity of model checking for several classes of formulae: It ranges from $\mathbf{\Delta_3^P}$ to **PSPACE** in the general case and from $\mathbf{\Delta_3^P}$ to $\mathbf{\Delta_4^P}$ for the most interesting subclasses, and roughly corresponds to solving extensive games with imperfect information.

We do also propose a version of ATLP for *imperfect information games* as "pure" ATLP is for *perfect information games* only. The resulting logic *Constructive Strategic Logic with plausibility* (CSLP) [7, 9] can be used in the same way as ATLP but now for perfect *and* imperfect information games. Moreover, the logic is more than just an independent combination of ATLP with epistemic operators, the plausibility concept allows to defined a neat doxastic notion, *rational beliefs*, on top of knowledge (similar to [4]). We show that beliefs satisfy axioms **KD45**. In summary, CSLP can be used to reason about rational play and rational beliefs under *uncertainty*.

The previous extension is about classical indistinguishability between states, however, there is another interesting angle to incomplete information. Where in ATL *the worst* possible response from the other agents is assumed we consider the case in which agents communication and cooperation abilities are limited such that it is not very likely that the "worst case" will happen. The presented logic ATL *with probabilistic success* [6] tries to soften the rigorous notion of success that underpins ATL and allows to reason about the likelihood that agents have a successful strategy to enforce their goals.

Undeniably, cooperation among agents plays a decisive role in strategic logics, however, in ATL it is only present implicitly. What we would like to analyze is *why* agents should cooperate with other agents. For this purpose we combine an argumentation-based approach to coalition formation [2] into the semantics of ATL. The proposed logic *Coalitional* ATL [3] allows to reason and to model the formation process of *rational* coalitions and their power.

Finally, we identify two further ingredients important for the modelling of rational agents. Firstly, we argue that *resources* play a decisive role in the selection of the right strategy as agents are usually confronted with a limited amount of them what should be reflected in the choice of strategies and in the selection of agents to cooperate with. For this purpose a combination of ATL with a variant of *Linear Logic* [8] is proposed where resources are treated as first-class citizens. Finally, we consider the *communication process* among agents in the more practical setting of agent oriented programming languages. We would like to note that both of the latter topics are part of our ongoing research.

3. CONCLUSIONS

We have presented several logics to model and to analyze rationality aspects in MAS; each of them suitable to be used for a specific aspect of rationality. Our main focus was on a model theoretic analysis where logics can be used to reason about a previously built *model*. This allows for the verification and specification of MAS. In consequence the model checking complexity was important throughout this thesis. In the case of ATL with limited ressources however we exemplarily showed[1] how these logics can also be used from a deductive point of view, e.g. as inference systems *within* agents.

4. REFERENCES

[1] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Lecture Notes in Computer Science*, 1536:23–60, 1998.

[2] L. Amgoud. An argumentation-based model for reasoning about coalition structures. In *ArgMAS*, pages 217–228, 2005.

[3] N. Bulling, C. Chesñevar, and J. Dix. Modelling coalitions: ATL+argumentation. In *Proceedings of AAMAS'08*, Estoril, Portugal, May 2008. ACM Press.

[4] N. Bulling and W. Jamroga. Agents, beliefs, and plausible behavior in a temporal setting. In *Proceedings of AAMAS'07*, pages 570–577, Honolulu, Hawaii, USA, 2007. ACM Press.

[5] N. Bulling, W. Jamroga, and J. Dix. Reasoning about temporal properties of rational play. *Annals of Mathematics and Artificial Intelligence*, 2009.

[6] N. Bulling and W.Jamroga. What agents can probably enforce. In *Proceedings of the International Workshop on Concurrency, Specification and Programming CS&P'08*, Gross Vaeter, Germany, 2008.

[7] N. Bulling and W.Jamroga. Rational play and rational beliefs under uncertainty. In *Proceedings of AAMAS'09*, Budapest, Hungary, may 2009.

[8] J.-Y. Girard. Linear logic: Its syntax and semantics. In J.-Y. Girard, Y. Lafont, and L. Regnier, editors, *Advances in Linear Logic (Proc. of the Workshop on Linear Logic, Cornell University, June 1993)*, number 222. Cambridge University Press, 1995.

[9] W. Jamroga and T. Ågotnes. Constructive knowledge: What agents can achieve under incomplete information. *Journal of Applied Non-Classical Logics*, 17(4):423–475, 2007.

[10] G. Weiss, editor. *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*. MIT Press: Cambridge, Mass, 1999.

[11] M. Wooldridge. *An Introduction to Multi Agent Systems.* John Wiley & Sons, 2002.

[1]More precisely, we will show this as part of our current research.