Modal logics for games and multi-agent systems

Thomas Ågotnes
Wojtek Jamroga

ESSLLI 2008
20th European Summer School in Logic, Language and Information
4–15 August 2008
Freie und Hansestadt Hamburg, Germany

Programme Committee. Enrico Franconi (Bolzano, Italy), Petra Hendriks (Groningen, The Netherlands), Michael Kaminski (Haifa, Israel), Benedikt Löwe (Amsterdam, The Netherlands & Hamburg, Germany) Massimo Poesio (Colchester, United Kingdom), Philippe Schlenker (Los Angeles CA, United States of America), Khalil Sima'an (Amsterdam, The Netherlands), Rineke Verbrugge (**Chair**, Groningen, The Netherlands).

Organizing Committee. Stefan Bold (Bonn, Germany), Hannah König (Hamburg, Germany), Benedikt Löwe (**chair**, Amsterdam, The Netherlands & Hamburg, Germany), Sanchit Saraf (Kanpur, India), Sara Uckelman (Amsterdam, The Netherlands), Hans van Ditmarsch (**chair**, Otago, New Zealand & Toulouse, France), Peter van Ormondt (Amsterdam, The Netherlands).

http://www.illc.uva.nl/ESSLLI2008/
esslli2008@science.uva.nl

INSTITUTE FOR LOGIC,
LANGUAGE AND COMPUTATION

ESSLLI 2008 is organized by the Universität Hamburg under the auspices of the *Association for Logic, Language and Information* (FoLLI). The *Institute for Logic, Language and Computation* (ILLC) of the *Universiteit van Amsterdam* is providing important infrastructural support. Within the Universität Hamburg, ESSLLI 2008 is sponsored by the Departments *Informatik*, *Mathematik*, *Philosophie*, and *Sprache, Literatur, Medien I*, the *Fakultät für Mathematik, Informatik und Naturwissenschaften*, the *Zentrum für Sprachwissenschaft*, and the *Regionales Rechenzentrum*. ESSLLI 2008 is an event of the *Jahr der Mathematik 2008*. Further sponsors include the *Deutsche Forschungsgemeinschaft* (DFG), the Marie Curie Research Training Site GLoRiClass, the European Chapter of the Association for Computational Linguistics, the *Hamburgische Wissenschaftliche Stiftung*, the Kurt Gödel Society, Sun Microsystems, the Association for Symbolic Logic (ASL), and the European Association for Theoretical Computer Science (EATCS). The official airline of ESSLLI 2008 is Lufthansa; the book prize of the student session is sponsored by *Springer Verlag*.

Thomas Ågotnes and Wojtek Jamroga

Modal logics for games and multi-agent systems

Course Material. 20th European Summer School in Logic, Language and Information (ESSLLI 2008), Freie und Hansestadt Hamburg, Germany, 4–15 August 2008

Modal Logics for Games and Multi-Agent Systems
Course outline

Thomas Ågotnes[1] and Wojtek Jamroga[2]
[1] Department of Computer Engineering, Bergen University College, Norway
[2] Department of Informatics, Clausthal University of Technology, Germany
tag@hib.no, wjamroga@in.tu-clausthal.de

Overview. Because agents in multi-agent systems (MAS) are often assumed to act rationally based on self interest, it seems appropriate to study such systems from an economic perspective. In particular, agents' strategies (and their outcomes) must be taken into account. On one hand, these issues have been studied extensively in game theory. On the other hand, formal logic is widely regarded as a foundation for specification, verification and reasoning about MAS. In this course, we study logical formalisations of games and agents in game-like scenarios.

The course alternates between two main tracks. The first is focused on Alternating-time Temporal Logic (ATL), one of the most prominent logics of cooperation and strategic ability. We introduce ATL, show its relationship to extensive games of perfect information, and discuss some important extensions of the logic, e.g. with uncertainty and/or rationality assumptions. In particular, we focus on imperfect information and the integration of strategic and epistemic modalities. The second track regards logical characterisations of game theoretic solution concepts, using different modal logics.

Background. Some basic familiarity with modal logic and game theory is assumed. Knowledge of epistemic and temporal logic, and game-theoretic solution concepts is a pro, but it is not required.

Program. The course consists of 10 lectures, 2 per day, each 40 minutes long (with a 10 minutes break between them). A preliminary program is given below.

1. **Introduction.** Multi-agent systems. Modal logic. Epistemic logic. Axioms and systems of modal logic. Correspondence theory.

2. **Coalition logic.** Strategic games and coalition logic (CL). Axiomatisation of CL.

3. **ATL.** Multi-step games and alternating-time temporal logic (ATL).

4. **More about ATL.** Axiomatisation; bisimulation; the role of memory; revocability of strategies.

5. **Strategic reasoning for imperfect information (part I).** Strategic reasoning for imperfect information scenarios. Problems with ATEL. Economic solution: ATLir.

6. **Strategic reasoning for imperfect information (part II).** General solution: CSL. Properties of constructive knowledge. Semantics for constructive normal form.
7. **Characterising solution concepts.** Non-cooperative games: characterising solution concepts in modal logic.
8. **Reasoning about rational play.** Reasoning about rational play in ATLP. Temporalized solution concepts.
9. **Axiomatisation of coalitional games (part I).** Cooperative games. Examples of games. Solution concepts. Axiomatisation in modal logic.
10. **Axiomatisation of coalitional games (part II).** Axiomatisation of coalitional games: completeness proof.

See the following individual lecture notes for references to reading material. Some material will be distributed as hand-outs at the course. Publications which may be of particular interest are:

1. W. van der Hoek and M. Pauly. Modal logic for games and information. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, pages 1077–1148. Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 2006.
2. R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002. Available at `http://www.eecs.berkeley.edu/~tah/Publications/alternating-time_temporal_logic.html`.
3. 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. Available at `http://www2.in.tu-clausthal.de/~wjamroga/research.php#publications`.

More course material, including updated notes, exercises, slides, and references will be placed on the course website:

`http://www2.in.tu-clausthal.de/~wjamroga/courses/GamesMAS2008ESSLLI/`

Lecture 1: Multi-agent systems and modal logics

The first lecture serves as an introduction of some fundamental concepts that we are going to use throughout this course.

Multi-agent systems. *Multi-agent systems* (*MAS*) are systems that involve several autonomous entities acting in the same environment. The entities are called *agents*. What is an agent? Despite numerous attempts to answer this question, there seems to be no conclusive definition. We argue that MAS are, most of all, a philosophical metaphor that induces a specific way of seeing the world, and makes us use agent-oriented vocabulary when describing the phenomena we are interested in. Thus, while some researchers present multi-agent systems as a new paradigm for computation or design, we believe that primarily multi-agent systems form a new paradigm for *thinking* and *talking* about the world, and assigning it a specific conceptual structure. The metaphor of a multi-agent system seems to build on the intuition that *we* are agents – and that other entities we study can be just like us to some extent. The usual properties of agents, like autonomy, pro-activeness etc., seem to be secondary: they are results of an introspection rather than primary assumptions we start with.

We note that this view of multi-agent systems comes close to the role of *logic* in both philosophy and computer science. Logic provides conceptual structures for *modeling* and *reasoning* about the world in a precise manner – and, occasionally, it also provides tools to do it automatically.

References: [Weiss, 1999; Wooldridge, 2002].

Modal logic. *Modal logic* is an extension of classical logic with new operators $\Box$ (*necessity*) and $\Diamond$ (*possibility*): $\Box p$ means that p is true in every possible scenario, while $\Diamond p$ means that p is true in at least one possible scenario. Let Π be the set of atomic propositions. Models of modal logics are called *Kripke structures* or *possible world models*, and include the set of *possible worlds* (or states) St, modal accessibility relation(s) $\mathcal{R}_1, ..., \mathcal{R}_k \subseteq St \times St$, and interpretation of propositions $\pi : \Pi \to 2^{St}$. Now, for a model $\mathcal{M}$ and world q in $\mathcal{M}$:

$$\mathcal{M}, q \models \Box_i \varphi \quad \text{iff} \quad \mathcal{M}, q' \models \varphi \text{ for all } q' \text{ such that } q\mathcal{R}_i q'$$
$$\mathcal{M}, q \models \Diamond_i \varphi \quad \text{iff} \quad \mathcal{M}, q' \models \varphi \text{ for some } q' \text{ such that } q\mathcal{R}_i q'$$

The actual accessibility relation can capture various dimensions of the reality (and therefore gives rise to different kinds of modal logics): knowledge ($\rightsquigarrow$ *epistemic logic*), beliefs ($\rightsquigarrow$ *doxastic logic*), obligations ($\rightsquigarrow$ *deontic logic*), actions ($\rightsquigarrow$ *dynamic logic*), time ($\rightsquigarrow$ *temporal logic*) etc. In particular, various aspects of agents (and agent systems) can be naturally captured within this generic framework.

References: [Blackburn *et al.*, 2001].

Reasoning about knowledge. The basic *epistemic logic* which we consider here involves modalities for individual agent's knowledge K_i, with $\mathrm{K}_i\varphi$ interpreted as "agent i knows that φ". Additionally, one can consider modalities for collective knowledge of groups of agents:

mutual knowledge ($E_A\varphi$: "everybody in group A knows that φ"), *common knowledge* ($C_A\varphi$: "all the agents in A know that φ, and they know that they know it etc."), and *distributed knowledge* ($D_A\varphi$: "if the agents could share their individual information, they would be able to recognize that φ"). Note that $E_A\varphi \equiv \bigwedge_{i\in A} K_i\varphi$, and hence the operators for mutual knowledge can be omitted from the language.

The formal semantics for the logic is based on *multi-agent epistemic models* of the type $\langle St, \sim_1, ..., \sim_k\rangle, \pi$, where St is a set of *epistemic states*, π is a valuation of propositions, and each $\sim_i \subseteq St \times St$ is an equivalence relation that defines the *indistinguishability of states* for agent i. Operators K_i are provided with the usual Kripke semantics given by the clause:

$$\mathcal{M}, q \models K_i\varphi \text{ iff } \mathcal{M}, q' \models \varphi \text{ for all } q' \text{ such that } q \sim_i q'$$

The accessibility relation corresponding to E_A is defined as $\sim_A^E = \bigcup_{i\in A} \sim_i$, and the semantics of E_A becomes

$$\mathcal{M}, q \models E_A\varphi \text{ iff } \mathcal{M}, q' \models \varphi \text{ for all } q' \text{ such that } q \sim_A^E q'.$$

Common knowledge C_A is given semantics in terms of the relation $\sim_A^C$ defined as the transitive closure of $\sim_A^E$:

$$\mathcal{M}, q \models C_A\varphi \text{ iff } \mathcal{M}, q' \models \varphi \text{ for all } q' \text{ such that } q \sim_A^C q'.$$

Finally, distributed knowledge $\mathsf{D_A}$ is given semantics in terms of the relation $\sim_A^D$ defined as $\bigcap_{i\in A} \sim_{\mathsf{i}}$, following the same pattern:

$$\mathcal{M}, q \models D_A\varphi \text{ iff } \mathcal{M}, q' \models \varphi \text{ for all } q' \text{ such that } q \sim_A^D q'.$$

In the lecture we mention some axiomatic properties of knowledge, and the correspondence between these axioms and structural properties of models. We also briefly discuss the logical omniscience problem, which is one of the main obstacles when modeling realistic (i.e., resource-bounded) agents with modal logic. Interestingly, this problem is inherent to programming epistemic states via logic programs, too.

References: [Fagin *et al.*, 1995; Halpern, 1995; van der Hoek and Verbrugge, 2002].

Lecture 2: Coalition logic

Strategic form games. Interactions between autonomous and rational agents acting strategically have been extensively studied in the field of *game theory*. The models used in game theory can be categorised into two types: *non-cooperative* games, in which the possible actions of individual players are taken as primitives, and *coalitional* (or *cooperative*) games, in which the possible joint actions of groups of players are taken as primitives. Don't be fooled by the terminology: both types of games can be used to reason both about cooperation and coalitions.

In the lecture we briefly introduce the most important models of the non-cooperative theory. Solution concepts for these games will be discussed in lecture 7, and coalitional games and their solution concepts in lectures 9 and 10.

A standard model in non-cooperative game theory is that of a *strategic game*. In a strategic game, it is assumed that each agent chooses her future actions (her strategy) once and for all at the beginning of the game, and that all agents do this simultaneously. Formally, a strategic game is a tuple $G = (N, \{\Sigma_i : i \in N\}, o, S, \{\succeq_i : i \in N\})$ where N is the set of agents, S is a set of states, Σ_i is the set of actions (or strategies) available to agent i, o associates an outcome state $o(a_1, \ldots, a_n) \in S$ with every tuple of actions $(a_1, \ldots, a_n) \in \times_{i \in N} \Sigma_i$, and $\succeq_i \subseteq S \times S$ is a preference relation (complete, reflexive, transitive) over the outcome states for agent i. Leaving out the preference relations, we get a *strategic game form* $G = (N, \{\Sigma_i : i \in N\}, o, S)$.

Cooperation in games. Henceforth, a *coalition* is simply a group of agents. An *effectivity function* models the powers agents can obtain by forming coalitions. Given a set of states X, we can say that a coalition C is *effective* for X if C can cooperate to ensure that the outcome will lie in X. Formally, an effectivity function for a set of players N over a set of states S is a function

$$E : \wp(N) \to \wp(\wp(S))$$

giving the sets of outcomes $E(C)$ for which each coalition C is effective. A game frame induces an effective function as follows:

$$X \in E_G(C) \Leftrightarrow \exists \sigma_C \forall \sigma_{N \setminus C} o(\sigma_C, \sigma_{N \setminus C}) \in X$$

– the class of effectivity functions induced from game frames in this way are called *playable* (called *α-effectivity functions* in social choice theory).

Coalition Logic. Modal logics of strategic ability form one of the fields where logic and game theory can successfully meet. Marc Pauly's *Coalition Logic* (CL) formalises reasoning about the abilities of coalitions. The language of CL extends propositional logic with a modality $[C]$ for each coalition C. The intended meaning of $[C]\varphi$ is that C can make φ come about.

Formally, the language is interpreted over *coalition models*

$$\mathcal{M} = (S, E, V)$$

where S is a set of states, $V : \Pi \to 2^S$ an interpretation of atomic propositions Π in the states, and

$$E : S \to (\wp(N) \to \wp(\wp(S)))$$

is a function such that for each $s \in S$, $E(s)$ is a playable effectivity function. The interpretation of the main construct of the language in a state s in a coalition model $\mathcal{M}$ is defined as follows:

$$\mathcal{M}, s \models [C]\phi \Leftrightarrow \phi^{\mathcal{M}} \in E(s)(C)$$

where $\phi^{\mathcal{M}} = \{s \in S : \mathcal{M}, s \models \phi\}$.

Observe from the discussion above that the sentences of Coalition Logic can alternatively and equivalently be seen as statements about strategic games (game forms).

In addition to introducing the language and semantics of CL, we will discuss its axiomatisation and mention other meta-logical properties.

References: [Pauly, 2001; Osborne and Rubinstein, 1994] .

Lecture 3: ATL

Alternating-time Temporal Logic (ATL) is one of the most important logic of time and strategies that has emerged in recent years. ATL is a generalization of the branching time temporal logic CTL (Coalition Tree Logic), in which path quantifiers (E: "there is a path", A: "for every path") are replaced with *strategic quantifiers* $\langle\langle A\rangle\rangle$. The formula $\langle\langle A\rangle\rangle\varphi$, where A is a coalition of agents, expresses the claim that A has a collective strategy to enforce φ. Moreover, ATL formulas include the usual temporal operators: $\bigcirc$ ("in the next state"), $\Box$ ("always from now on"), $\Diamond$ ("sometime in the future"), and $\mathcal{U}$ ("until").

Several semantics have been introduced for ATL, of which the one based on *concurrent game structures* (*CGS*) is perhaps the most popular. A concurrent game structure $\mathcal{M} = \langle \mathbb{A}\mathrm{gt}, St, Act, d, o, \pi\rangle$ includes: a nonempty finite set of agents $\mathbb{A}\mathrm{gt} = \{1, ..., k\}$; a nonempty set of global states St of the system; a set of (atomic) actions Act; function $d : \mathbb{A}\mathrm{gt} \times St \to 2^{Act}$ defining the set of actions available to an agent in a state; a valuation of propositions $\pi : St \to 2^{\Pi}$; and o is a (deterministic) transition function which assigns the outcome state $q' = o(q, \alpha_1, \dots, \alpha_k)$ to state q and a tuple of actions $\langle\alpha_1, \dots, \alpha_k\rangle$ that can be executed by $\mathbb{A}\mathrm{gt}$ in q. Note that, as concurrent game structures do not allow to represent any kind of uncertainty, ATL can be seen as a logic for reasoning about agents with *perfect information* about the current state of the game.

A *strategy* of agent a is a conditional plan that specifies what a is going to do for every 'possible situation'. For an agent with no implicit recall of the past, a strategy can be thus represented by function $s_a : St \to Act$ such that $s_a(q) \in d_a(q)$, i.e. a function that specifies a valid choice of action for each *state* of the system. For agents with perfect recall, this would be $s_a : St^+ \to Act$ such that $s_a(q_0q_1...q_i) \in d_a(q_i)$, i.e. one that specifies a valid choice of action for each *history* of the game. A *collective strategy* S_A for a group of agents A is a tuple of strategies, one per agent from A. Then, by $out_M(q, S_A)$ (or just $out(q, S_A)$) we denote the set of all paths that may result from agents A executing strategy S_A from state q onward. The semantics of strategic quantifiers is defined as follows:

$\mathcal{M}, q \models \langle\langle A\rangle\rangle\varphi$ iff there is a strategy S_A, such that $\mathcal{M}, \lambda \models \varphi$ for every $\lambda \in out(q, S_A)$;

$\mathcal{M}, \lambda \models \bigcirc\varphi$ iff $\lambda[1..\infty] \models \varphi$;

$\mathcal{M}, \lambda \models \Box\varphi$ iff $\lambda[i..\infty] \models \varphi$ for all $i \geq 0$;

$\mathcal{M}, \lambda \models \varphi\,\mathcal{U}\,\psi$ iff $\lambda[i..\infty] \models \varphi$ for some $i \geq 0$, and $\lambda[j..\infty] \models \varphi$ for all $0 \leq j \leq i$.

Additionally, the "sometime" modality can be defined as: $\Diamond\varphi \equiv \top\,\mathcal{U}\,\varphi$.

We observe that ATL embeds two important logics. First, the branching time logic CTL can be seen as a strategic logic with a very limited set of strategic quantifiers (as the CTL path quantifiers can be embedded in ATL with the following definitions: $\mathsf{E}\varphi \equiv \langle\langle \mathbb{A}\mathrm{gt}\rangle\rangle\varphi$, $\mathsf{A}\varphi \equiv \langle\langle\varnothing\rangle\rangle\varphi$). Second, Coalition Logic can be seen as the "next"-fragment of ATL, with the following embedding of the central modality of CL: $[A]\varphi \equiv \langle\langle A\rangle\rangle\bigcirc\varphi$.

The above clauses define the semantics of the full version of alternating-time logic, usually called ATL* for historical reasons. It must be noted, however, that the typical variant of ATL, used in the literature, is restricted to formulae in which every temporal operator is immediately preceded by exactly one cooperation modality. We will usually refer to the restricted variant as "vanilla" ATL.

Towards the end of the lecture, we briefly mention some complexity results regarding the satisfiability and model checking problems for ATL, and compare them with analogous results for temporal logics.

References: [Alur *et al.*, 2002].

Lecture 4: More about ATL

In this lecture we continue our discussion of ATL and its logical and meta-logical properties.

We begin by presenting a sound and complete axiomatisation of "vanilla" ATL, by Goranko and van Drimmelen.

The concept of *bisimulation* is central in modal logic, and is for example very useful in order to study the key meta-logical property *expressiveness*. Bisimulation is a relationship between semantic structures, and it is typically the case that bisimilar structures satisfy exactly the same formulae. Then, we immediately know something about the expressiveness of the logical language: if two structures are bisimilar and one of them has some property while the other has not, then that property is not expressible in the logical language. The following is an adaption [Ågotnes *et al.*, 2007] of the standard modal logic notion of a bisimulation to CGSs.

First, let $D(q) = d_1(q) \times \cdots \times d_k(q)$ denote the set of *joint actions* in q, and when $\vec{a}_C \in D(q, C)$ let $next_{\mathcal{M}}(q, \vec{a}_C) = \{o(q, \vec{b}) : \vec{b} \in D(q), a_i = b_i \text{ for all } i \in C\}$ denote the set of possible next states in CGS $\mathcal{M}$ when coalition C choose actions $\vec{a}_C$.

Given two CGSs, $\mathcal{M}_1 = (\mathbb{A}\text{gt}, St_1, Act_1, \Pi, \pi_1, d_1, o_1)$ and $\mathcal{M}_2 = (\mathbb{A}\text{gt}, St_2, Act_2, \Pi, \pi_2, d_2, o_2)$, and a set of agents $C \subseteq \mathbb{A}\text{gt}$, a relation $\beta \subseteq St_1 \times St_2$ is a *C-bisimulation between $\mathcal{M}_1$ and $\mathcal{M}_2$*, denoted $\mathcal{M}_1 \rightleftarrows_\beta^C \mathcal{M}_2$, iff for all states $q_1 \in St_1$ and $q_2 \in St_2$, $q_1 \beta q_2$ implies that

Local harmony $\pi_1(q_1) = \pi_2(q_2)$;

Forth For all joint actions $\vec{a}_C^1 \in D_1(q_1, C)$ for C, there exists a joint action $\vec{a}_C^2 \in D_2(q_2, C)$ for C such that for all states $s_2 \in next_{\mathcal{M}_2}(q_2, \vec{a}_C^2)$, there exists a state $s_1 \in next_{\mathcal{M}_1}(q_1, \vec{a}_C^1)$ such that $s_1 \beta s_2$;

Back Likewise, for 1 and 2 swapped.

A relation β is a *bisimulation between $\mathcal{M}_1$ and $\mathcal{M}_2$*, denoted $\mathcal{M}_1 \rightleftarrows_\beta \mathcal{M}_2$, if β is a C-bisimulation between $\mathcal{M}_1$ and $\mathcal{M}_2$ for every $C \subseteq \mathbb{A}\text{gt}$.

Using this notion of bisimulation, we can characterise the expressiveness of ATL as mentioned above.

Here, it is relevant to take into account a semantic issue mentioned in lecture 3: *the role of memory*. Two different ways to define strategies are, first, as as mappings from states to actions, and second, as mappings from sequences of states (histories) to actions. We call the first type *memoryless* strategies, modelling agents who base their actions on the current state only, and the second *full* strategies, modelling agents who base their strategies on the history of states. The semantics of ATL can be defined using either; by restricing the quantification ("there is a strategy..") in the semantic clauses to the appropriate class of strategies.

We can now show that:

- For "vanilla" ATL, memory does not matter: the interpretation of a formula (its truth value in a state of a structure) is the same whether we use only memoryless strategies or whether we use full strategies.

- The language of "vanilla" ATL cannot discern between bisimilar CGSs (neither with memoryless nor with full strategies)
- ATL* is strictly more expressive than "vanilla" ATL; satisfaction of ATL* formulae is not invariant under bisimulations.
- For ATL* memory matters: the interpretation of a formula depends on the amount of recall available.

A final theme is *revocability* of strategies. Observe the following property of ATL: when a formula of the form

$$\langle\langle C \rangle\rangle \Box \varphi$$

is evaluated, the joint strategy chosen by C (in the evaluation of the quantifier $\langle\langle C \rangle\rangle$) does not play any role in the evaluation of the subformula φ. But what if φ is itself a cooperation formula, e.g. if $\varphi \equiv \langle\langle C \rangle\rangle \bigcirc \psi$? In the lecture we discuss whether an alternative semantics could be useful, and look at logics of strategic ability under *irrevocable* strategies. To implement such strategies, the semantics of ATL can be modified to use *model updates*. The update $M \dagger S_C$ of model M by a (memoryless) strategy S_C for coalition C is defined by removing transitions which are not compatible with the strategy. The interpretation of the cooperation modalities can then be defined as follows:

$$M, q \models \langle\langle C \rangle\rangle \Box \phi \Leftrightarrow \exists S_C \forall \lambda \in out_{M \dagger S_C}(q, S_C)\ \forall j \geq 0 (M \dagger S_C, \lambda[j] \models \phi)$$

and so on. The semantics is indeed different from the standard definition (and, now, memory matters).

References: [Goranko and van Drimmelen, 2003; Ågotnes *et al.*, 2007; van der Hoek *et al.*, 2005; Brihaye *et al.*, 2008; Chatterjee *et al.*, 2007; Pinchinat, 2007].

Lectures 5 & 6: Strategic reasoning for imperfect information

ATL does not take into account the epistemic limitations of the agents; it assumes that every agent has *complete information* about the global state of the system. The *Alternating-time Temporal Epistemic Logic* ATEL was introduced in [van der Hoek and Wooldridge, 2002] as a straightforward combination of the multi-agent epistemic logic and ATL in order to formalize reasoning about the interaction of knowledge and abilities of agents and coalitions. ATEL enables expressing various modes and nuances of interaction between knowledge and strategic abilities, e.g.:

$\langle\langle A\rangle\rangle\varphi \rightarrow \mathrm{E}_A\langle\langle A\rangle\rangle\varphi$; $\mathrm{C}_A\langle\langle A\rangle\rangle\varphi \rightarrow \langle\langle A\rangle\rangle\mathrm{C}_A\varphi$; $\langle\langle i\rangle\rangle\varphi \rightarrow \mathrm{K}_i\neg\langle\langle \mathbb{A}\mathrm{gt}\setminus\{i\}\rangle\rangle\neg\varphi$.

The models of ATEL are *concurrent epistemic game structures* (CEGS):

$$\mathcal{M} = \langle \mathbb{A}\mathrm{gt}, St, Act, d, o, \pi, \sim_1, ..., \sim_k\rangle$$

combining the CGS-based models for ATL and the multi-agent epistemic models. The semantics of ATEL takes the semantic clauses for ATL and those from epistemic logic, and combines them in a straightforward way.

While the logic ATEL extends smoothly both ATL and epistemic logic, and provides a rich formalism for reasoning about knowledge and strategic abilities, it also raises a number of conceptual problems. Most importantly, one would expect that an agent's ability to achieve property φ should imply that the agent has enough control and knowledge to *identify* and *execute* a strategy that enforces φ. A number of approaches have been proposed to overcome this problem. Most of the solutions agree that only *uniform* strategies (i.e., strategies that specify the same choices in indistinguishable states) are really executable. However, in order to identify a successful strategy, the agents must consider not only the courses of action, starting from the current state of the system, but also from states that are indistinguishable from the current one. There are many cases here, especially when group epistemics is concerned: the agents may have common, ordinary, or distributed knowledge about a strategy being successful, or they may be hinted the right strategy by a distinguished member (the "leader"), a subgroup ("headquarters committee") or even another group of agents ("consulting company").

Most existing solutions (ATL_{ir} [Schobbens, 2004], ETSL [van Otterloo and Jonker, 2004]) treat only some of the cases (albeit in an elegant way), while others (ATOL [Jamroga and van der Hoek, 2004], "Feasible ATEL" [Jonker, 2003]) offer a more general treatment of the problem at the expense of an overblown logical language. Recently, a new, non-standard semantics for ability under imperfect information has been proposed in the form of *Constructive Strategic Logic* (*CSL*) [], which we believe to be both intuitive, general and elegant. In CSL, formulae are interpreted over *sets of states* rather than single states. We write $M, Q \models \langle\langle A\rangle\rangle\varphi$ to express the fact that A must have a strategy which is successful for all "opening" states from Q. New epistemic operators $\mathbb{K}_i, \mathbb{E}_A, \mathbb{C}_A, \mathbb{D}_A$ for "practical" or "constructive" knowledge yield the set of states for which a single evidence (i.e., a successful strategy) should be presented (instead of checking if the required property holds in each of the states separately, like standard epistemic operators do). Let $\hat{\mathcal{K}} = \mathbb{C}, \mathbb{E}, \mathbb{D}$ and $\mathcal{K} = \mathrm{C}, \mathrm{E}, \mathrm{D}$, respectively. Now:

$\mathcal{M}, Q \models \langle\langle A\rangle\rangle\varphi$ iff there is S_A such that $\mathcal{M}, \lambda \models \varphi$ for every $\lambda \in \cup_{q\in Q}\ out(q, S_A)$;

$\mathcal{M}, Q \models \hat{\mathcal{K}}_A\varphi$ iff $\mathcal{M}, \{q' \mid \exists_{q \in Q}\ q \sim_A^{\mathcal{K}} q'\} \models \varphi$.

We point out that in CSL:

1. $\mathbb{K}_a\langle\langle a\rangle\rangle\varphi$ refers to agent a having a strategy *"de re"* to enforce φ (i.e. having a successful uniform strategy and knowing the strategy);
2. $K_a\langle\langle a\rangle\rangle\varphi$ refers to agent a having a strategy *"de dicto"* to enforce φ (i.e. knowing only that *some* successful uniform strategy is available);
3. $\langle\langle a\rangle\rangle\varphi$ expresses that agent a has a uniform strategy to enforce φ *from the current state* (but not necessarily even knows about it).

Thus, $\mathbb{K}_i\langle\langle i\rangle\rangle\varphi$ captures the notion of i's knowing *how to play* to achieve φ, while $\mathrm{K}_i\langle\langle i\rangle\rangle\varphi$ refers to knowing only *that a successful play is possible.* This extends naturally to abilities of coalitions.

In the lecture, we will also discuss some formal properties of constructive knowledge, and present a "canonical" form of formulae that allows to define the semantics entirely in terms of single states without losing expressivity of the language.

References: [Jamroga and Ågotnes, 2007; van der Hoek and Wooldridge, 2003; Schobbens, 2004].

Lecture 7: Characterising solution concepts

In lecture 2 we discussed game theory and introduced strategic games. A *solution* for a particular class of games is a description of the outcomes that may emerge. The *Nash equilibrium* is a solution concept which is one of the most important concepts in game theory. A collection of strategies, one for each agent, form a Nash equilibrium iff each strategy is the best response by that agent to the other strategies. A *weakly dominant* strategy is a strategy for an agent which is at least as good as any other strategy (no matter what strategies the other players play) for that agent.

We have seen how we can use logics such as Coalition Logic and ATL to express properties of game-like structures. But we haven't yet discussed an obviously interesting problem: formalising solution concepts. In this lecture we discuss how we can express solution concepts and related properties of strategic games in ATL-like logics. The properties we have discussed until now have been of the "what an agent (or a coalition) *can* do". Solution concepts, however, are concerned with what rational agents *will* do. And here is a shortcoming with ATL: there is no built-in mechanism for representing *preferences*. A simple way to extend ATL to incorporate preferences is (following [Baltag, 2002]) to add:

- Atomic propositions of the form $u_i \geq d$ where i is an agent and $d \in D$ and D is a finite domain, meaning that the utility of agent i is greater than or equal to d

With this kind of atomic propositions, it is straightforward to "simulate" a strategic game by a pointed CGS (M, q) (actions correspond to strategies, etc.). We then say that a pointed CGS (M, q) *corresponds* to a game Γ. It turns out, however, that even with this extension it is still difficult to characterise solution concepts such as the Nash equilibrium in ATL. A possible solution is to further extend ATL with:

- A *counterfactual* operator $C_i(\sigma_i, \varphi)$ where σ_i is (a name standing for) a strategy for agent i and φ is a formula, meaning "if i would play strategy σ, then φ".

Note that this promotes *stategies* to first-class citizens of the logical language. The semantics of the counterfactual expressions can be defined using model updates (lecture 4):

$$M, q \models C_i(\sigma_i, \varphi) \Leftrightarrow (M \dagger [\![\sigma_i]\!], q \models \varphi)$$

where $[\![\sigma]\!]$ is the strategy denoted by σ.

Weak dominance can now be logically characterised as follows:

$$wd_i(\alpha) \;\equiv\; \bigwedge_{v \in D} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

$$WD_i(\alpha) \;\equiv\; \bigwedge_{\beta \in \Upsilon_j} C_j(\beta, wd_i(\alpha)).$$

where Υ_j is the set of all strategy names for agent j. For any finite game Γ, strategy α is weakly dominant for i iff $M, q \models WD_i(\alpha)$, whenever (M, q) corresponds to Γ (why isn't $wd_i(\alpha)$ enough?).

Moving on to Nash equlibria, the fact that i's best response to k playing α_k is α_i can be formalised as:

$$BR_i(\alpha_k, \alpha_i) \equiv \quad C_k(\alpha_k, \bigwedge_{v \in D} ((\langle\langle i \rangle\rangle \bigcirc (u_i \geq v)) \rightarrow C_i(\alpha_i, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

and the fact that (α_1, α_2) is a Nash equilibrium can be expressed as follows:

$$NE(\alpha_1, \alpha_2) \equiv BR_1(\alpha_2, \alpha_1) \wedge BR_2(\alpha_1, \alpha_2)$$

For any finite game Γ, (α_1, α_2) is a Nash equilibrium iff $M, q \models NE(\alpha_1, \alpha_2)$, whenever M, q corresponds to Γ.

References: [van der Hoek *et al.*, 2005; Harrenstein *et al.*, 2003; Osborne and Rubinstein, 1994].

Lecture 8: Reasoning about rational play

For many games the number of all possible outcomes is excessive, although only some of them "make sense". We need a notion of rationality to discard the "less sensible" ones, and determine what should happen had the game been played by ideal players. Game theory defines a number of rationality criteria through so called *solution concepts*. We note that, while some solution concepts see rationality as a property of individual strategies (e.g., dominant strategy, iterated undominated strategy), the most prominent concepts (Nash equilibrium, Pareto optimality) refer to whole *strategy profiles*, i.e., combinations of strategies from all the agents in $\mathbb{A}\text{gt}$.

We showed in the previous lecture how game-theoretical solution concepts can be characterized in an ATL-like logic. Now, we discuss how one can use such characterizations to reason about the behavior and abilities of agents who conform with a given rationality criterion. To this end, we present the logic of ATLP ("ATL with Plausibility") that extends ATL with constructs for defining and reasoning about rational (or plausible) play of agents. $\mathbf{Pl}_A$ assumes that agents in A play rationally; this means, that the agents can only use strategy profiles that are *plausible* in the given model. In particular, $\mathbf{Pl}$ ($\equiv \mathbf{Pl}_{\mathbb{A}\text{gt}}$) imposes rational behavior on all agents in the system. Similarly, $\mathbf{Ph}$ ($\equiv \mathbf{Pl}_\emptyset$) disregards plausibility assumptions, and refers to all *physically* available scenarios. The model update operator ($\mathbf{set\text{-}pl}\ \omega$) allows to define (or redefine) the set of plausible strategy profiles to the ones described by plausibility term ω (in this sense, it implements *revision* of plausibility). Operator ($\mathbf{refn\text{-}pl}\ \omega$) enables *refining* the set of plausible strategy profiles, i.e. selecting a subset of the hitherto plausible profiles. With ATLP, we can for example say that $\mathbf{Pl}\,\langle\!\langle\emptyset\rangle\!\rangle\Box(\mathsf{closed}\wedge\mathbf{Ph}\,\langle\!\langle guard\rangle\!\rangle\bigcirc\neg\mathsf{closed})$: "It is plausible that the emergency door will always remain closed, but the guard retains the physical ability to open them"; or $(\mathbf{set\text{-}pl}\ \omega_{NE})\mathbf{Pl}\,\langle\!\langle a\rangle\!\rangle\Diamond\neg\mathsf{jail}_\mathsf{a}$: "Suppose that only playing Nash equilibria is rational, then agent a can plausibly reach a state where it is out of prison".

To define the semantics of ATLP, we extend CGS's to *concurrent game structures with plausibility*. Apart from an actual set of plausible strategy profiles Υ, a *concurrent game structure with plausibility* (CGSP) must specify the denotation of plausibility terms ω. It is defined via a *plausibility mapping* that maps each term to a set of strategy profiles.

Let $A, B \subseteq \mathbb{A}\text{gt}$ be groups of agents. We define collective strategy s_A to be *B-consistent* with a set of strategy profiles Υ iff the B's part of s_A appears as a part of some profile in Υ. Function $out(q, s_A, B)$ returns the set of paths that can result from the execution of s_A from state q on if we assume that only B-consistent strategy profiles are played. Now, the semantics of ATLP formulae can be given as below. Note that the satisfaction relation has an additional parameter: namely, the set of agents B to whom the rationality/plausibility assumptions apply.

$M, q \models_B p$ iff $p \in \pi(q)$;

$M, q \models_B \neg\varphi$ iff $M, q \not\models_B \varphi$;

$M, q \models_B \varphi \wedge \psi$ iff $M, q \models_B \varphi$ and $M, q \models_B \psi$;

$M, q \models_B \langle\!\langle A\rangle\!\rangle\gamma$ iff there is strategy s_A, B-consistent with Υ, such that $M, \lambda \models_B \gamma$ for all $\lambda \in out(q, s_A, B)$;

$M, \lambda \models_B \bigcirc\gamma$ iff $M, \lambda[1..\infty] \models_B \gamma$ (for $\Box$ and $\mathcal{U}$ analogously);

$M, q \models_B \mathbf{Pl}_A\, \varphi$ iff $M, q \models_A \varphi$

$M, q \models_B (\textbf{set-pl}\ \omega)\varphi$ iff $M', q \models_B \varphi$ where the new model M' is equal to M but the new set $\Upsilon_{M'}$ of plausible strategy profiles of M' is set to $[\![\omega]\!]^q_M$.

$M, q \models_B (\textbf{refn-pl}\ \omega)\varphi$ iff $M', q \models_B \varphi$ where M' is equal to M but $\Upsilon_{M'}$ set to $\Upsilon_M \cap [\![\omega]\!]^q$.

The "absolute" satisfaction relation $\models$ is given by $\models_\emptyset$.

During this lecture, we will also point out that temporal operators enable defining agents' outcomes in a game by temporal patterns rather than simply payoffs obtained at the end of the game. This gives rise to what we call *temporalized solution concepts* which can be used to define rationality assumptions in interaction scenarios that are not bounded in time.

References: [Bulling *et al.*, 2008].

Lectures 9 & 10: Axiomatisation of coalitional games

In lecture 2 we introduced non-cooperative games, and we discussed basic solution concepts in lecture 7. We move now to the theory of coalitional (or cooperative) games.

A *coalitional game* (without transferable payoff) is an $(m+3)$-tuple:

$$\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \dots, \sqsupseteq_m \rangle$$

where:

- $N = \{1, \dots, m\}$ is a non-empty set of *agents*;
- Ω is a non-empty set of *outcomes*;
- $V : (2^N \setminus \emptyset) \rightarrow 2^\Omega$ is the *characteristic function* of Γ, which for every non-empty coalition C defines the choices $V(C)$ available to C, so that $\omega \in V(C)$ means C can choose outcome ω; and
- $\sqsupseteq_i \subseteq \Omega \times \Omega$ is a complete, reflexive, and transitive *preference relation*, for each agent $i \in N$.

Solution concepts of coalitional games are concerned with such questions such as "which coalitons will form?", or, in other words, which coalitions are *stable*. The most well known solution concept is the *core*. The core of a coalitional game is the set of outcomes which can be chosen by the grand coalition N such that there does not exists a coalition $C \subseteq N$ which can choose another outcome which is better for all the agents in C. A key property of a coalitional game is whether the core is empty or not. Other solution concepts include *stable sets* and *the bargaining set*.

We want to be able to express properties of coalitional games in a logical language – in particular solution related properties such as non-emptyness of the core or core membership for some outcome. One possibility would be to have atomic propositions of the form $\omega \succeq_i \omega'$, where ω and ω' are explicit names of outcomes, and a modality $\langle C \rangle$ interpreted by the characteristic function in the spirit of Coalition Logic, such that for example $\langle C \rangle \omega$ would be true iff $\omega \in V(C)$. Core membership for ω could then be expressed as:

$$CM(\omega) \equiv \langle N \rangle \omega \wedge \neg \left[\bigvee_{C \subseteq N} \bigvee_{\omega' \in \Omega} (\langle C \rangle \omega') \wedge \bigwedge_{i \in C} (\omega' \succ_i \omega) \right]$$

and non-emptyness of the core as:

$$CNE \equiv \bigvee_{\omega \in \Omega} CM(\omega)$$

There are some disadvantages to such a solution: the expressions above depend on the set of outcomes Ω, and are thus different for games with different outcomes. Furthermore, the expressions are not very succinct – and in fact, for games with infinitely many outcomes they are not even well-formed formulae.

An alternative, which does not have these problems, is *Modal Coalitional Game Logic (MCGL)* [Ågotnes *et al.*, 2008]. MCGL is a normal modal logic. Above we discussed a way of using atomic propositions for the preference relations, and a modality for the characteristic function as in Coalition Logic. The approach in MCGL is, roughly speaking, the exact opposite: the characteristic function is modelled using atomic propositions while modalities are used for the preference relation. The main constructs of the language is an atom p_C and a modality $\langle C^s \rangle$ for each coalition C. The language is interpreted in the context of an outcome in a coalitional game. p_C means that C can choose the current outcome, while $\langle C^s \rangle \varphi$ means that there is some outcome which is strictly better than the current one for all the agents in C, in which φ is true (the s stands for "strictly"). Formally, when ω is an outcome in Γ (there is also a special modality $\langle D \rangle$):

- $\Gamma, w \models p_C$ iff $w \in V(C)$
- $\Gamma, w \models \langle C^s \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$ and not $w \sqsupseteq_i v$, and $\Gamma, v \models \phi$
- $\Gamma, w \models \langle D \rangle \phi$ iff $\Gamma, v \models \phi$ for some $v \neq w$

We write $[C^s]$ for the dual: $[C^s]\varphi \equiv \neg \langle C^s \rangle \neg \varphi$. Now, we can express core membership as follows:

$$MCM \equiv p_N \wedge \bigwedge_{C \subseteq N} [C^s] \neg p_C$$

We have that $(\Gamma, \omega) \models MCM$ iff ω is in the core of Γ.

Non-emptyness of the core can be expressed as follows:

$$MCNE \equiv MCM \vee \langle D \rangle MCM$$

$(\Gamma, \omega) \models MCNE$ iff $\Gamma \models MCNE$ iff the core of Γ is non-empty.

After introducing coalitional games, will will use most of these two lectures on discussing MCGL. In particular, we will go through its completeness proof, which is based on a technique by de Rijke [de Rijke, 1993] (see also [Blackburn *et al.*, 2001]) using the difference modality $\langle D \rangle$ as well as "converse" modalities.

References: [Ågotnes *et al.*, 2008; 2006; Osborne and Rubinstein, 1994; de Rijke, 1993; Blackburn *et al.*, 2001].

References

[Ågotnes *et al.*, 2006] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. On the logic of coalitional games. In P. Stone and G. Weiss, editors, *Proceedings of the Fifth International Conference on Autonomous Agents and Multiagent Systems (AA-MAS)*, pages 153–160, Hakodate, Japan, May 2006. ACM Press.

[Ågotnes *et al.*, 2008] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. Reasoning about coalitional games, 2008. Manuscript, to appear.

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

[Baltag, 2002] A. Baltag. A logic for suspicious players. *Bulletin of Economic Research*, 54(1):1–46, 2002.

[Blackburn *et al.*, 2001] P. Blackburn, M. de Rijke, and V. Venema. *Modal Logic*. Cambridge Univ. Press, 2001.

[Brihaye *et al.*, 2008] Thomas Brihaye, Arnaud Da Costa, François Laroussinie, and Nicolas Markey. ATL with strategy contexts. in preparation, 2008.

[Bulling *et al.*, 2008] N. Bulling, W.Jamroga, and J. Dix. Reasoning about rational agents in ATLP. Technical Report IfI-08-03, Clausthal University of Technology, 2008.

[Chatterjee *et al.*, 2007] Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Strategy logic. In Luís Caires and Vasco Thudichum Vasconcelos, editors, *CONCUR*, volume 4703 of *Lecture Notes in Computer Science*, pages 59–73. Springer, 2007.

[de Rijke, 1993] M. de Rijke. *Extended Modal Logic*. PhD thesis, University of Amsterdam, 1993.

[Emerson, 1990] E. A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 995–1072. Elsevier Science Publishers, 1990.

[Fagin *et al.*, 1995] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press: Cambridge, MA, 1995.

[Goranko and van Drimmelen, 2003] V. Goranko and G. van Drimmelen. Complete axiomatization and decidability of the Alternating-time Temporal Logic. Submitted, 2003.

[Halpern, 1995] J. Y. Halpern. Reasoning about knowledge: a survey. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *The Handbook of Logic in Artificial Intelligence and Logic Programming, Volume IV*, pages 1–34. Oxford University Press, 1995.

[Harrenstein *et al.*, 2003] B.P. Harrenstein, W. van der Hoek, J.-J. Meyer, and C. Witteveen. A modal characterization of Nash equilibrium. *Fundamenta Informaticae*, 57(2–4):281–321, 2003.

[Jamroga and van der Hoek, 2004] W. Jamroga and W. van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 63(2–3):185–219, 2004.

[Jamroga and Ågotnes, 2007] 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.

[Jamroga, 2003] W. Jamroga. Some remarks on alternating temporal epistemic logic. In B. Dunin-Keplicz and R. Verbrugge, editors, *Proceedings of Formal Approaches to Multi-Agent Systems (FAMAS 2003)*, pages 133–140, 2003.

[Jonker, 2003] G. Jonker. Feasible strategies in Alternating-time Temporal Epistemic Logic. Master thesis, University of Utrecht, 2003.

[Osborne and Rubinstein, 1994] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press: Cambridge, MA, 1994.

[Pauly, 2001] M. Pauly. *Logic for Social Software*. PhD thesis, University of Amsterdam, 2001.

[Pinchinat, 2007] Sophie Pinchinat. A generic constructive solution for concurrent games with expressive constraints on strategies. In Kedar S. Namjoshi, Tomohiro Yoneda, Teruo Higashino, and Yoshio Okamura, editors, *5th International Symposium on Automated Technology for Verification and Analysis*, volume 4762 of *Lecture Notes in Computer Science*, pages 253–267. Springer-Verlag, 2007.

[Schobbens, 2004] P. Y. Schobbens. Alternating-time logic with imperfect recall. *Electronic Notes in Theoretical Computer Science*, 85(2), 2004.

[van der Hoek and Pauly, 2006] W. van der Hoek and M. Pauly. Modal logic for games and information. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, pages 1077–1148. Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 2006.

[van der Hoek and Verbrugge, 2002] W. van der Hoek and R. Verbrugge. Epistemic logic: A survey. *Game Theory and Applications*, 8:53–94, 2002.

[van der Hoek and Wooldridge, 2002] W. van der Hoek and M. Wooldridge. Tractable multiagent planning for epistemic goals. In C. Castelfranchi and W.L. Johnson, editors, *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-02)*, pages 1167–1174. ACM Press, New York, 2002.

[van der Hoek and Wooldridge, 2003] W. van der Hoek and M. Wooldridge. Cooperation, knowledge and time: Alternating-time Temporal Epistemic Logic and its applications. *Studia Logica*, 75(1):125–157, 2003.

[van der Hoek *et al.*, 2004] W. van der Hoek, S. van Otterloo, and M. Wooldridge. Preferences in game logics. In *Proceedings of AAMAS-04*, 2004.

[van der Hoek *et al.*, 2005] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proceedings of AAMAS'05*, pages 157–164, 2005.

[van Otterloo and Jonker, 2004] S. van Otterloo and G. Jonker. On Epistemic Temporal Strategic Logic. *Electronic Notes in Theoretical Computer Science*, XX:35–45, 2004. Proceedings of LCMAS'04.

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

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

[Ågotnes *et al.*, 2007] T. Ågotnes, V. Goranko, and W. Jamroga. Alternating-time temporal logics with irrevocable strategies. In D. Samet, editor, *Proceedings of TARK XI*, pages 15–24, 2007.

[Ågotnes, 2006] Thomas Ågotnes. Action and knowledge in alternating-time temporal logic. *Synthese*, 149(2):377–409, 2006. Section on Knowledge, Rationality and Action.