Reasoning about Strategic Abilities: Agents with Truly Perfect Recall

Nils Bulling, Delft University of Technology
Wojciech Jamroga, Institute of Computer Science, Polish Academy of Sciences
Matei Popovici, POLITEHNICA University of Bucharest

In alternating-time temporal logic $\mathsf{ATL}^\star$, 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 paper, we investigate the alternative semantics for $\mathsf{ATL}^\star$ where the past is not forgotten in nested games. We show that the standard semantics of $\mathsf{ATL}^\star$ coincides with the "truly perfect recall" semantics for agents with perfect information and in case of so called "objective" abilities under uncertainty. On the other hand, the two semantics differ significantly for the most popular ("subjective") notion of ability under imperfect information. The same applies to the standard vs. "truly perfect recall" semantics of $\mathsf{ATL}^\star$ with persistent strategies. We compare the relevant variants of $\mathsf{ATL}^\star$ by looking at their their expressive power, sets of validities, and tractability of model checking.

CCS Concepts:• **Theory of computation** → **Modal and temporal logics;** *Logic and verification;*• **Computing methodologies** → **Multi-agent systems;**

Additional Key Words and Phrases: alternating-time temporal logic, multi-player games, reasoning about knowledge and time

1. INTRODUCTION

The *alternating-time temporal logic* $\mathsf{ATL}^\star$ and its fragment ATL [Alur et al. 1997; 2002] are logics which allow for reasoning about strategic interactions in multi-agent systems (MAS). The main idea is to extend the framework of temporal logic with the game-theoretic notion of *strategic ability*. Hence, $\mathsf{ATL}^\star$ enables to express statements about what agents (or groups of agents) can achieve. For example, $\langle\!\langle a \rangle\!\rangle \Diamond \mathsf{win_a}$ says that agent a has the ability to eventually win no matter what the other agents do, while $\langle\!\langle a, b \rangle\!\rangle \Box \mathsf{safe}$ expresses that agents a and b together can force the system to always remain in a safe state. Such properties can be useful for specification, verification and reasoning about interaction in agent systems. They have become especially relevant due to active development of algorithms and tools for verification where the "correctness" property is given in terms of strategic ability [Alur et al. 2000; Alur et al. 2001; Kacprzak and Penczek 2004; Lomuscio and Raimondi 2006; Chen et al. 2013; Huang and van der Meyden 2014; Busard et al. 2014; Pilecki et al. 2014; Lomuscio et al. 2015; Busard et al. 2015; Busard 2017; Jamroga et al. 2017; Jamroga et al. 2018]. Still, when verifying a system, one must first of all have a clear idea *what* is to be verified. An important challenge for model checking MAS is to define the correctness property in the right way. This means choosing the right language and the right semantics, one which accurately captures agents' abilities in a given context. The challenge is not only theoretical. When designing a system, specifying requirements, or verifying its properties, one must choose between many semantic variants of $\mathsf{ATL}^\star$ that start from different assumptions about the capabilities of agents. For instance, agents may be able to observe the full state of the system or only parts of it (perfect vs. imperfect information), and they may base their decisions on the current state only, or on the entire history of the game (perfect vs. imperfect recall) [Schobbens 2004; Jamroga and van der Hoek 2004]. Intermediate cases

 1539-9087/2010/03-ART39 $15.00
DOI: 0000001.0000001

of finite-memory strategies have also been studied [Vester 2013]. Moreover, agents can have objective or subjective ability to achieve their goals [Bulling and Jamroga 2014], their strategies can come with or without long-term commitment [Ågotnes et al. 2007; Brihaye et al. 2009], they can be assumed to play rationally [Bulling et al. 2008], be equipped with bounded resources [Alechina et al. 2004; Alechina et al. 2009; Alechina et al. 2010; 2011; Bulling and Farwer 2010b; 2010a], a cost-free mechanism for broadcasting information within a team [Dima et al. 2010; Guelev et al. 2011], and so on.

In this paper, we focus on the commonly accepted perfect recall semantics of strategic ability, proposed in [Alur et al. 2002; Schobbens 2004]. We point out that, for nested strategic modalities, it interprets formulae of $\mathsf{ATL}^\star$ in a counterintuitive way. Then, we study a modified the semantics that avoids the problem. Most importantly, we formally investigate the difference between the standard and the modified semantics in terms of valid sentences, expressive power, and complexity of model checking.

1.1. Contribution: Analysis of Perfect Recall in Logics of Strategic Ability

We begin our analysis by observing that the standard perfect recall semantics of $\mathsf{ATL}^\star$ has a counterintuitive flavor: despite using perfect recall strategies, agents may not have access to all of their past observations. More precisely, agents *forget* their past observations once they proceed to realize a sub-goal in the game. As an example, consider the formula $\langle\langle b, c\rangle\rangle \Diamond \langle\langle a, b\rangle\rangle \bigcirc \mathsf{married_{ab}}$ which expresses that Bob and Charles have a joint strategy to ensure that, at some point in the future, Alice and Bob will be able to get married. Agents' abilities rely on their knowledge; in case of perfect recall, one would assume that each agent can use all their past observations to determine their subsequent actions. However, the standard semantics of $\mathsf{ATL}^\star$ interprets the subformula $\langle\langle a, b\rangle\rangle \bigcirc \mathsf{married_{ab}}$ *in the original model*. This amounts to assuming that Bob, when looking for his best strategy to make $\langle\langle \boldsymbol{a}, \boldsymbol{b}\rangle\rangle \bigcirc \mathsf{married_{ab}}$ true, must ignore (or forget) all the observations that he has made while executing his strategy for $\langle\langle \boldsymbol{b}, \boldsymbol{c}\rangle\rangle \Diamond \langle\langle a, b\rangle\rangle \bigcirc \mathsf{married_{ab}}$.

The above feature has some logical and conceptual reasons. $\mathsf{ATL}^\star$ was designed to extend the branching-time temporal logic $\mathsf{CTL}^\star$ which in turn was built on the assumption that "the future behavior [of a concurrent program] depends only upon the current state, and not upon how that state was reached" so that it "reflects the essential properties of genuine concurrent programs" [Emerson and Halpern 1986]. In logical terms, this means stronger compositionality of the semantics, as the current state alone suffices to determine the truth value of the formula. In modeling terms, this amounts to assuming Markovian behavior of the system. The main problem is that, for agents with perfect recall of the past, the dynamics of their mental states is *not* Markovian.

In this paper, we study a modified semantics of $\mathsf{ATL}^\star$ where formulae are interpreted in finite sequences of states rather than single states of the system. We call the semantics *truly perfect recall* to emphasize that decisions within a strategy can refer to the whole history of observations made by the agents. We show that the modified semantics offers a significantly different view of agents' abilities from the original semantics of $\mathsf{ATL}^\star$. More precisely, we prove that if agents have imperfect information then $\mathsf{ATL}^\star$ with truly perfect recall differs from $\mathsf{ATL}^\star$ with standard perfect recall in terms of expressive power as well as valid sentences. The same can be shown for variants of $\mathsf{ATL}^\star$ that allow agents to irrevocably commit to their strategies [Ågotnes et al. 2007; Brihaye et al. 2009]. We also point out that truly perfect recall makes model checking $\mathsf{ATL}^\star$ harder than in the standard semantics.[1] Analogously to [Bulling and Jamroga 2014], we conclude that the truly perfect recall semantics corresponds to a different *class of games*, and allows for expressing different properties of those games, than the "classic" variants of $\mathsf{ATL}^\star$ from [Alur et al. 2002; Schobbens 2004].

[1] That is, the complexity significantly increases for decidable fragments of the problem.

1.2. Motivation

The work reported in this paper concerns modeling, specification, and reasoning about strategic abilities in multi-agent systems. We believe that, when doing any kind of reasoning, it is of utmost importance to *get the model and the property right*. It is important for modelers to understand the meaning of the formulae they write down, and thus there is a need to refine existing semantics if they produce counterintuitive interpretation of the formulae.

In $\mathsf{ATL}^\star$, the interaction between agents is usually understood as a "game" where the "proponents" (the coalition A in the formula $\langle\langle A\rangle\rangle\varphi$) may want to pursue the goal characterized by the temporal subformula φ. They win if there exists a collective strategy such that φ is guaranteed even for the most damaging response of the "opponents" (i.e., the agents outside A).[2] An important subclass of the games is formed by ones where the goal φ refers to abilities in another, nested game. Notice that strategic abilities in such games have a two-level structure: they refer to existence of a strategy that enables (or prevents) another strategy. We believe that the template corresponds to many important properties of multi-agent systems. In particular, reasoning about *policies* (e.g., security policies, social policies, etc.) often follows this conceptual structure. In Section 1.3, we present several motivating examples to make the intuition more concrete.

To solve any kind of game, the modeler must decide what capabilities he or she will associate with the players. In particular, one needs to decide what kind of memory (or recall) the proponents have. Typical options are: perfect recall (agents in A remember all the past observations), finite or bounded memory (limited by a finite or bounded number of memory cells, or a restricted time window), and memoryless agents (whose memory is fully encoded in their current local state). Each option has its uses, and fits different scenarios. The only constraint is conceptual consistency. For instance, a semantics based on the perfect recall assumption should not admit a formula saying that agent a knows a strategy to guarantee p in the next step ($\langle\langle a\rangle\rangle\bigcirc\mathsf{p}$), and yet she will not know that p holds, even if she follows the same strategy ($\neg\langle\langle a\rangle\rangle\bigcirc K_a\mathsf{p}$). This is the issue that we focus on in this paper.

1.3. Motivating Examples

The main point that we raise concerns the semantics of formulae that involve nested strategic modalities. That is, formulae of $\mathsf{ATL}^\star$ that specify agents' ability to endow someone with (or deprive of) the ability to achieve a given goal. In other words, we are concerned with specifications that address the existence of strategies which enable (or disable) other strategies. At the first glance, such formulae may seem rather esoteric, and not likely to be encountered in specifications of actual systems. Below we present a number of examples which demonstrate the opposite, namely that the ability to influence abilities can be extremely important. Hence, the right semantics of nested strategic modalities is also of utmost importance for the soundness of reasoning and verification.

Example 1.1 (*Coercion in voting*). Consider a very simple voting scenario with two agents: the voter v and the coercer c. The voter casts her vote for a selected candidate $i \in \{1,\dots,n\}$ (action $vote_i$). Upon exit from the polling station, the voter can hand in a proof of how she voted to the coercer (action $give$) or refuse to hand in the proof (action ng). The proof may be a certified receipt from the election authorities, a picture of the ballot taken with a smartphone, etc.: anything that the coercer will consider believable. After that, the coercer can either punish the voter (action pun) or not punish (action np). A simple model of the scenario for $n = 2$ is shown in Figure 1. Proposition $\mathsf{vote_i}$ labels the states where the voter has already voted for candidate i, and proposition end indicates that the election is over. Proposition pun labels the states where the voter has been punished. The indistinguishability relation for the coercer is depicted by dotted lines.

Suppose that the coercer favors candidate 1, and wants to force the voter to vote for that candidate. The formula $\langle\langle c\rangle\rangle\Box\big(\mathsf{end} \to (\mathsf{pun} \leftrightarrow \neg\mathsf{voted_1})\big)$ expresses that c has a strategy to make sure that,

[2]In game-theoretic terms, this can be understood as a 2-player zero-sum extensive form game with binary payoffs over infinite runs, where the rationality criterion is given by maxmin.

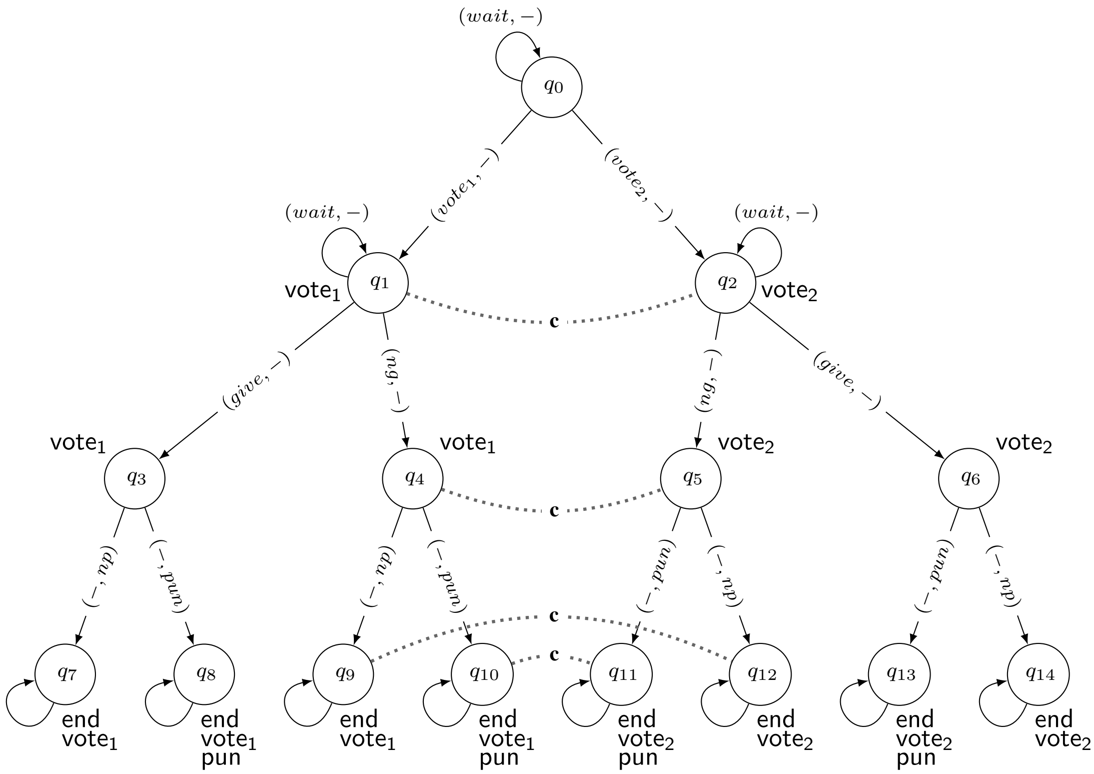


Fig. 1. A simple model of voting and coercion

whenever the election has come to an end, the voter is punished if, and only if, she did not vote for 1. Based on that, we can use the formula

$$\langle\langle v\rangle\rangle\Box\Big(\neg\langle\langle c\rangle\rangle\Box\big(\mathsf{end}\rightarrow(\mathsf{pun}\leftrightarrow\neg\mathsf{voted}_1)\big)\Big)$$

to specify that the voter can always prevent coercion.[3]

Example 1.2 (*Social fairness*). The Ministry of Education should strive to ensure fair access to university education across the population, e.g., by funding scholarships and stipends for students, subsidizing infrastructure in underdeveloped regions, etc. The objective is that everybody should be given an opportunity to study, provided that they choose to do so, and can conceivably make it. Let $\mathsf{E}\Diamond\psi$ mean "there is at least one possible path where ψ eventually holds," i.e., it is at least conceivable that ψ can become true at some future moment. Also, suppose proposition $\mathsf{study_s}$ models the situation where s is admitted at the university. Then, $\mathsf{E}\Diamond\mathsf{study_s}$ expresses that there exists an objective possibility that a is admitted. On the other hand, $\langle\langle s\rangle\rangle\Diamond\mathsf{study_s}$ expresses that s has the ability to get admitted, no matter what other agents do. Now, the requirement that "the Ministry should be able to provide fair access to university education" can be captured by the following formula:

$$\langle\langle ministry\rangle\rangle\Box\bigwedge_{s\in\mathbb{A}\mathrm{gt}}\Big(\mathsf{E}\Diamond\mathsf{study_s}\rightarrow\langle\langle s\rangle\rangle\Diamond\mathsf{study_s}\Big).$$

[3]The symbols, notational conventions, and semantics of formulae will be properly introduced in Section 2. Still, we hope that the reader can – already at this point – get an idea of the kind of abilities that we study in the paper.

Example 1.3 (*Robot soccer*). A possible task specification for a member of a RoboCup team is:

$$\langle\langle a\rangle\rangle\Diamond\Big(\mathsf{possession_b}\wedge\langle\langle b\rangle\rangle\Diamond\mathsf{goal_b}\Big).$$

That is, the robot a should seek a strategy to pass the ball to its teammate b (i.e., make proposition $\mathsf{possession_b}$ true) in such a way that b can score a goal for their team (i.e. make $\mathsf{goal_b}$ true).

Example 1.4 (*TCP*). The following specification describes an objective of the congestion control mechanism in the TCP protocol:

$$\langle\langle sender\rangle\rangle\Box(\mathsf{send\_pkt}\rightarrow\langle\langle receiver\rangle\rangle\Diamond\mathsf{ack\_pkt})$$

expressing that a sender can maintain that whenever a packet is sent (i.e. proposition $\mathsf{send\_pkt}$ is true), the receiver has a strategy to eventually acknowledge it (i.e. to make $\mathsf{ack\_pkt}$ true). Such a strategy may include actions such as replying with a delay, so that the receiver buffer does not overflow.

Example 1.5 (*Distribution of cryptographic keys*). Public key cryptography is based on generation of a pair of keys (sk_a, pk_a) where sk_a is agent a's secret key, and pk_a is his public key. The secret key is known only by a, whereas the public key is openly available to everybody (e.g., posted on the web). They serve dual cryptographic functions, i.e., messages encrypted by pk_a can be decrypted only using sk_a, and vice versa. The keys can be used for either communication or authentication. If another agent wants to send a private message to a, she can encrypt the message with pk_a and send it to a (who is the only agent possessing the key to decrypt it). If a wants to authenticate himself to another agent, he sends her a message plus its copy encrypted with his secret key (typically the communication is supposed to be private, so both parts are additionally encrypted with b's public key). When b decrypts the second part with a's public key and it matches the first part, then the message must have originated from a. However, this only works when the process of key distribution is trustworthy, i.e., when a and b know that the public keys pk_b, pk_a indeed come from b and a, respectively.

Suppose $K_i m$ denotes that agent i knows or reads the content of message m. Then, the formula $\langle\langle a\rangle\rangle(\Diamond K_b m\wedge\Box\neg K_c m)$ expresses that a can ensure that b will eventually read the message ($\Diamond K_b m$) and at the same time prevent c from ever reading the message ($\Box\neg K_c m$). Here c is a potentially malitious agent.

If we are mainly interested in communication then the goal of key exchange between agents a and b can be specified as:

$$\langle\langle a,b\rangle\rangle\Diamond\Box\bigwedge_{m}\bigwedge_{c\neq a,b}\Big(\langle\langle a\rangle\rangle(\Diamond K_b m\wedge\Box\neg K_c m)\wedge\langle\langle b\rangle\rangle(\Diamond K_a m\wedge\Box\neg K_c m)\Big).$$

In what follows, we will use a simplified working example to illustrate the main definitions in an intuitive way. We point out, however, that the working example shares the most important feature with the above motivating scenarios, in the sense that it asks about existence of a strategy to provide an agent with (or deprive of) specific strategic ability.

1.4. Related Work

An important strand in research on ATL^* emerged in quest of the "right" semantics for strategic ability for a specific setting. ATL was combined with epistemic logic [van der Hoek and Wooldridge 2003; Jamroga and van der Hoek 2004], and several semantic variants were defined for various assumptions about agents' memory [Schobbens 2004; Jamroga and van der Hoek 2004; Ågotnes and Walther 2009; Vester 2013] and available information [Schobbens 2004; Jamroga and van der Hoek 2004; Ågotnes 2006; Jamroga and Ågotnes 2007], cf. also [Bulling and Jamroga 2014; Ågotnes et al. 2015] for a broader discussion. Moreover, many conceptual extensions have been considered, e.g., with explicit reasoning about strategies [van der Hoek et al. 2005; Walther et al. 2007; Chatterjee et al. 2010; Mogavero et al. 2010a; Mogavero et al. 2014; Huang and van der Meyden 2014;

Belardinelli 2014; Berthon et al. 2017], bounded resources [Alechina et al. 2009; Alechina et al. 2010; Bulling and Farwer 2010b; Alechina et al. 2017], rationality assumptions in the form of game-theoretic solution concepts [Bulling et al. 2008], mechanisms for coordination within teams [Hawke 2010; van Ditmarsch and Knight 2014], and persistent commitment to strategies [Ågotnes et al. 2007; Brihaye et al. 2009]. Another strand of papers considered variants of coalitional ability where members of the team were assumed to have a unified view of the state of the system, either by sharing their information at no cost throughout the game [Guelev and Dima 2008; Dima et al. 2010; Guelev et al. 2011], or by aggregating their uncertainty at each step [Diaconu and Dima 2012].

Several papers have come close to what we investigate here. The dynamics of knowledge in the seminal book [Fagin et al. 1995] was interpreted in infinite runs that included both the past and the future part of temporal path. A very similar, history-based semantics for $\mathsf{ATL}^\star$ with imperfect information and perfect recall (called ATEL-R*) was in fact considered as early as [Jamroga and van der Hoek 2004], but it was not studied further. Later, [Bulling and Dix 2010] also used a semantic relation that referred to the history of events. The focus of that work, however, was coalition formation, and the history-based semantics was used to keep track of the satisfaction of agents' goals. Furthermore, [Mogavero et al. 2010b] used history-based semantics to study a variant of $\mathsf{ATL}^\star$ where coalition A can enforce property ψ in state q if, on all plays enforced by A from q, ψ is true when evaluated *from the beginning of the game*. This differs from our work as follows. First, we only use histories to propagate past observations to strategies that are witnesses to nested strategic modalities, and not to evaluate the "winning condition" (our path subformulae are purely future-oriented). Secondly, [Mogavero et al. 2010b] look only at the perfect information setting, whereas we consider the cases of both perfect and imperfect information (with and without strategic commitment). Accordingly, our results are very much different. While [Mogavero et al. 2010b] show that their "relentful $\mathsf{ATL}^\star$" has the same expressive power and model checking complexity as standard $\mathsf{ATL}^\star$, our "$\mathsf{ATL}^\star$ with truly perfect recall" has incomparable expressive power and different model checking complexity. Moreover, it generates a different set of validities than standard $\mathsf{ATL}^\star$.

History-based semantics for $\mathsf{ATL}^\star$ with imperfect information was also used in several papers by Dima and his coauthors [Dima et al. 2010; Guelev et al. 2011; Diaconu and Dima 2012], and the focus of those papers was very close to what we study here. The differences are as follows. First, we look at the "truly perfect recall" semantics of $\mathsf{ATL}^\star$ in abstract concurrent game structures, whereas Dima et al. define the semantics in interpreted systems of infinite runs with perfect recall. Secondly, the semantics of coalitional play in [Dima et al. 2010; Guelev et al. 2011; Diaconu and Dima 2012] is seriously restricted by assuming that each coalition uses a single, aggregated indistinguishability relation; in fact, it can be argued that this amounts to treating coalitions as single agents in disguise, cf. [Kaźmierczak et al. 2014]. Thirdly, we focus on *comparing* the new semantics to the standard semantics, while Dima et al. concentrate on development of axiomatic systems and model checking algorithms for their variants of $\mathsf{ATL}^\star$. Fourthly, we also investigate the impact of truly perfect recall in the case when persistent strategic commitment is allowed.[4]

Finally, [Belardinelli et al. 2017] use the same semantics as we do here. Again, the difference lies in the purpose of the studies, and the obtained resuls. While we investigate the differences between the standard, Markovian semantics of $\mathsf{ATL}^\star$ and the "no forgetting" semantics, the focus of [Belardinelli et al. 2017] is on model checking – in particular, looking for decidable subclasses of the generally undecidable problem.

1.5. Structure of the Article

The paper is structured as follows. We introduce the classic variants of alternating-time temporal logic with perfect recall and perfect/imperfect information in Section 2. In Section 3, we point out the "forgetting" phenomenon in the classical semantics of $\mathsf{ATL}^\star$, and present the truly perfect recall semantics that avoids it. In Section 4 we compare the expressive powers of the variants of $\mathsf{ATL}^\star$

[4] A variant of $\mathsf{ATL}^\star$ that combines imperfect information, truly perfect recall, and persistent strategies has been considered in [Guelev and Dima 2012]. We discuss the relationship to our work in Section 7.2.

with the standard vs. truly perfect recall semantics. In Section 5 we do the same with respect to the validity sets induced by the two semantics. In Section 6, we briefly address the impact of truly perfect recall on the computational complexity of model checking for decidable fragments of the problem. Finally, Section 7 looks at the difference between standard perfect recall and truly perfect recall in the presence of persistent strategies. We conclude our work in Section 8, and discuss some directions for future research.

The material presented in this article is based on the conference papers [Bulling et al. 2013; 2014]. It extends the conference versions with detailed proofs, carefully constructed motivating and working examples, and formal analysis of truly perfect recall under strategic commitment. Moreover, we extend the analysis to other "modes" of ability under imperfect information – in particular, we present a surprising result for the *objective* interpretation of ability. We also add a brief discussion on the impact of truly perfect recall on the effectiveness and complexity of model checking $\mathsf{ATL}^\star$.

2. REASONING ABOUT STRATEGIC ABILITY

In this section, we briefly recall the main concepts behind $\mathsf{ATL}^\star$ and its variants.

2.1. Syntax of Alternating-Time Temporal Logic

$\mathsf{ATL}^\star$ [Alur et al. 1997; 2002] can be seen as a generalization of the branching time logic $\mathsf{CTL}^\star$, with the path quantifiers E and A being replaced by *strategic modalities* $\langle\!\langle A \rangle\!\rangle$. The formula $\langle\!\langle A \rangle\!\rangle\gamma$ expresses that group A has a *collective strategy* to enforce the temporal property γ where γ can include the temporal operators $\bigcirc$ ("next"), and $\mathcal{U}$ ("until"). Formally, let Π be a countable set of atomic propositions, and $\mathbb{A}\mathrm{gt}$ be a finite nonempty set of agents. The language of $\mathsf{ATL}^\star$ is given by the following grammar:

$$\begin{array}{l} \varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\!\langle A \rangle\!\rangle\gamma, \\ \gamma ::= \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc\gamma \mid \gamma\,\mathcal{U}\,\gamma, \qquad \text{where } A \subseteq \mathbb{A}\mathrm{gt} \text{ and } p \in \Pi. \end{array}$$

We define "sometime in the future" as $\Diamond\gamma \equiv \top\,\mathcal{U}\,\gamma$ and "always in the future" as $\Box\gamma \equiv \neg\Diamond\neg\gamma$. Formulae φ and γ are called *state* and *path formulae* of $\mathsf{ATL}^\star$, respectively. State formulae constitute the language of $\mathsf{ATL}^\star$. By requiring that each temporal operator is immediately preceded by a strategic modality, we obtain the sub-language ATL; for example, $\langle\!\langle A \rangle\!\rangle\Diamond p$ is an ATL formula but $\langle\!\langle A \rangle\!\rangle\Diamond\Box p$ and $\langle\!\langle A \rangle\!\rangle(\Diamond p \wedge \Diamond r)$ are not.

2.2. Models: Imperfect Information Concurrent Game Structures

We interpret $\mathsf{ATL}^\star$ formulae over *imperfect information concurrent game structures* (iCGS) [van der Hoek and Wooldridge 2003; Schobbens 2004]. An iCGS is given by a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o, \{\sim_a \mid a \in \mathbb{A}\mathrm{gt}\}\rangle$ consisting of a nonempty finite set of all agents $\mathbb{A}\mathrm{gt} = \{1, \ldots, k\}$, a nonempty set of states St, a set of atomic propositions Π and their valuation $\pi : \Pi \to 2^{St}$, and a nonempty finite set of (atomic) actions Act. Function $d : \mathbb{A}\mathrm{gt} \times St \to 2^{Act}$ defines nonempty sets of actions available to agents at each state; we will usually write $d_a(q)$ instead of $d(a,q)$. Function o is a (deterministic) transition function that assigns the outcome state $q' = o(q, \alpha_1, \ldots, \alpha_k)$ to each state q and tuple of actions or *action profile* $\langle \alpha_1, \ldots, \alpha_k \rangle$ such that $\alpha_i \in d_i(q)$ for $1 \leq i \leq k$. Finally, each $\sim_a \subseteq St \times St$ is an equivalence relation that represents the indistinguishability of states from agent a's perspective.[5] We assume that agents have identical choices in indistinguishable states ($d_a(q) = d_a(q')$ whenever $q \sim_a q'$). We also assume that collective knowledge is interpreted in the sense of "everybody knows', i.e., $\sim_A = \bigcup_{a \in A} \sim_a$. For mathematical completeness, we define $\sim_\emptyset$ as the identity relation. We will use $[q]_A = \{q' \mid q \sim_A q'\}$ to refer to A's epistemic image of state q. Note that the perfect information concurrent game structures

[5]The relations capture *observational* indistinguishability. The *knowledge* that an agent collects by means of subsequent observations is not encoded in the model but rather in the constraints on strategies that the agent is allowed to play, see also the remark after Example 2.1.

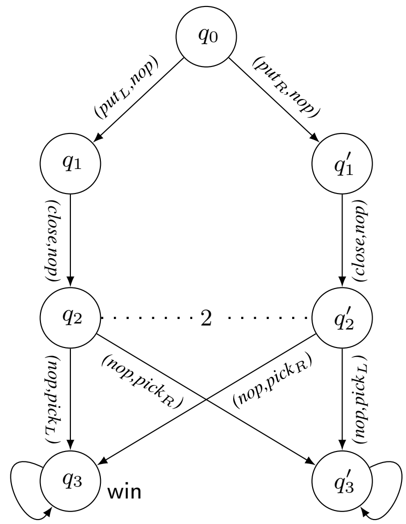


Fig. 2. The iCGS M_1 describing the *shell* game. Tuples (α_1, α_2) represent the action profiles. α_1 denotes an action of player 1—the shuffler—and action α_2 of player 2—the guesser. The dotted line represents 2's indistinguishability relation; State q_3 is labelled with the only proposition win. For example, when the guesser plays action $pick_R$ in state q_2 the game proceeds to state q'_3. nop indicates the "do nothing" action.

(CGS) from [Alur et al. 2002] can be seen as a specific type of iCGS that assumes each $\sim_a$ to be the identity relation.

Example 2.1 (*Shell game*). Consider model M_1 in Figure 2 that depicts a simple version of the *shell game*. There are two players: the *shuffler* 1 and the *guesser* 2. Initially, the shuffler places a ball in one of two shells (the left or the right). The shells are open, and the guesser can see the location of the ball. Then the shuffler turns the shells over, so that the ball becomes hidden. The guesser wins if he picks up the shell containing the ball. Formally: $\mathbb{A}\text{gt} = \{1, 2\}$, $St = \{q_0, q_1, q_2, q_3, q'_1, q'_2, q'_3\}$, $\Pi = \{\mathsf{win}\}$, $\pi(q_3) = \{\mathsf{win}\}$, $Act = \{put_L, put_R, pick_L, pick_R, close, nop\}$, $d_1(q_0) = \{put_L, put_R\}$, $d_1(q_1) = d_1(q'_1) = \{close\}$, $d_1(q_2) = d_1(q'_2) = d_2(q_0) = d_2(q_1) = d_2(q'_1) = \{nop\}$, $d_1(q_2) = d_1(q'_2) = \{pick_L, pick_R\}$, $d_i(q_3) = d_i(q_4) = \{nop\}$ for $i \in \mathbb{A}\text{gt}$. $q \sim_i q$ for all $q \in St$ and $i \in \mathbb{A}\text{gt}$. Also, $q_2 \sim_2 q'_2$. The function o is illustrated in Figure 2. Obviously, this is a very simplified version of the shell game as the shuffler does not even shuffle the shells; he simply places the ball in one of them and closes them. However, the example is rich enough to point out the limitations of the standard semantics of $\mathsf{ATL}^\star$.

Two remarks are in order. First, the relation $\sim_a$ encodes a's (in)ability to distinguish pairs of states, based on the qualities encapsulated in those states. That is, $q \sim_a q'$ iff q and q' *look* the same to a, independent of the history of events that led to them. If one assumes that the agent has external memory that allows her to remember the history of past events, this must be represented by an indistinguishability relation on *histories*, introduced in the next paragraph. Secondly, in order to describe an actual game, we also need to fix the initial state of an iCGS. A pair (M, q) consisting of an iCGS M and a state of M is called a *pointed iCGS*.

A *history* h is a finite sequence of states $q_0 q_1 \ldots q_n \in St^+$ which results from the execution of subsequent transitions; that is, there must be an action profile connecting q_i with q_{i+1} for each $i = 0, \ldots, n-1$. Two histories $h = q_0 q_1 \ldots q_n$ and $h' = q'_0 q'_1 \ldots q'_m$ are *indistinguishable for agent* a (denoted $h \approx_a h'$) iff $n = m$ and $q_i \sim_a q'_i$ for $i = 0 \ldots n$. This corresponds to the notion of *synchronous perfect recall* in temporal-epistemic logic [Fagin et al. 1995]. We also extend the indistinguishability relation over histories $\approx_a$, to groups: $\approx_A = \bigcup_{a \in A} \approx_a$. We write $h \circ h'$ to refer to the concatenation of the histories h, h' and *last*(h) to refer to the first and last state from history h, respectively. $\Lambda_M^{fin}(q)$ is the set of all histories in model M starting from state q, and $\Lambda_M^{fin} = \bigcup_{q \in St} \Lambda_M^{fin}(q)$ is the set of all histories in model M.

A *path* $\lambda = q_0 q_1 q_2 \ldots$ is an infinite sequence of states such that there is a transition from each q_i to q_{i+1}. We write $h \circ \lambda$, where $h = q_0' q_1' \ldots q_n'$ to refer to the path $q_0' q_1' \ldots q_n' q_0 q_1 q_2 \ldots$ obtained by concatenating h and λ, provided that there is a transition from q_n' to q_0. We use $\Lambda_M(q)$ to refer to the set of paths in M that start in state q, and define $\Lambda_M := \bigcup_{q \in St_M} \Lambda_M(q)$ to be the set of paths in M. We use $\lambda[i]$ to denote the ith position on path λ (starting from $i = 0$), $\lambda[i,j]$ (with $j \geq i$) to denote the history $q_i \ldots q_j$, and $\lambda[i,\infty]$ to denote the subpath of λ starting from i. Whenever the model is clear from context, we shall omit the subscript. As we will see later, the semantics of formulae is defined over paths. The truth, however, essentially depends on the sequence of propositional labels of each state and not on the name of the state. Hence, we say that two paths λ and λ' are *propositionally equivalent*, in notation $\lambda \equiv \lambda'$ if, and only if, $\lambda[i] \in \pi(p)$ iff $\lambda'[i] \in \pi(p)$ for all $i \in \mathbb{N}$ and $p \in \Pi$.

2.3. Strategies and Their Outcomes

A *strategy* of agent a is a conditional plan that specifies what a is going to do in each situation. It makes sense, from a conceptual and computational point of view, to distinguish between two types of strategies: an agent may base its decision on the current state or on the whole *history* of events that have happened. In this paper, we consider only the latter case. A *perfect information* strategy (*I-strategy* for short) is a function $s_a : St^+ \to Act$ such that $s_a(q_0 \ldots q_n) \in d_a(q_n)$ for all $q_0 \ldots q_n \in St^+$. An *imperfect information strategy* (*i-strategy*) must be additionally *uniform*, in the sense that $h \approx_a h'$ implies $s_a(h) = s_a(h')$. A *collective x-strategy* s_A with $x \in \{I, i\}$, is a tuple of *x*-strategies, one per agent in A. In particular, for imperfect information, each individual strategy s_a of s_A must be uniform. We also note that uniformity restricts each strategy component of s_A with respect to the corresponding individual indistinguishability relation $\sim_a$ and not to the one corresponding to the group: $\sim_A$. We use $s_A|_a$ to denote agent a's part of the collective strategy s_A, and $s_\emptyset$ to denote the empty profile which is the only strategy of the empty coalition.

The function $out_M(h, s_A)$ returns the set of all paths in M starting with history h, that can occur when s_A is executed. Formally:

$out_M(h, s_A) = \{h \circ \lambda = q_0 q_1 q_2 \ldots \mid$ such that for each $i \geq |h|$ there exists $\langle \alpha_{a_1}^{i-1}, \ldots, \alpha_{a_k}^{i-1} \rangle$ such that $\alpha_a^{i-1} \in d_a(q_{i-1})$ for every $a \in \mathbb{A}\mathrm{gt}$, $\alpha_a^{i-1} = s_A|_a(q_0 q_1 \ldots q_{i-1})$ for every $a \in A$, and $o(q_{i-1}, \alpha_{a_1}^{i-1}, \ldots, \alpha_{a_k}^{i-1}) = q_i\}$.

Example 2.2 (*Strategies and outcomes*). Consider model M_1 from Example 2.1, coalition $A = \{1, 2\}$, and its collective strategy $s_A = (s_1, s_2)$ where $s_1(q_0) = put_L$, $s_1(q_0 q_1) = s_1(q_0 q_1') = close$ and $s_2(q_0 q_1 q_2) = pick_L$. The values of $s_a(h)$ are unimportant for all the other combinations of $h \in \Lambda_{M_1}^{fin}$ and $a \in A$. If s_A is executed in state q_0, we have $out_M(q_0, s_A) = \{q_0 q_1 q_2 q_3^\omega\}$.

In the same model, consider strategy $s_A' = (s_1', s_2')$, where s_1' assigns an arbitrary permissible action to each history, $s_2'(q_0 q_1 q_2) = s_2'(q_0 q_1' q_2') = pick_L$ and $s_2'(h) = nop$ for all other histories. If $q_0 q_1 q_2$ is the current game development (the initial state was q_0, followed by q_1 and q_2), then $out_M(q_0 q_1 q_2, s_A') = \{q_0 q_1 q_2 q_3^\omega\}$. At the same time, if $s_2'(q_2) = s_2'(q_2') = pick_L$ and $s_2'(h) = nop$ for all other histories, then — if the game starts in q_2 — we have: $out_M(q_2, s_A') = \{q_2 q_3^\omega\}$.

Function $plays_M^x(h, s_A)$ returns the set of paths which agents from A consider possible if the game started with h and strategy s_A is executed. For perfect information, $plays_M^I(h, s_A) = out_M(h, s_A)$. For imperfect information, $plays_M^i(h, s_A)$ includes also the paths that A *think* might occur, i.e., ones starting from histories that are indistinguishable from A's point of view:

$$plays_M^x(h, s_A) = \begin{cases} out_M(h, s_A) & \text{for } x = I \\ \bigcup_{h \approx_A h'} out_M(h', s_A) & \text{for } x = i \end{cases}$$

Example 2.3 (*Subjective outcome of a strategy*). Let us revisit the strategies presented in Example 2.2. Recall that $out_M(q_2, s'_A) = \{q_2 q_3^\omega\}$, and observe that $out_M(q'_2, s'_A) = \{q'_2 q_3'^\omega\}$. Hence $plays_M^i(q_2, s'_A) = \{q_2 q_3^\omega, q'_2 q_3'^\omega\}$: the history of the game is q_2 and since coalition A cannot distinguish q_2 from q'_2, both the "*winning*" and "*losing*" paths are considered possible when s'_A is executed.

Note that the above definitions of functions *out* and *plays* are slightly more general than the ones from [Alur et al. 2002; Schobbens 2004; Bulling and Jamroga 2014]: outcome paths are constructed given an initial *sequence of states* rather than a single state. This will prove convenient when we define the truly perfect recall semantics of $\mathsf{ATL}^\star$ in Section 3.

2.4. Standard Perfect Recall Semantics

Let M be an iCGS and $\lambda \in \Lambda_M$. The *(standard perfect recall) semantics* of $\mathsf{ATL}^\star$ can be defined via relation $\models_x$, parameterized with $x \in \{i, I\}$:

$M, \lambda \models_x p$ iff $\lambda[0] \in \pi(p)$ (where $p \in \Pi$);
$M, \lambda \models_x \neg\varphi$ iff $M, \lambda \not\models_x \varphi$;
$M, \lambda \models_x \varphi_1 \wedge \varphi_2$ iff $M, \lambda \models_x \varphi_1$ and $M, \lambda \models_x \varphi_2$;
$M, \lambda \models_x \langle\!\langle A \rangle\!\rangle\varphi$ iff there is a collective x-strategy s_A such that, for each $\lambda' \in plays_M^x(\lambda[0], s_A)$, we have $M, \lambda' \models_x \varphi$;
$M, \lambda \models_x \bigcirc\varphi$ iff $M, \lambda[1, \infty] \models_x \varphi$;
$M, \lambda \models_x \varphi_1 \,\mathcal{U}\, \varphi_2$ iff there is $i \in \mathbb{N}_0$ such that $M, \lambda[i, \infty] \models_x \varphi_2$ and for all $0 \le j < i$, we have that $M, \lambda[j, \infty] \models_x \varphi_1$.

Also, for a state q and a state formula φ, we define $M, q \models_x \varphi$ iff $M, \lambda \models_x \varphi$ for all $\lambda \in \Lambda_M(q)$.

The logical system with the syntax given by $\mathsf{ATL}^\star$ and the semantics by $\models_x$ will be referred to as $\mathsf{ATL}^\star_\mathsf{x}$. That is, $\mathsf{ATL}^\star_\mathsf{I}$ refers to the language of $\mathsf{ATL}^\star$ with the perfect information semantics $\models_I$, whereas $\mathsf{ATL}^\star_\mathsf{i}$ is $\mathsf{ATL}^\star$ with the imperfect information semantics $\models_i$. Moreover, formula φ is *valid* in $\mathsf{ATL}^\star_\mathsf{x}$ iff $M, q \models_x \varphi$ for all M and states q in M.

Example 2.4 (*Shell game ctd.*). Consider the iCGS M_1 from Figure 2, and assume q_2 is the initial state of the game. It is easy to see that $M_1, q_2 \models_I \langle\!\langle 2 \rangle\!\rangle\Diamond\mathsf{win}$: under perfect information, the guesser can win by choosing the left shell in q_2. On the other hand, $M_1, q_2 \not\models_i \langle\!\langle 2 \rangle\!\rangle\Diamond\mathsf{win}$: under imperfect information, the guesser has no *uniform* strategy that succeeds from both q_2 and q'_2. Finally, if the game begins in q_0 then the guesser can win ($M_1, q_0 \models_i \langle\!\langle 2 \rangle\!\rangle\Diamond\mathsf{win}$) by using the i-strategy: "play $pick_L$ (resp. $pick_R$) after history $q_0 q_1 q_2$ (resp. $q_0 q'_1 q'_2$)". The strategy is uniform as both histories are distinguishable for the guesser[6].

Informally, $M, \lambda \models_I \langle\!\langle A \rangle\!\rangle\varphi$ holds iff there exists a collective I-strategy s_A such that φ holds on all outcome paths that result from executing s_A after history $\lambda[0]$. Thus, in the case of standard $\mathsf{ATL}^\star$ (with perfect information), the history is always limited to the *current state*, and thus the previous states of the play are completely ignored. Notice also that $M, q \models_i \langle\!\langle A \rangle\!\rangle\varphi$ requires A to have a single strategy that is successful in *all* states indistinguishable from q for *any* member of the coalition.

Remark 2.5 (*Embedding path quantifiers in* $\mathsf{ATL}^\star$). The path quantifiers of $\mathsf{CTL}^\star$ can be translated to $\mathsf{ATL}^\star$ as follows: $\mathsf{A}\varphi \equiv \langle\!\langle \emptyset \rangle\!\rangle\varphi$ ("for all paths, φ") and $\mathsf{E}\varphi \equiv \neg\mathsf{A}\neg\varphi$ ("there is a path

[6]We note that the guesser has no memoryless strategy (i.e. a strategy that assigns actions to states only) to win, as such a strategy had to assign the same choices to q_2 and q'_2.

such that φ"). We also note that, when all the agents have perfect information, the existential path quatifier can be equivalently expressed by $\mathsf{E}\varphi \equiv \langle\langle \mathbb{A}\mathrm{gt} \rangle\rangle \varphi$.

Remark 2.6 (*Reasoning about knowledge*). We note in passing that epistemic operators $K_a\varphi$ ("a knows that φ") can be expressed in $\mathsf{ATL}_\mathsf{i}^\star$. The standard semantics of K_a is given by:

$$M, q \models K_a\varphi \quad \text{ iff } M, q' \models \varphi \text{ for every } q' \text{ such that } q \sim_a q'.$$

It is easy to see that $M, \lambda[0] \models K_a\varphi$ iff $M, \lambda \models_i \langle\langle a \rangle\rangle \varphi\, \mathcal{U}\, \varphi$. Equivalently, $K_a\varphi$ can be expressed in $\mathsf{ATL}_\mathsf{i}^\star$ as $\langle\langle a \rangle\rangle \bot\, \mathcal{U}\, \varphi$.

2.5. Other Semantics of Strategic Ability under Imperfect Information

The concept of *ability* has received considerable attention in philosophy, logic, and artificial intelligence over the last 70 years, including the seminal works of Ryle [Ryle 1949], McCarthy and Hayes [McCarthy and Hayes 1969], Chellas [Chellas 1969], Moore [Moore 1977; 1985], and Belnap and Perloff [Belnap and Perloff 1988]. The term corresponds fairly closely to its everyday usage: that is, ability means the capability to do things, and to bring about states of affairs. Originally, the focus was mainly on characterizing the abilities of human agents. Shortly afterwards, researchers in computer science and artificial intelligence got also interested in the notion of *what machines can achieve*. The 70 years of research can be briefly summarized as: *there is no single formal characterization of ability*; in fact, the term can have many different flavours depending on the context and the scenario to which it is applied.

Roughly speaking, the possible interpretations of A's ability to bring about property φ, formalized by formula $\langle\langle A \rangle\rangle \varphi$, can be grouped in four categories:

(1) There exists a specification of A's behavior σ_A (not necessarily executable) such that φ holds in every execution of s_A;
(2) There is an executable strategy s_A such that φ holds in every execution of s_A *(A have objective ability to enforce φ)*;
(3) A know that there is an executable strategy s_A such that φ holds in every execution of s_A, but they do not necessarily know the strategy itself *(A have a strategy "de dicto" to enforce φ).*
(4) There is an executable strategy s_A such that A know that φ holds in every execution of s_A *(A have a strategy "de re", or the subjective ability to enforce φ).*

Out of those, cases 2 and 4 correspond to the most prominent variants of ability in the literature. The latter (subjective ability) captures the capabilities of self-sufficient agents and teams, and is closest to the everyday use of the term, especially in relation to humans. The former (objective ability) characterizes an important aspect of abilities ascribed to machines and computer programs. We focus on $\mathsf{ATL}_\mathsf{i}^\star$ formalizations of the two notions, and refer the interested reader to [Ågotnes et al. 2015] for a more detailed discussion and further bibliography.

2.5.1. Variants of Subjective Ability. The semantics of $\mathsf{ATL}_\mathsf{i}^\star$, presented in the previous subsections, is based on two fundamental assumptions. First, a coalitional strategy consists of individual strategies, each of them to be executed by a single member of the coalition. Thus, executability of a strategy (encoded by the notion of *uniform strategy*) is based on what individual agents know in a given situation. Secondly, the coalition knows which combination of individual strategies it should select if there is one that succeeds from all the situations they consider possible. That is, given a collective strategy, we look at all the outcome paths starting from histories indistinguishable from the current one.

Unfortunately, there is no single interpretation of collective knowledge, and hence also collective indistinguishability for a group of agents A. The best known possibilities are: mutual knowledge E_A (also known as "everybody knows"), common knowledge C_A, and distributed knowledge D_A [Fagin et al. 1995]. The indistinguishability relations on states are lifted from individual epistemic relations $\sim_a$ as follows:

— $\sim_A^D = \bigcap_{a\in A} \sim_a$, i.e., distributed knowledge aggregates individual certainty within A,
— $\sim_A^E = \bigcup_{a\in A} \sim_a$, i.e., "everybody knows" aggregates uncertainty,[7]
— $\sim_A^C$ is the transitive closure of $\sim_A^E$; that is, common knowledge aggregates uncertainty, plus uncertainty about uncertainty, etc.

Additionally, we take $\sim_\emptyset^C$ and $\sim_\emptyset^D$ to coincide with the identity relation.

In the same way, one can define indistinguishability relations on histories. $\approx_A^D$ is the intersection of all $\approx_a$ relations for $a \in A$, $\approx_A^E$ is their union, and $\approx_A^C$ is the transitive closure of $\approx_A^E$. It is easy to see that $\approx_A = \approx_A^E$, i.e., the semantics of $\langle\!\langle A\rangle\!\rangle$ with imperfect information in Section 2.4 uses the "everybody knows" type of indistinguishability. This follows the "canonical" semantics given by Schobbens [Schobbens 2004], and fits the most frequent interpretation of strategic ability in the literature. Still, the choice of $\approx_A^E$ is somewhat arbitrary. One may ask what happens if we change the semantics of $\mathsf{ATL}_\mathsf{i}^\star$ so that the members of A look at outcome paths they perceive as possible in the sense of distributed or common knowledge. In particular, would it lead to different properties of the resulting logic? To study the question formally, we define the semantic relations $\models_{i_D}$ and $\models_{i_C}$ as follows. Let $\mathit{plays}^{i_D}(h, s_A) = \bigcup_{h\approx_A^D h'} \mathit{out}(h', s_A)$ and $\mathit{plays}^{i_C}(h, s_A) = \bigcup_{h\approx_A^C h'} \mathit{out}(h', s_A)$. Then, $\models_{i_D}$ is defined according to the template in Section 2.4, replacing the clause for $\langle\!\langle A\rangle\!\rangle\varphi$ in the following way:

$M, \lambda \models_{i_D} \langle\!\langle A\rangle\!\rangle\varphi$ iff there is a collective i-strategy s_A such that, for each $\lambda' \in \mathit{plays}_M^{i_D}(\lambda[0], s_A)$, we have $M, \lambda' \models_{i_D} \varphi$.

Relation $\models_{i_C}$ is constructed analogously.

2.5.2. Objective Ability under Uncertainty. Objective ability refers to existence of a strategy that is guaranteed to succeed from the perspective of an external observer with complete information about the current state of the model. That is, only the paths starting from the current global state are relevant, which can be formalized by $\mathit{plays}^{i_o}(h, s_A) = \mathit{plays}^I(h, s_A) = \mathit{out}(h, s_A)$. Similarly to the previous semantic variants, relation $\models_{i_o}$ is defined by replacing the clause for $\langle\!\langle A\rangle\!\rangle\varphi$ as follows:

$M, \lambda \models_{i_o} \langle\!\langle A\rangle\!\rangle\varphi$ iff there is a collective i-strategy s_A such that, for each $\lambda' \in \mathit{plays}_M^{i_o}(\lambda[0], s_A)$, we have $M, \lambda' \models_{i_o} \varphi$.

Throughout the paper, we will keep the semantics of ability from Section 2.4 as the baseline. We will also study to what extent small shifts in the semantics (along the lines discussed in this section) influence the formal properties of interest.

3. STRATEGIES WITH TRULY PERFECT RECALL

We have already seen that, in $\mathsf{ATL}^\star$, strategies are synthesised with respect to the current state of the game ($\lambda[0]$), and that "*previous events*" influence neither the strategy selection nor the resulting paths. In this section we illustrate how this leads to the forgetting phenomenon. We also introduce a "*no-forgetting*" semantics for $\mathsf{ATL}^\star$.

3.1. Agents with Standard Perfect Recall Forget

In the standard semantics of $\mathsf{ATL}^\star$ agents "forget" information about the past, even if they are assumed to have perfect recall. For instance, in formula $\langle\!\langle a\rangle\!\rangle\Diamond\langle\!\langle b\rangle\!\rangle\Box\mathsf{p}$, agent b must start collecting observations from scratch when executing her strategy to bring about $\Box\mathsf{p}$. The history of the game determined by the first strategic modality is not taken into account. This leads to counterintuitive effects, as the following example shows.

Example 3.1 (*Forgetting in perfect recall*). On one hand, $M_1, q_0 \models_i \langle\!\langle 2\rangle\!\rangle\Diamond\mathsf{win}$, that is, the guesser has a uniform strategy to win the shell game starting in q_0. On the other hand, $M_1, q_2 \models_i$

[7]Notice that $\sim_A^E$ is equal to the $\sim_A$ relation, introduced in Section 2.2.

$\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$. As the shuffler in q_0 can easily enforce the future state to be q_2, we obtain that $M_1, q_0 \models_i \langle\langle 1\rangle\rangle\Diamond\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$. Thus, in (M_1, q_0), the guesser has the ability to win no matter what the shuffler does, and *at the same time* the shuffler has a strategy to deprive the guesser of the ability.

3.2. ATL* with Truly Perfect Recall

To get rid of this "forgetting" behavior, we will use the *truly perfect recall semantics* of $\mathsf{ATL}^\star$, captured by relation $\models_x^{nf}$, where $x \in \{i, I\}$. Again, the index indicates whether we refer to the abilities based on perfect (I) or imperfect information strategies (i). Formulae are interpreted over triples consisting of a model, a path and an index $k \in \mathbb{N}_0$ which points to the current position on the infinite path. Intuitively, the subhistory of the path up to k encodes the past, and the subpath starting after k, the future. The crucial part of this semantics is that the agents always remember the sequence of the past events — and they can *learn* from those events.

Definition 3.2 (*Truly perfect recall semantics for* $\mathsf{ATL}^\star$). Let M be an iCGS, $\lambda \in \Lambda_M$ and $k \in \mathbb{N}_0$. The *truly perfect recall semantics* of $\mathsf{ATL}^\star$ $\models_x^{nf}$, parameterized with $x \in \{i, I\}$, is defined as follows:

$M, \lambda, k \models_x^{nf} p$ iff $\lambda[k] \in \pi(p)$ for $p \in \Pi$;
$M, \lambda, k \models_x^{nf} \neg\varphi$ iff $M, \lambda, k \not\models_x^{nf} \varphi$;
$M, \lambda, k \models_x^{nf} \varphi_1 \wedge \varphi_2$ iff $M, \lambda, k \models_x^{nf} \varphi_1$ and $M, \lambda, k \models_x^{nf} \varphi_2$;
$M, \lambda, k \models_x^{nf} \langle\langle A\rangle\rangle\varphi$ iff there exists a collective x-strategy s_A such that, for all $\lambda' \in$ *plays*$_M^x(\lambda[0, k], s_A)$, we have $M, \lambda', k \models_x^{nf} \varphi$;
$M, \lambda, k \models_x^{nf} \bigcirc\varphi$ iff $M, \lambda, k+1 \models_x^{nf} \varphi$
$M, \lambda, k \models_x^{nf} \varphi_1 \mathcal{U} \varphi_2$ iff there exists $i \geq k$ such that $M, \lambda, i \models_x^{nf} \varphi_2$ and $M, \lambda, j \models_x^{nf} \varphi_1$ for all $k \leq j < i$.

Remark 3.3. We note that the semantics of $\mathsf{ATL}^\star$ encoded by $\models_i^{nf}$ is based on the "everybody knows" type of indistinguishability. Similarly to Section 2.5, one can define analogous semantics for agents who consider outcome paths starting from the indistguishable states in the sense of common or distributed knowledge. Objective abilities of agents with truly perfect recall can be also of interest. Formally, relations $\models_x^{nf}$ for $x = i_D, i_C, i_o$ are defined by replacing the clause for $\langle\langle A\rangle\rangle\varphi$ in Definition 3.2 as follows:

$M, \lambda, k \models_x^{nf} \langle\langle A\rangle\rangle\varphi$ iff there exists a collective i-strategy s_A such that, for all $\lambda' \in$ *plays*$_M^x(\lambda[0, k], s_A)$, we have $M, \lambda', k \models_x^{nf} \varphi$;

The logical system with the syntax given by $\mathsf{ATL}^\star$ and the semantics by $\models_x^{nf}$ will be referred to as $\mathsf{ATL}^\star_{\mathsf{nf,x}}$. Given a state formula φ and a history h, we define $M, h \models_x^{nf} \varphi$ iff $M, \lambda, k \models_x^{nf} \varphi$ for all $\lambda \in \Lambda_M$ such that $\lambda[0, k] = h$. A state formula φ is *valid* in $\mathsf{ATL}^\star_{\mathsf{nf,x}}$ iff $M, q \models_x^{nf} \varphi$ for all models M and states q (note that states can be seen as a special kind of histories); and *satisfiable* if such a pair (M, q) exists.

The new semantics differs from the standard semantics of $\mathsf{ATL}^\star$ only in that it keeps track of the history by incorporating it into each path. Instead of building paths starting in the current state of the game ($\lambda[0]$ in the standard semantics), we look at paths λ that describe the play from the very beginning. $\lambda[0, k-1]$ represents the sequence of past states (excluding the current one), $\lambda[k]$ is the current state, and $\lambda[k+1, \infty]$ is the future part of the play. We illustrate the semantics by the following example.

Example 3.4 (*Shell game ctd.*). Consider the pointed iCGS (M_1, q_0) from Figure 2 again. Whatever the shuffler does in the first two steps, the guesser can adapt its action (in q_2 and q_2') to win the game. In particular, the i-strategy s_2 from Example 2.4 can be used to demonstrate that for all $\lambda \in$ *plays*$^i(q_0, s_2)$—for every strategy of 2—we have $M_1, \lambda, 0 \models_i^{nf} \Diamond\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$. As a consequence, $M_1, q_0 \models_i^{nf} \neg\langle\langle 1\rangle\rangle\Diamond\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$.

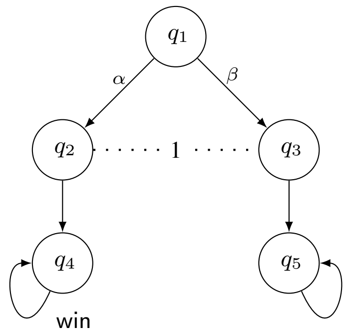


Fig. 3. The iCGS M_f with a single player (1) and two possible actions (α and β) in state q_1, which lead to q_2 and q_3, respectively. States q_2 and q_3 are indistinguishable. The player forgets her previous action if a subformula is evaluated in q_2 or q_3. This happens because, traditionally, paths record only states and not actions executed by the player.

Thus, the truly perfect recall semantics gives the intended result for M_1. Note that agents can still forget the *actions* that they have performed. Essentially, our notion of true perfect recall is rooted in the way in which temporal paths are typically defined: as sequences of states rather than sequences of interleaved states and action profiles.

Remark 3.5 (*Forgetting in the truly perfect recall semantics*). In Figure 3, a single-player iCGS M_f is shown. First, we observe that $M_f, q_1 \models_i^{nf} \langle\!\langle 1 \rangle\!\rangle \Diamond \mathsf{win}$. The player's strategy is s_1 where $s_1(q_1) = \alpha$. The actions assigned to all other histories by s_1 are unimportant. We note that $plays^i_{M_f}(q_1, s_1) = \{q_1 q_2 q_4^\omega\}$ contains a unique path.

However, it also holds that $M_f, q_1 \not\models_i^{nf} \langle\!\langle 1 \rangle\!\rangle \bigcirc \langle\!\langle 1 \rangle\!\rangle \bigcirc \mathsf{win}$. When synthesising a strategy s_1' for the subformula $\langle\!\langle 1 \rangle\!\rangle \bigcirc \mathsf{win}$, we have $plays^i_{M_f}(q_1 q_2, s_1') = out_{M_f}(q_1 q_2, s_1') \cup out_{M_f}(q_1 q_3, s_1')$, since $q_1 q_2 \approx_1 q_1 q_3$.

The player remembers the history, *but not the actions which have been played, including her own*. This is the case since, in $\mathsf{ATL}^\star$, histories are simply sequences of states and not sequences of interleaved action profiles and states. We consider this as a purely technical issue, as the last performed action profile can be encoded within a state whenever the need arises. Then, the modeler can define explicitly which agents can observe what actions.

3.3. Standard vs. True Perfect Recall: When Are They Different?

The definitions of $\models_x$ and $\models_x^{nf}$ *look* different, but that does not necessarily mean that they produce different evaluations of $\mathsf{ATL}^\star$ formulae. Here, we examine whether it is indeed the case.

3.3.1. Perfect Information. As the difference between $\mathsf{ATL}^\star_\mathsf{x}$ and $\mathsf{ATL}^\star_{\mathsf{nf,x}}$ lies in the "forgetting" of past observations when evaluating nested formulae, it comes as no real surprise that the two semantics coincide for perfect information. Agents with perfect information always precisely know the current global state of the system, and thus they cannot be uncertain about anything, including their own observations.

THEOREM 3.6. *For all iCGSs M, paths $\lambda \in \Lambda_M$, and $\mathsf{ATL}^\star$ formulae φ we have that $M, \lambda, 0 \models_I^{nf} \varphi$ iff $M, \lambda \models_I \varphi$.*

PROOF. Let $h \circ \lambda$ be an arbitrary path in Λ_M with $k = |h| - 1$ and $|h| \geq 1$. First, we observe that:

$$plays^I_M(h, s_A) = out_M(h, s_A) \qquad (1)$$

for arbitrary collective I-strategies s_A. Furthermore, for all collective I-strategies s_A and histories h, there is a collective I-strategy s'_A such that:

$$out_M(h, s'_A) = \{h \circ \lambda \mid last(h) \circ \lambda \in out_M(last(h), s_A)\} \tag{2}$$

and also, for all collective I-strategies s'_A and histories h there is a collective strategy s_A such that:

$$out_M(last(h), s'_A) = \{last(h) \circ \lambda \mid h \circ \lambda \in out_M(h, s_A)\} \tag{3}$$

Informally, in (2) s'_A makes the same decisions as s_A given that history h has already taken place, while in (3) s'_A makes the same decisions after $last(h) \circ \lambda$ as s_A would do after history $h \circ \lambda$.

Now, we prove the stronger statement: $M, h \circ \lambda, k \models_I^{nf} \varphi$ iff $M, last(h) \circ \lambda \models_I \varphi$, for all $h \circ \lambda \in \Lambda_M$ and all ATL*-formulae φ, such that $k = |h| - 1$, $|h| \geq 1$. The proof is done by induction over the formula structure of φ.

Base cases: The case for $\varphi = p$ is straightforward. $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ does not contain cooperation modalities. By (1-3) we have that the following statements are equivalent:

— $M, h \circ \lambda, k \models_I^{nf} \langle\langle A \rangle\rangle \gamma$
— there exists s_A such that for all $h \circ \lambda' \in plays_M^I(h, s_A)$ we have $M, h \circ \lambda', k \models_I^{nf} \gamma$
— there exists s_A such that for all $last(h) \circ \lambda' \in plays_M^I(last(h), s_A)$ we have $M, last(h) \circ \lambda' \models_I \gamma$
— $M, last(h) \circ \lambda \models_I \langle\langle A \rangle\rangle \gamma$.

Induction hypothesis: Let φ be a formula. Then, the statement is true for all strict (state) subformulae of φ.
Induction step: The cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ contains cooperation modalities.

Let $\langle\langle B_i \rangle\rangle \psi_i$ be an outermost ATL*-subformula in γ i.e. there is no other cooperation modality in γ which strictly contains $\langle\langle B_i \rangle\rangle \psi_i$. Let:

$$H_i = \{(h \circ \lambda, k) \mid M, h \circ \lambda, k \models_I^{nf} \langle\langle B_i \rangle\rangle \psi_i\}$$
$$L_i = \{\lambda \mid M, \lambda \models_I \langle\langle B_i \rangle\rangle \psi_i\}$$

We observe that the complement of H_i (resp. L_i) with respect to Λ_M^{fin} is precisely the set $\{(h \circ \lambda, k) \mid M, h \circ \lambda, k \models_I^{nf} \neg\langle\langle B_i \rangle\rangle \psi_i$ (resp. $\{\lambda \mid M, \lambda \models_I \neg\langle\langle B_i \rangle\rangle \psi_i\}$). By induction hypothesis, $H_i = \{(h \circ \lambda, k) \mid last(h) \circ \lambda \in L_i\}$ (4). Direction "$\Rightarrow$“: Let s_A be a witnessing strategy for $M, h \circ \lambda, k \models_I^{nf} \langle\langle A \rangle\rangle \gamma$, i.e. $\forall \lambda' \in out_M(h, s_A)$ we have $M, \lambda', k \models_I^{nf} \gamma$. By (4) and (2) it follows that $M, last(h) \circ \lambda \models_I \gamma$. Direction "$\Leftarrow$“ follows exactly the same argument. □

3.3.2. Imperfect Information. As we have seen, the logics ATL$_\mathsf{I}^\star$ and ATL$_{\mathsf{nf,I}}^\star$ for perfect information are equivalent. However, the two semantics differ in the imperfect information case. To see this, consider model M_1 and state q_0 from Example 3.1. Let $\varphi \equiv \langle\langle 1 \rangle\rangle \Diamond \neg \langle\langle 2 \rangle\rangle \Diamond \mathsf{win}$. In Examples 3.1 and 3.4 we have shown that $M_1, q_0 \models_i \varphi$ but $M_1, q_0 \not\models_i^{nf} \varphi$. As a consequence, we obtain the following.

PROPOSITION 3.7. *There is an iCGS M, a state q in M, and an ATL* formula φ such that $M, q \models_i \varphi$ and $M, q \not\models_i^{nf} \varphi$.*

3.3.3. Other Notions of Subjective Ability. Proposition 3.7 shows that the semantics presented in Section 3.2 assigns different truth values than the standard one recalled in Section 2.4. This is at least the case when the set of paths, considered possible from the point of view of coalition A, is based on the “everybody knows” relation $\approx_A^E$. What happens if some other notion of group indistinguishability is used instead? As it turns out, the pattern is still the same.

PROPOSITION 3.8. *There is an iCGS M, a state q in M, and an ATL* formula φ such that $M, q \models_{i_D} \varphi$ and $M, q \not\models_{i_D}^{nf} \varphi$. The same applies to $\models_{i_C}$ vs. $\models_{i_C}^{nf}$.*

PROOF. We observe that the semantics of $\mathsf{ATL}^\star_i$, $\mathsf{ATL}^\star_{i_D}$, and $\mathsf{ATL}^\star_{i_C}$ coincide for formulae that use only singleton coalitions. The same holds for the semantics of the single-agent fragments of $\mathsf{ATL}^\star_{i,\mathsf{nf}}$, $\mathsf{ATL}^\star_{i_D,\mathsf{nf}}$, and $\mathsf{ATL}^\star_{i_C,\mathsf{nf}}$. This is because the relations $\approx^E_{\{a\}}$, $\approx^C_{\{a\}}$, and $\approx^D_{\{a\}}$ are exactly the same for any $a \in \mathbb{A}\mathrm{gt}$. In consequence, the model and the formula in the proof of Proposition 3.7 can be used here as well. □

Thus, the standard and true perfect recall semantics produce different evaluations of formulae, regardless of the notion of group knowledge that defines the set of initial states, relevant for the coalition. One could expect the same pattern if we only look at the executions starting from the actual state, i.e., in the case of objective ability. Surprisingly, the intuition turns out wrong.

3.3.4. Objective Ability under Imperfect Information

THEOREM 3.9. *For all iCGSs M, paths $\lambda \in \Lambda_M$, and $\mathsf{ATL}^\star$ formulae φ we have that $M, \lambda, 0 \models^{nf}_{i_o} \varphi$ iff $M, \lambda \models_{i_o} \varphi$.*

PROOF. We use the *tree unfoldings for objective ability* from [Bulling and Jamroga 2014, Section 4.2]. Briefly, $T_o(M, q)$ is the iCGS where the states are given by the finite histories in $\Lambda^{fin}_M(q)$, and the indistinguishability relations are defined by $\approx_a$. It was proved in [Bulling and Jamroga 2014, Theorem 2] that $T_o(M, q), q$ always satisfies the same formulae of $\mathsf{ATL}^\star_i$ as M, q. Note also that in $T_o(M, q)$ the notions of standard and true perfect recall coincide. Thus, we get the following chain of equivalences (remember that φ is a state formula!):

$$M, \lambda \models_{i_o} \varphi \text{ iff } T_o(M, \lambda[0]), \lambda \models_{i_o} \varphi \text{ iff } T_o(M, \lambda[0]), \lambda, 0 \models^{nf}_{i_o} \varphi \text{ iff } M, \lambda, 0 \models^{nf}_{i_o} \varphi.$$

The last equivalence follows easily from the construction of the semantic relation $\models^{nf}_{i_o}$. □

The result is quite surprising. The change from standard perfect recall to true perfect recall consists in two factors. First, strategies for nested strategic modalities may be evaluated in a tighter set of execution paths. Secondly, the notion of uniformity changes, because it is based on a tighter indistinguishability relation. Note that the former does not apply in case of objective ability, but the latter is still in place. Theorem 3.9 shows that the language of $\mathsf{ATL}^\star$ is not expressive enough to discern between uniform strategies based on standard vs. true perfect recall.

3.3.5. Summary. Thus, perfect recall and truly perfect recall coincide for agents with perfect information and for objective abilities under imperfect information. On the other hand, they are different for subjective abilities of agents with imperfect information, regardless of the type of epistemic neighborhood that is used in the semantics to provide the initial states for execution paths. How big is the impact? At the first glance, the change is not necessarily substantial. We will address the question formally in the next sections, and show that assuming "no forgetting" in interpretation of nested modalities changes the class of properties definable by formulae of $\mathsf{ATL}^\star$, as well as the set of valid sentences of the logic.

4. TRULY PERFECT RECALL: EXPRESSIVITY

We now proceed to show that the seemingly small change in semantics has important consequences for the resulting logics. We prove that the forgetting and truly perfect recall variants of $\mathsf{ATL}^\star$ differ in the properties they allow to express. We will look at which properties of iCGSs can be expressed in $\mathsf{ATL}^\star_{\mathsf{x}}$ and $\mathsf{ATL}^\star_{\mathsf{nf},\mathsf{x}}$, respectively (where $x \in \{i, I\}$). To do this, we briefly recall the notions of distinguishing power and expressive power (for more details, see e.g. [Clarke and Schlingloff 2001, Chapter 21]).

Definition 4.1 (*Distinguishing and expressive power*). Consider two logical systems $\mathsf{L}_1 = (\mathcal{L}_1, \models_1)$ and $\mathsf{L}_2 = (\mathcal{L}_2, \models_2)$ over the same class of models $\mathcal{M}$ (in our case, the class of iCGSs). By $[\![\varphi]\!]_\models = \{(M, q) \mid M, q \models \varphi\}$, we denote the class of pointed models that satisfy φ according to $\models$. Likewise, $[\![\varphi, M]\!]_\models =$

$\{q \mid M, q \models \varphi\}$ is the set of states (or, equivalently, pointed models) that satisfy φ in a given structure M.

We say that L_2 is *at least as expressive as* L_1 (written: $\mathsf{L}_1 \preceq_e \mathsf{L}_2$) iff for every formula $\varphi_1 \in \mathcal{L}_1$ there exists $\varphi_2 \in \mathcal{L}_2$ such that $[\![\varphi_1]\!]_{\models_1} = [\![\varphi_2]\!]_{\models_2}$. Moreover, L_2 is *at least as distinguishing as* L_1 (written: $\mathsf{L}_1 \preceq_d \mathsf{L}_2$) iff for every model M and formula $\varphi_1 \in \mathcal{L}_1$ there exists $\varphi_2 \in \mathcal{L}_2$ such that $[\![\varphi_1, M]\!]_{\models_1} = [\![\varphi_2, M]\!]_{\models_2}$.[8] L_1 and L_2 are equally expressive (resp. equally distinguishing) iff $\mathsf{L}_2 \preceq_x \mathsf{L}_1$ and $\mathsf{L}_1 \preceq_x \mathsf{L}_2$ where $x = e$ (resp. $x = d$). Finally, we say that L_2 is *strictly more distinguishing than* L_1 (written: $\mathsf{L}_1 \prec_d \mathsf{L}_2$) iff L_2 is at least as distinguishing, but not equally distinguishing to L_1. The definition of "strictly more expressive" is analogous.

Thus, expressive power refers to the general definability of properties by formulae of a given logical system. In contrast, distinguishing power captures the ability to discern between particular models. Note that $\mathsf{L}_1 \preceq_e \mathsf{L}_2$ implies $\mathsf{L}_1 \preceq_d \mathsf{L}_2$ but the converse is not true. For example, it is known that CTL has the same distinguishing power as CTL*, but strictly less expressive power [Clarke and Schlingloff 2001].

4.1. Comparing Expressivity for Perfect Information

Below is an immediate consequence of Theorem 3.6, again highlighting that both semantics coincide for agents with perfect information.

THEOREM 4.2. $\mathsf{ATL}^\star_\mathsf{I}$ *and* $\mathsf{ATL}^\star_{\mathsf{nf,I}}$ *are equally expressive and have the same distinguishing power.*

4.2. Imperfect Information

In what follows, we compare the expressiveness of the truly perfect recall variant of $\mathsf{ATL}^\star_{\mathsf{nf,i}}$ with that of its "forgetting" counterpart $\mathsf{ATL}^\star_\mathsf{i}$.

4.2.1. Truly Perfect Recall Does Not Embed Standard Perfect Recall

Example 4.3. Consider the models in Figure 4. We have that $M_2, a_0 \models_i^{nf} \langle\!\langle 1 \rangle\!\rangle \bigcirc \langle\!\langle 2 \rangle\!\rangle \bigcirc \mathsf{win}$ but $M_2', a_0 \not\models_i^{nf} \langle\!\langle 1 \rangle\!\rangle \bigcirc \langle\!\langle 2 \rangle\!\rangle \bigcirc \mathsf{win}$. In model M_2, player 2 can learn the state of the game after the first move (1 plays α in a_0); this is not the case in M_2'. Under the truly perfect recall semantics the two models are distinguishable, however, the models cannot be distinguished in $\mathsf{ATL}^\star_\mathsf{i}$.

To better understand the construction of M_2 and M_2', let us start with the models $\overline{M_2'}^{\downarrow}$ (Figure 4) and $\overline{M_2}^{\downarrow}$. The latter is simply $\overline{M_2'}^{\downarrow}$ where $a_0 \not\sim_2 b_0$. We note that $(\overline{M_2}^{\downarrow}, a_0)$ and $(\overline{M_2'}^{\downarrow}, a_0)$ can be distinguished in $\mathsf{ATL}^\star_\mathsf{i}$: we have $\overline{M_2}^{\downarrow}, a_0 \models_i \langle\!\langle 1,2 \rangle\!\rangle \Diamond \mathsf{win}$ but $(\star)$ $\overline{M_2'}^{\downarrow}, a_0 \not\models_i \langle\!\langle 1,2 \rangle\!\rangle \Diamond \mathsf{win}$ — in $\overline{M_2'}$, coalition $\{1,2\}$ has precisely the same (imperfect) information as player 2, and cannot distinguish a_0a_1 from b_0b_1.

We construct $M_2^{\downarrow}$ and $M'^{\downarrow}_2$ by adding transitions (ϵ, μ) and (μ, μ) to both $\overline{M_2}^{\downarrow}$ and $\overline{M_2'}^{\downarrow}$. $M'^{\downarrow}_2$ is shown in Figure 4. Now, $(\star)$ no longer holds: player 2 may play μ in a_0a_1 as well as in b_0b_1. Thus, player 1's action (μ or ϵ) leads to the winning state.

But even in this setup, $(M_2^{\downarrow}, a_0)$ and $(M_2'^{\downarrow}, a_0)$ may still be distinguished in $\mathsf{ATL}^\star_\mathsf{i}$, by e.g. $\langle\!\langle 2 \rangle\!\rangle \Box \neg \mathsf{win}$: in $M_2^{\downarrow}$ player 2 can ensure that the winning state is never reached, by playing α in a_0a_1, which is not true in $M_2'^{\downarrow}$, since $a_0a_1 \sim_2 b_0b_1$. Even if the second player does not determine the next-state from a_0, $a_0 \not\sim_2 b_0$ means that he can prevent winning, no matter what 1 does.

To solve this issue, we need to add more options for player 1 in a_0. In M_2, $\langle\!\langle 2 \rangle\!\rangle \Box \neg \mathsf{win}$ does not hold in M_2 (nor in M_2'). For instance, if player 1 plays μ, player 2's former strategy no longer prevents winning.

[8] Equivalently: for every pair of pointed models that can be distinguished by some $\varphi_1 \in \mathcal{L}_1$ there exists $\varphi_2 \in \mathcal{L}_2$ that distinguishes these models.

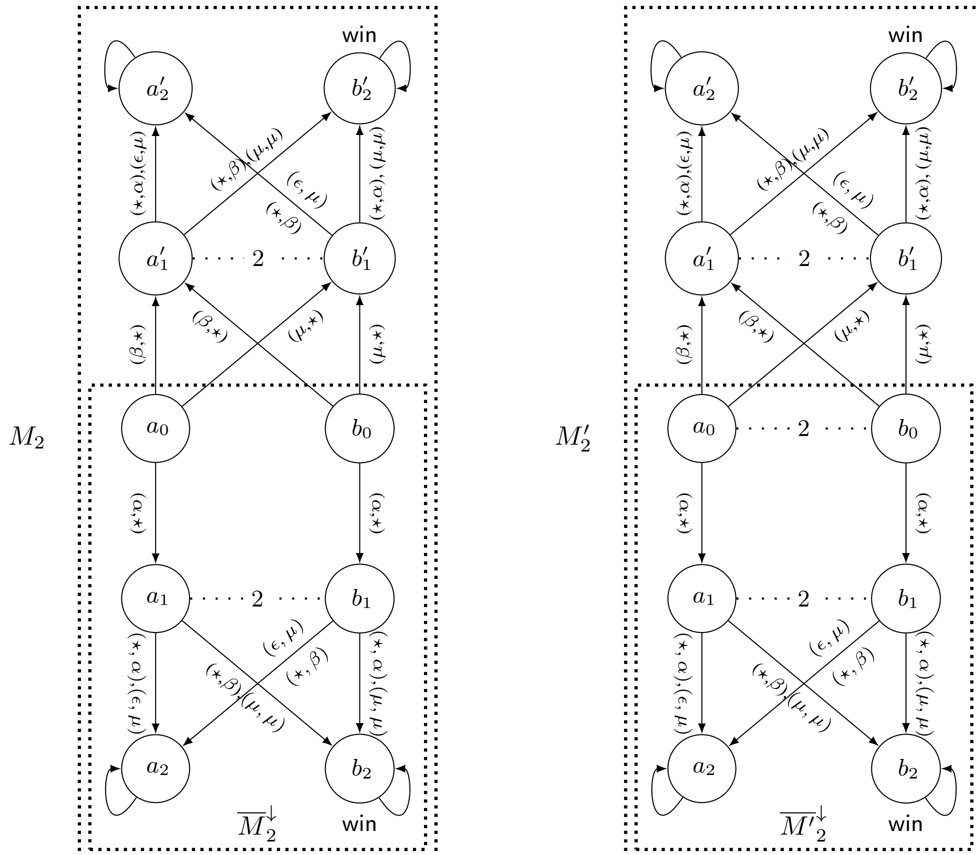


Fig. 4. Models M_2, M'_2 and their submodels $\overline{M}_2^{\downarrow}$, $\overline{M'}_2^{\downarrow}$. All the models include two players 1 and 2. Action tuples (α_1, α_2) give the action of player 1 (α_1) and of player 2 (α_2). The only difference between M_2 and M'_2 is that in model M_2 player 2 can also not distinguish a_0 and b_0.

The following proposition is based on the fact that no $\mathsf{ATL}_\mathsf{i}^\star$ formula distinguishes M_2 and M'_2 from Example 4.3. The complete proof is given in Appendix A.1.

PROPOSITION 4.4. *There are pointed iCGSs which satisfy the same* $\mathsf{ATL}_\mathsf{i}^\star$*-formulae, but can be distinguished in* $\mathsf{ATL}_{\mathsf{nf,i}}^\star$. *Thus,* $\mathsf{ATL}_{\mathsf{nf,i}}^\star \not\preceq_d \mathsf{ATL}_\mathsf{i}^\star$*, and hence also* $\mathsf{ATL}_{\mathsf{nf,i}}^\star \not\preceq_e \mathsf{ATL}_\mathsf{i}^\star$.

4.2.2. Standard Perfect Recall Does Not Embed True Perfect Recall. Next, we investigate whether $\mathsf{ATL}_{\mathsf{nf,i}}^\star$ is at least as distinguishing as $\mathsf{ATL}_\mathsf{i}^\star$.

Example 4.5. Let us consider the models iCGSs M_3 and M'_3 shown in Figure 5. There is an $\mathsf{ATL}_\mathsf{i}^\star$-formula that can distinguish both models: $M_3, a_0 \models_i \langle\!\langle 1 \rangle\!\rangle \bigcirc \langle\!\langle 1 \rangle\!\rangle \bigcirc \mathsf{win}$ and $M'_3, a_0 \not\models_i \langle\!\langle 1 \rangle\!\rangle \bigcirc \langle\!\langle 1 \rangle\!\rangle \bigcirc \mathsf{win}$. In the latter case player 1 "forgets" after the first step that the game has started in state a_0, and cannot distinguish states a_1 from a_2 when evaluating the nested formula. It is easy to see that there is no uniform strategy that wins from both a_1, a_2.

This leads to the following result:

PROPOSITION 4.6. *There are pointed iCGSs which satisfy the same* $\mathsf{ATL}_{\mathsf{nf,i}}^\star$ *formulae, but can be distinguished in* $\mathsf{ATL}_\mathsf{i}^\star$. *Thus,* $\mathsf{ATL}_\mathsf{i}^\star \not\preceq_d \mathsf{ATL}_{\mathsf{nf,i}}^\star$*, and hence also* $\mathsf{ATL}_\mathsf{i}^\star \not\preceq_e \mathsf{ATL}_{\mathsf{nf,i}}^\star$.

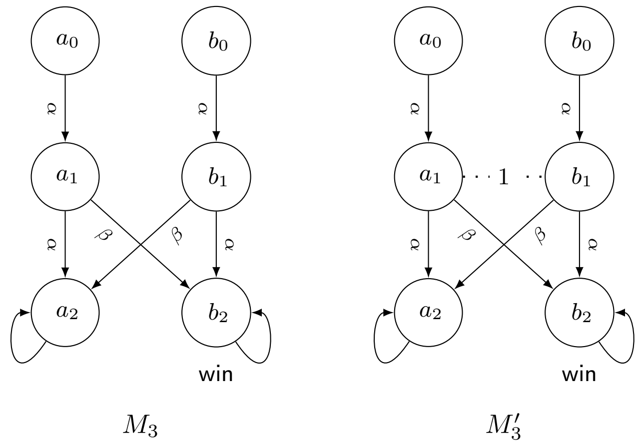


Fig. 5. Models M_3 and M'_3

PROOF. We consider models M_3 and M'_3 from Figure 5. We prove $M_3, h \models_i^{nf} \varphi$ iff $M'_3, h \models_i^{nf} \varphi$ for all $h \in \Lambda^{fin}_{M_3}(a_0)$ and all $\varphi \in \mathsf{ATL}^\star$, by induction over the formula structure of φ.

Base cases. Case $\varphi = p$ is straightforward. Case $\varphi = \langle\!\langle A \rangle\!\rangle \gamma$ where γ does not contain cooperation modalities. It is sufficient to observe that: ($\star$) $plays^i_{M_3}(h, s_A) = plays^i_{M'_3}(h, s_A)$ for all collective strategies s_A and histories $h \in \Lambda^{fin}_{M_3}(a_0)$.

Induction step. Cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\!\langle A \rangle\!\rangle \gamma$ where γ contains cooperation modalities: Let $\langle\!\langle B_i \rangle\!\rangle \psi_i$ (with $i = 1 \dots k$) be an outermost $\mathsf{ATL}^\star$-subformula in γ. By induction hypothesis, we have: $M_3, h \models_i^{nf} \langle\!\langle B_i \rangle\!\rangle \psi_i$ iff $M'_3, h \models_i^{nf} \langle\!\langle B_i \rangle\!\rangle \psi_i$ — the same histories h satisfy $\langle\!\langle B_i \rangle\!\rangle \psi_i$ in both models. It follows (by ($\star$)) that s_A is a witnessing strategy for $M_3, h \models_i^{nf} \langle\!\langle A \rangle\!\rangle \gamma$ iff s_A is a witnessing strategy for $M'_3, h \models_i^{nf} \langle\!\langle A \rangle\!\rangle \gamma$.

To conclude the proof, we note that both models can be distinguished in $\mathsf{ATL}^\star_\mathsf{i}$ as shown in Example 4.5. □

4.2.3. Final Result. As an immediate consequence, we obtain the theorem below.

THEOREM 4.7. *The logics* $\mathsf{ATL}^\star_\mathsf{i}$ *and* $\mathsf{ATL}^\star_{\mathsf{nf},\mathsf{i}}$ *have incomparable distinguishing and expressive powers.*

4.3. Other Notions of Subjective and Objective Ability

For the other semantics of subjective ability, we observe that all the models used in Section 4.2 include only uncertainty for agent 2. In consequence, the collective indistinguishability for coalitions coincides for mutual and common knowledge. Formally: $\sim^E_{\{1\}} = \sim^C_{\{1\}} = \sim_1$, $\sim^E_{\{2\}} = \sim^C_{\{2\}} = \sim_2$, and $\sim^E_{\{1,2\}} = \sim^C_{\{1,2\}} = \sim_2$; the same applies to indistinguishability of histories. Thus, whenever two of the models satisfy the same formulae of $\mathsf{ATL}^\star_i$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i}$), they must also satisfy the same formulae of $\mathsf{ATL}^\star_{i_C}$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i_C}$). Also, whenever two of the models are discerned by a formula of $\mathsf{ATL}^\star_i$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i}$), they must also be discerned by a formula of $\mathsf{ATL}^\star_{i_C}$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i_C}$). In consequence, the same models and formulae can be used to demonstrate the following.

THEOREM 4.8. *The logics* $\mathsf{ATL}^\star_{i_C}$ *and* $\mathsf{ATL}^\star_{\mathsf{nf},i_C}$ *have incomparable distinguishing and expressive powers.*

For the i_D-semantics, the argument is similar. The interpretation of strategic modalities for a singleton coalitions is the same in the i_D-semantics and in the i-semantics. Moreover, the relation

$\sim^D_{\{1,2\}}$ is the identity relation, i.e., it is equivalent to perfect information. Thus, the i_D-semantics of $\langle\!\langle 1,2\rangle\!\rangle\varphi$ in the models used in Section 4.2 is equivalent to its perfect information semantics, and hence expressible equivalently by $\neg\langle\!\langle\emptyset\rangle\!\rangle\neg\varphi$ in the i-semantics (cf. Remark 2.5). This finally implies that, whenever two of the models satisfy the same formulae of $\mathsf{ATL}^\star_i$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i}$), they must also satisfy the same formulae of $\mathsf{ATL}^\star_{i_D}$ (resp. $\mathsf{ATL}^\star_{\mathsf{nf},i_D}$).

Furthermore, the discerning formulae in the proofs of Section 4.2 use only abilities of singleton coalitions. Thus, whenever one of the formulae discern two models in the i-semantics (resp. nf, i-semantics), it must also discern the models according to the i_D-semantics (resp. nf, i_D-semantics). In consequence, the models and formulae from Section 4.2 can be used to demonstrate the following.

THEOREM 4.9. *The logics* $\mathsf{ATL}^\star_{i_D}$ *and* $\mathsf{ATL}^\star_{\mathsf{nf},i_D}$ *have incomparable distinguishing and expressive powers.*

On the other hand, since the perfect recall and truly perfect recall semantics coincide for objective ability under imperfect information (Theorem 3.9), we get the following as immediate corollary:

THEOREM 4.10. $\mathsf{ATL}^\star_{i_o}$ *and* $\mathsf{ATL}^\star_{\mathsf{nf},i_o}$ *are equally expressive and have the same distinguishing power.*

5. VALIDITIES

Another way of comparing two logical systems is to compare their sets of valid sentences, that is, the general properties that hold in every model according to the given semantics.

Intuitively, each formula can be interpreted as a game property. Such properties describe the abilities of agents and their groups, that possibly hold for some games, and do not hold in the others. While expressiveness concerns which game properties are definable in the language of the logic, validities are properties that universally hold. Thus, by comparing validity sets of different logical systems, we can compare the general properties of games induced by the underlying semantics (cf. [Bulling and Jamroga 2014]).

Given a semantics *sem*, we use $Val(\mathsf{ATL}^\star_{sem})$ to denote the set of valid sentences of $\mathsf{ATL}^\star_{sem}$, and $Sat(\mathsf{ATL}^\star_{sem})$ to denote the set of satisfiable sentences of $\mathsf{ATL}^\star_{sem}$. Intuitively, each formula $\varphi \in Val(\mathsf{ATL}^\star_{sem})$ describes an *invariant property* or *game rule* of $\mathsf{ATL}^\star_{sem}$. In this section, we investigate the relationship between $Val(\mathsf{ATL}^\star_x)$ and $Val(\mathsf{ATL}^\star_{x,nf})$ for $x \in \{I, i\}$. In particular, a result of the form $Val(\mathsf{ATL}^\star_{sem_1}) \subsetneq Val(\mathsf{ATL}^\star_{sem_2})$ means that the game rules of $\mathsf{ATL}^\star_{sem_2}$ are a strict *specialization* of those of $\mathsf{ATL}^\star_{sem_1}$.

Finally, we recall that $Sat(\mathsf{ATL}^\star_{sem_2}) \subsetneq Sat(\mathsf{ATL}^\star_{sem_1})$ iff $Val(\mathsf{ATL}^\star_{sem_1}) \subsetneq Val(\mathsf{ATL}^\star_{sem_2})$. Thus, any result comparing the validity sets of $\mathsf{ATL}^\star_{sem_1}$ and $\mathsf{ATL}^\star_{sem_2}$ immediately implies the dual characterization of satisfiable sentences.

5.1. Perfect Information and Objective Ability under Imperfect Information

The following result is a direct corollary of Theorem 3.6.

THEOREM 5.1. $Val(\mathsf{ATL}^\star_{\mathsf{I}}) = Val(\mathsf{ATL}^\star_{\mathsf{nf,I}})$.

Similarly, the following is an immediate consequence of Theorem 3.9:

THEOREM 5.2. $Val(\mathsf{ATL}^\star_{i_o}) = Val(\mathsf{ATL}^\star_{\mathsf{nf},i_o})$.

5.2. Imperfect Information

Now we will compare the validity sets of $\mathsf{ATL}^\star_{\mathsf{nf,i}}$ and $\mathsf{ATL}^\star_{\mathsf{i}}$.

5.2.1. Epistemic Tree Unfoldings. First, we introduce a class $\mathcal{T}$ of iCGSs for which the semantics of $\mathsf{ATL}^\star_{\mathsf{i}}$ and $\mathsf{ATL}^\star_{\mathsf{nf,i}}$ coincide. Members of $\mathcal{T}$ are *infinite epistemic trees* obtained by applying an *unfolding procedure* on arbitrary models M, with respect to a given *initial state* q. For the moment, suppose that all states in M can be distinguished from q. Then, the model M can be unfolded from

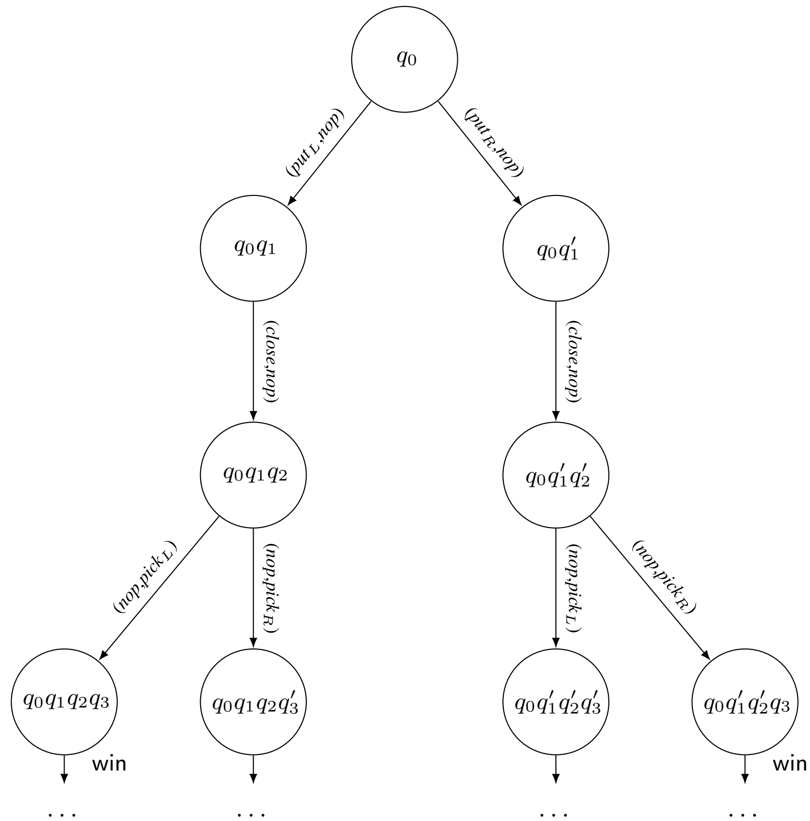


Fig. 6. The no-forgetting epistemic tree unfolding of the iCGS M_1 from Figure 2

q to an infinite tree $T(M,q)$ whose states correspond to histories in M. Two nodes h and h' in a tree are linked by an epistemic relation belonging to an agent a, written as $h \sim_a h'$ if, and only if, $h \approx_a h'$ in M.

Now, it might be the case that there are states indistinguishable from q. They have to be considered as well. Let $Q' = \{q' \in M \mid q \sim_{\mathbb{A}\mathrm{gt}} q'\}$ be the set of all states indistinguishable from q for some agent from $\mathbb{A}\mathrm{gt}$. For each state $\hat{q} \in Q'$ we construct the unfolding $T(M,\hat{q})$ as described above. Moreover, we introduce epistemic links *between* these trees. For each two histories h and h' in each of these trees we define $h \sim_a h'$ if, and only if, $h \approx_a h'$ in M. The resulting model, i.e., the collection of all those trees plus the inter-tree epistemic relations, is denoted by $T^{nf}(M,q)$. We note in passing that the unfolding was already considered in [Bulling and Jamroga 2014, Example 7] and shown to be insufficient for $\mathsf{ATL}^\star_\mathsf{i}$ with the standard semantics. With respect to the truly-perfect recall semantics, however, the unfolding does the right thing.

Definition 5.3 (*No-forgetting epistemic tree unfolding*). Consider an iCGS $M = (\mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o, \{\sim_a \mid a \in \mathbb{A}\mathrm{gt}\})$ and a state q of M. We construct the iCGS $T(M,q)$ as follows:

$$T(M,q) = \langle \mathbb{A}\mathrm{gt}^{T,q}, St^{T,q}, \Pi^{T,q}, \pi^{T,q}, Act^{T,q}, d^{T,q}, o^{T,q}, \{\sim_a^{T,q} \mid a \in \mathbb{A}\mathrm{gt}^{T,q}\}\rangle$$

where:

— $\mathbb{A}\mathrm{gt}^{T,q} = \mathbb{A}\mathrm{gt}$, $\Pi^{T,q} = \Pi$, $Act^{T,q} = Act$, $St^{T,q} = \{h \mid h \in \Lambda_M^{fin}\}$: the sets of agents, propositions and actions of $T(M,q)$ coincide with those of M. Each state in $T(M,q)$ is a history of M which starts in state q.
— $\pi^{T,q}(p) = \{h \mid \mathit{last}(h) \in \pi(p)\}$, $\forall p \in \Pi$; each state h from $T(M,q)$ is labelled with the same propositions as $\mathit{last}(h)$ in M;
— $d_a^{T,q}(h) = d_a(\mathit{last}(h))$, $\forall h \in St^{T,q}$; the set of allowed actions in each state h of $T(M,q)$ is the same as that in each state $\mathit{last}(h)$ of M;
— $q = o(\mathit{last}(h), \alpha_1, \ldots, \alpha_k) \iff o^{T,q}(h, \alpha_1, \ldots, \alpha_k) = h \circ q, \forall h \in St^{T,q}$; each transition from h to $h \circ q'$ in $T(M,q)$ corresponds to a transition from $\mathit{last}(h)$ to q' in M;
— $h \sim_a^{T,q} h'$ iff $h \approx_a h'$ in M, $\forall h, h' \in St^{T,q}$; two states h and h' are indistinguishable for agent a in $T(M,q)$ if the histories h and h' are indistinguishable for a in M.

Let $Q' = \{q' \mid q \sim_{\mathbb{A}\mathrm{gt}} q'\}$. The *no-forgetting epistemic tree unfolding* of M, denoted $T^{nf}(M,q)$, is the iCGS:

$$T^{nf}(M,q) = \langle \mathbb{A}\mathrm{gt}^{T^{nf},q}, St^{T^{nf},q}, \Pi^{T^{nf},q}, \pi^{T^{nf},q}, Act^{T^{nf},q}, d^{T^{nf},q}, o^{T^{nf},q}, \{\sim_a^{T^{nf},q} \mid a \in \mathbb{A}\mathrm{gt}^{T^{nf},q}\}\rangle$$

obtained as follows:

— $\mathbb{A}\mathrm{gt}^{T^{nf},q} = \mathbb{A}\mathrm{gt}^{T,q}$, $\Pi^{T^{nf},q} = \Pi^{T,q}$, $Act^{T^{nf},q} = Act^{T,q}$, $St^{T^{nf},q} = \bigcup_{q \in Q'} St^{T,q}$
— $\pi^{T^{nf}}(p) = \bigcup_{q \in Q'} \pi^{T,q}(p)$, $\forall p \in \Pi$;
— $d_a^{T^{nf},q}(h) = d_a^{T,q}(h)$, $\forall h \in St^{T^{nf},q}$;
— $h \sim_a^{T^{nf},q} h'$ iff $h \approx_a h'$, $\forall h, h' \in St^{T^{nf},q}$;

Example 5.4 (*No-forgetting epistemic tree unfolding*). In Figure 6, we show the no-forgetting epistemic tree unfolding $T^{nf}(M_1, q_0)$ of model M_1 from Figure 2. Note that there are no epistemic links between different states of $T^{nf}(M_1, q_0)$, since all corresponding histories are distinguishable by the second agent in M_1. Moreover, in this particular case: $T^{nf}(M_1, q_0) = T(M_1, q_0)$.

Figure 7 illustrates the submodel $M_2^{\uparrow}$ of M_2 from Example 4 (on the left), as well as $T^{nf}(M_2^{\uparrow}, a_0)$ (right). The unfolding is obtained by first constructing the components $T(M_2^{\uparrow}, a_0)$ and $T(M_2^{\uparrow}, b_0)$, and then adding epistemic links between all indistinguishable states of the latter two models.

5.2.2. True Perfect Recall Inherits Validities from Standard Perfect Recall. The following proposition first establishes that the unfolding is truth-preserving in the truly perfect recall semantics. Hence, each sentence which is true with respect to a given state q and model M, is also true with respect to the unfolding of M from q. At the same time, the truly perfect recall and standard $\mathsf{ATL}^{\star}_{\mathsf{i}}$-semantics coincide over no-forgetting epistemic tree unfoldings.

PROPOSITION 5.5. *$M, q \models_i^{nf} \varphi$ iff $T^{nf}(M,q), q \models_i^{nf} \varphi$ iff $T^{nf}(M,q), q \models_i \varphi$, for all $\mathsf{ATL}^{\star}$-formulae, iCGSs M and states q.*

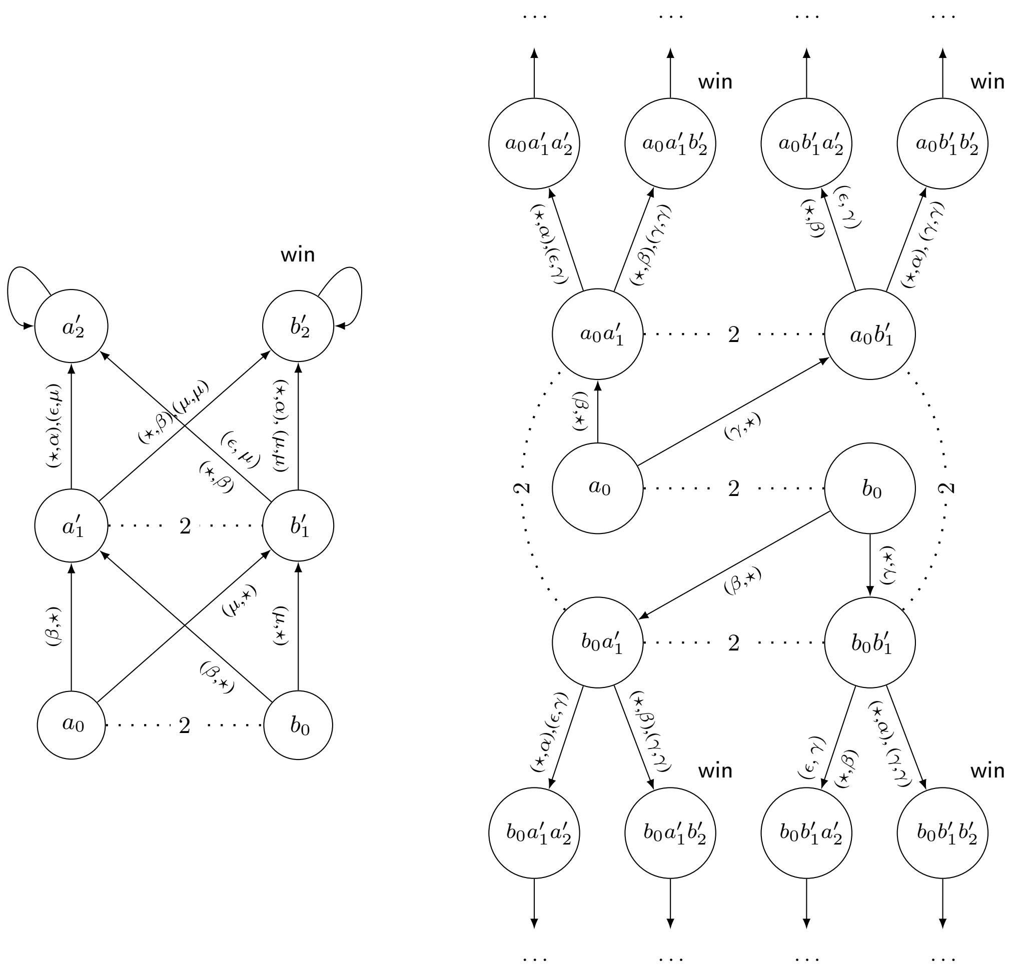
The proof is presented in Appendix A.2. This result is key for obtaining the following:

PROPOSITION 5.6. $\mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{i}}) \subseteq \mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{nf,i}})$.

PROOF. We prove that $Sat(\mathsf{ATL}^{\star}_{\mathsf{nf,i}}) \subseteq Sat(\mathsf{ATL}^{\star}_{\mathsf{i}})$. Let $\varphi \in Sat(\mathsf{ATL}^{\star}_{\mathsf{nf,i}})$. Thus, there exist M, q such that $M, q \models_i^{nf} \varphi$. By Proposition 5.5, $T^{nf}(M,q) \models_i \varphi$. Hence $\varphi \in Sat(\mathsf{ATL}^{\star}_{\mathsf{i}})$. □

5.2.3. The Converse Does Not Hold. We now show that there exists a sentence which is valid in $\mathsf{ATL}^{\star}_{\mathsf{nf,i}}$, but not in $\mathsf{ATL}^{\star}_{\mathsf{i}}$. Informally, such a sentence expresses that whenever an agent a has the ability to maintain p, then the agent preserves this ability in some next-state of the game.

PROPOSITION 5.7. $\mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{nf},i}) \not\subseteq \mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{i}})$.



$M_2^{\uparrow}$ $\qquad$ $T^{nf}(M_2^{\uparrow}, a_0)$

Fig. 7. The submodel $M_2'^{\uparrow}$ of M_2' (Example 4), is shown on the left. On the right, we have the epistemic tree unfolding $T^{nf}(M_2'^{\uparrow}, a_0)$. The epistemic links between states $a_0a'_1$, $b_0b'_1$ and $b_0a'_1$, $a_0b'_1$ have been omitted.

PROOF. We consider the formula $\varphi \equiv \langle\langle a\rangle\rangle\Box \mathsf{p} \rightarrow \mathsf{E}\bigcirc \langle\langle a\rangle\rangle\Box\mathsf{p}$ where $\mathsf{E}\bigcirc\psi \equiv \neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\psi$. Informally, $\mathsf{E}\bigcirc\psi$ says that there is a path on which ψ holds in the next moment. Thus, φ expresses that, if agent a has the ability to maintain p forever, then he retains the ability in at least one successor of the current state. We will show that φ is valid in $\mathsf{ATL}^{\star}_{\mathsf{nf},i}$, but not in $\mathsf{ATL}^{\star}_{\mathsf{i}}$.

For the former, suppose that $M, q \models^{nf}_i \langle\langle a\rangle\rangle\Box\mathsf{p}$, let s_a be a witnessing strategy and $\lambda = qq_1q_2\ldots \in \mathit{plays}^i_M(q, s_a)$. Since $\mathit{plays}^i_M(qq_1, s_a) \subseteq \mathit{plays}^i_M(q, s_a)$ we have $M, \lambda, 1 \models^{nf}_i \langle\langle a\rangle\rangle\Box\mathsf{p}$ and also $M, \lambda, 0 \models^{nf}_i \bigcirc\langle\langle a\rangle\rangle\Box\mathsf{p}$. Thus, $M, \lambda, 0 \models^{nf}_i \mathsf{E}\bigcirc\langle\langle a\rangle\rangle\Box\mathsf{p}$,

For the latter, we use model M'_3 in Figure 5, extended with proposition p which holds everywhere except for b_2. That is, $M'_3, q \models \mathsf{p}$ iff $M'_3, q \not\models \mathsf{win}$. Clearly, $M'_3, a_0 \models_i \langle\langle 1\rangle\rangle\Box\mathsf{p}$. On the other hand, $M'_3, a_0 \not\models_i \mathsf{E}\bigcirc\langle\langle 1\rangle\rangle\Box\mathsf{p}$. In consequence, φ is not valid in $\mathsf{ATL}^{\star}_{\mathsf{i}}$, which concludes the proof. $\Box$

5.2.4. Final Result. It is immediate from Propositions 5.6 and 5.7 that $\mathsf{ATL}^\star_{\mathsf{nf,i}}$ describes a more specific class of games than $\mathsf{ATL}^\star_{\mathsf{i}}$: games in which players do not forget past events:

THEOREM 5.8. $Val(\mathsf{ATL}^\star_{\mathsf{i}}) \subsetneq Val(\mathsf{ATL}^\star_{\mathsf{nf,i}})$. *Consequently, also* $Sat(\mathsf{ATL}^\star_{\mathsf{nf,i}}) \subsetneq Sat(\mathsf{ATL}^\star_{\mathsf{i}})$.

Thus, games of truly perfect recall can be seen as a special subclass of games with "standard" perfect recall, as captured by the original semantics of $\mathsf{ATL}^\star_{\mathsf{i}}$ in [Schobbens 2004].

5.3. Other Notions of Subjective Ability

For the types of subjective ability that look at the paths starting from the common knowledge neighborhood (resp. distributed knowledge neighborhood) of the current state, we observe that:

(1) The inclusion result (Proposition 5.6) carries over to the $\models_{i_C}$ (resp. $\models_{i_D}$) semantics. A closer inspection of the proof of Proposition 5.5 (invariance under unfolding) reveals that it never uses any specific properties of the "everybody knows" relation $\approx_A$. Thus, one can replace every occurrence of $plays^i$ with $plays^i_C$ (resp. $plays^i_D$) in the proof, and obtain a proof that:
— $M, q \models^{nf}_{i_C} \varphi$ iff $T^{nf}(M, q), q \models^{nf}_{i_C} \varphi$ iff $T^{nf}(M, q), q \models_{i_C} \varphi$, and
— $M, q \models^{nf}_{i_D} \varphi$ iff $T^{nf}(M, q), q \models^{nf}_{i_D} \varphi$ iff $T^{nf}(M, q), q \models_{i_D} \varphi$.
In consequence, we get that $Val(\mathsf{ATL}^\star_{i_C}) \subseteq Val(\mathsf{ATL}^\star_{\mathsf{nf},i_C})$ and $Val(\mathsf{ATL}^\star_{i_D}) \subseteq Val(\mathsf{ATL}^\star_{\mathsf{nf},i_D})$ by the same argument as in the proof of Proposition 5.6.

(2) The formula $\varphi \equiv \langle\langle a \rangle\rangle \Box \mathsf{p} \to \mathsf{E}\bigcirc \langle\langle a \rangle\rangle \Box \mathsf{p}$, used in the proof of Proposition 5.7, contains only singleton coalitions. Thus, its truth value is the same no matter which subjective semantics is used (they coincide for abilities of individual agents). In consequence, one can use the argument in the proof of Proposition 5.7 to show that φ is a validity of $\mathsf{ATL}^\star_{\mathsf{nf},i_C}$ and $\mathsf{ATL}^\star_{\mathsf{nf},i_D}$, but not of $\mathsf{ATL}^\star_{i_C}$ and $\mathsf{ATL}^\star_{i_D}$. This in turn implies that $Val(\mathsf{ATL}^\star_{\mathsf{nf},i_C}) \not\subseteq Val(\mathsf{ATL}^\star_{i_C})$ and $Val(\mathsf{ATL}^\star_{\mathsf{nf},i_D}) \not\subseteq Val(\mathsf{ATL}^\star_{i_D})$ by the same argument as in the proof of Proposition 5.6.

As a result, we obtain the following analogue of Theorem 5.8.

THEOREM 5.9.

(*1*) $Val(\mathsf{ATL}^\star_{i_C}) \subsetneq Val(\mathsf{ATL}^\star_{\mathsf{nf},i_C})$. *Consequently, also* $Sat(\mathsf{ATL}^\star_{\mathsf{nf},i_C}) \subsetneq Sat(\mathsf{ATL}^\star_{i_C})$.
(*2*) $Val(\mathsf{ATL}^\star_{i_D}) \subsetneq Val(\mathsf{ATL}^\star_{\mathsf{nf},i_D})$. *Consequently, also* $Sat(\mathsf{ATL}^\star_{\mathsf{nf},i_D}) \subsetneq Sat(\mathsf{ATL}^\star_{i_D})$.

6. MODEL CHECKING

Model checking is the problem of establishing whether a logical formula is satisfied in a given structure. The complexity of model checking is especially insightful for the development of verification tools for particular logics. Decision problems for games with imperfect information and perfect recall are known to be computationally hard even for 2-player games, and undecidable when coalitional strategies are considered [Peterson and Reif 1979; Pnueli and Rosner 1990; Peterson et al. 2001; Berwanger and Kaiser 2010; Dima and Tiplea 2011]. On the other hand, some decidable cases have also been identified [Chatterjee et al. 2007; Berwanger and Kaiser 2010; Berwanger et al. 2011; Guelev et al. 2011; Berwanger and Mathew 2014; Berwanger et al. 2015], most notably the verification of abilities of single individual agents. In this section, we point out that the truly perfect recall semantics shares (un)decidability of model checking with the standard perfect recall variant, but it makes verification more costly in the decidable cases.

Note: the purpose of this section is not to establish new technical results, but to shortly review the impact of "no forgetting" from the computational perspective. Because of that, we mostly focus on the complexity of model checking for ATL (without "$\star$") which is the syntactic fragment of $\mathsf{ATL}^\star$ where every strategic modality is immediately followed by a temporal operator, and every temporal operator is immediately preceded by a strategic modality.

6.1. Verification of $\mathsf{ATL}^{\star}_{\mathsf{nf,i}}$ vs. $\mathsf{ATL}^{\star}_{\mathsf{i}}$

Since the two semantics coincide for perfect information, we only consider the case of imperfect information. We begin by quoting a rather pessimistic result for verification of alternating-time specifications with imperfect information and perfect recall.

THEOREM 6.1 ([DIMA AND TIPLEA 2011]). *Model checking $\mathsf{ATL}_{\mathsf{i}}$ (and hence also $\mathsf{ATL}^{\star}_{\mathsf{i}}$) is undecidable.*

The undecidability result was proved by a reduction of the halting problem that employed a 3-player iCGS and a formula with no nested cooperation modalities [Dima and Tiplea 2011]. We recall that, when no nested cooperation modalities are present, the $\mathsf{ATL}^{\star}_{\mathsf{nf,i}}$ semantics coincides with that of $\mathsf{ATL}^{\star}_{\mathsf{i}}$. Thus, we get the following as a consequence.

COROLLARY 6.2. *Model checking of $\mathsf{ATL}_{\mathsf{nf,i}}$ (and hence also $\mathsf{ATL}^{\star}_{\mathsf{nf,i}}$) is undecidable.*

What about decidable fragments of the problem? Restricting the class of models to turn-based structures will not work, as the proof in [Dima and Tiplea 2011] can be adapted to yield a turn-based model in the reduction of the halting problem. On the other hand, [Guelev et al. 2011] proposed an effective algorithm for model checking ATL with truly perfect recall and coalitions whose strategies are based on the distributed knowledge relation within the coalition. This is in turn equivalent to model checking the "singleton fragment" of ATL with truly perfect recall, i.e., the fragment of $\mathsf{ATL}_{\mathsf{nf,i}}$ with strategic modalities restricted to coalitions of size at most one [Kaźmierczak et al. 2014]). More formally:

THEOREM 6.3 ([GUELEV ET AL. 2011]). *Model checking of the singleton fragment of $\mathsf{ATL}_{\mathsf{nf,i}}$ is decidable and can be done in nonelementary time with respect to the size of the model and the length of the formula. For formulae of strategic depth*[9] *of at most k, the model checking problem is in* $\mathbf{kEXPTIME}$.

The lower bounds can be derived from the following result for model checking of temporal-epistemic logic with perfect recall.

THEOREM 6.4 ([SHILOV ET AL. 2004; SHILOV ET AL. 2006]). *Model checking CTLK with perfect recall is decidable with nonelementary upper and lower bounds with respect to the size of the model and the length of the formula. For formulae of knowledge depth*[10] *of at most k, the model checking problem is* $\mathbf{kEXPTIME}$*-complete.*

From this, we obtain the following characterization of complexity for the singleton fragments of $\mathsf{ATL}_{\mathsf{i}}$ and $\mathsf{ATL}_{\mathsf{nf,i}}$.

THEOREM 6.5. *Model checking of the singleton fragment of $\mathsf{ATL}_{\mathsf{nf,i}}$ is complete in nonelementary time with respect to the size of the model and the length of the formula. For formulae of strategic depth of at most k, the model checking problem is* $\mathbf{kEXPTIME}$*-complete.*

PROOF. The upper bound is straightforward from Theorem 6.3.

For the lower bound, observe that every formula of CTLK with perfect recall can be equivalently translated into $\mathsf{ATL}_{\mathsf{nf,i}}$ by replacing $\mathsf{A}\varphi$ with $\langle\langle\emptyset\rangle\rangle\varphi$ and $K_a\varphi$ with $\langle\langle a\rangle\rangle\bot\,\mathcal{U}\,\varphi$. By Theorem 6.4, we get the hardness result.[11] □

THEOREM 6.6. *Model checking of the singleton fragment of $\mathsf{ATL}_{\mathsf{i}}$ is* $\mathbf{EXPTIME}$*-complete with respect to the size of the model and the length of the formula. It remains* $\mathbf{EXPTIME}$*-complete for formulae of bounded length.*

[9] The maximum number of nested strategic operators different than $\langle\langle\emptyset\rangle\rangle$. The fact that $\langle\langle\emptyset\rangle\rangle$ can be excluded from the count follows from the state-splitting construction: it only produces exponential blowup of the state space for nonempty coalitions.
[10] The maximum number of nested epistemic operators.
[11] The idea of the proof was hinted to us by Catalin Dima in a personal communication.

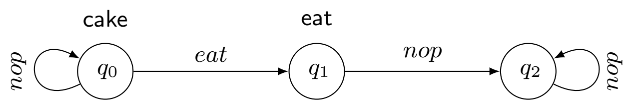


Fig. 8. Have the cake or eat it? The cake dilemma: CGS M_8

PROOF. We proceed recursively (bottom-up), starting with subformulae that contain no nested strategic modalities, and replacing them with fresh atomic propositions that hold in exactly the same subset of states. By Theorem 6.3, the procedure runs in exponential time. The lower bound follows from Theorem 6.4. $\square$

Thus, assuming truly perfect recall under imperfect information changes the verification complexity for worse. On the other hand, it simplifies the underlying tree unfoldings (cf. Section 5.2), which can potentially make verification easier, especially for simple (and short) formulae of ATL. Moreover, model checking $\mathsf{ATL}_{\mathsf{nf,i}}$ is no harder than verification of temporal-epistemic logic with perfect recall [van der Meyden and Shilov 1999; Shilov et al. 2004; Shilov et al. 2006]. Thus, $\mathsf{ATL}_{\mathsf{nf,i}}$ buys the expressivity of strategic operators for no extra computational price. We also note that the increase in complexity is due to the perfect recall assumption itself, and *not* due to the interaction of strategic modalities with truly perfect recall.

6.2. Other Semantics of Ability under Imperfect Information

Since mutual knowledge, common knowledge, and distributed knowledge coincide for single agents, the same results follow for the subjective ability based on distributed knowledge and common knowledge neighborhoods, i.e., ATL_{i_D}, ATL_{nf,i_D}, ATL_{i_C} and ATL_{nf,i_C}. More precisely:

THEOREM 6.7.

(1) *Model checking the full logics of* ATL_{i_D}, ATL_{i_C}, ATL_{nf,i_D}, *and* ATL_{nf,i_C} *is undecidable.*
(2) *Model checking of the singleton fragment of* ATL_{i_D} *and* ATL_{i_C} *is* $\mathbf{EXPTIME}$*-complete with respect to the size of the model and the length of the formula. It remains* $\mathbf{EXPTIME}$*-complete for formulae of bounded length.*
(3) *Model checking of the singleton fragment of* ATL_{nf,i_D} *and* ATL_{nf,i_C} *is complete in nonelementary time with respect to the size of the model and the length of the formula. For formulae of strategic depth of at most* k*, the model checking problem is* $\mathbf{kEXPTIME}$*-complete.*

Interestingly, since ATL_{i_o} and ATL_{nf,i_o} coincide, we get that true perfect recall does not increase the complexity of verification for objective abilities:

THEOREM 6.8.

(1) *Model checking* ATL_{i_o} *is undecidable.*
(2) *Model checking of the singleton fragment of* ATL_{i_o} *and* ATL_{nf,i_o} *is* $\mathbf{EXPTIME}$*-complete.*

This suggests that reasoning about objective abilities under imperfect information can be a valuable alternative to the more popular, subjective semantics also from the complexity point of view.

7. TRULY PERFECT RECALL UNDER STRATEGIC COMMITMENT

We pointed out in Section 3 that, when proceeding from a higher-level goal to a subgoal, the semantics of $\mathsf{ATL}^\star$ "forgets" past observations of agents—even if the agents are assumed to have perfect recall. A similar feature was observed in [Ågotnes et al. 2007; Brihaye et al. 2009] with respect to agents' strategies: they do not persist from outer to nested strategic modalities. Strategies in $\mathsf{ATL}^\star$ are *revocable* in the sense that an agent is not bound by her strategy anymore when proceeding from the main game to a subgame. In many cases, this makes the meaning of $\mathsf{ATL}^\star$ specifications counterintuitive.

Consider formula $\varphi \equiv \langle\!\langle a \rangle\!\rangle \Box \langle\!\langle a \rangle\!\rangle \bigcirc \mathsf{eat}$ which says that Alice has a strategy to maintain forever her ability to eat the cake. The formula is easily satisfiable in $\mathsf{ATL}^\star$; a simple model for φ is presented in Figure 8. However, the only way for Alice to keep her ability to eat the cake is by never eating the cake, which is somewhat paradoxical. If Alice executes the strategy, she will deprive herself of the ability she wants to maintain in the first place.

Also, consider the formula $\langle\!\langle c \rangle\!\rangle \Box \langle\!\langle a, b \rangle\!\rangle \Diamond \mathsf{married_{ab}}$ which expresses that Charlie can provide Alice and Bob with the ability to get married. One can imagine that Charlie can achieve that, e.g., by becoming the local superintendent registrar or a local priest, and granting every marriage request from his friends. However, the $\mathsf{ATL}^\star$ interpretation of the formula is that Alice and Bob must find a strategy for $\Diamond \mathsf{married_{ab}}$ *against every possible behavior of Charlie*, despite the fact that Charlie did select his strategy in order to make $\langle\!\langle a, b \rangle\!\rangle \Diamond \mathsf{married_{ab}}$ true.

We will now briefly recall two variants of $\mathsf{ATL}^\star$ that were proposed to handle specifications where persistence of strategies is important [Ågotnes et al. 2007; Brihaye et al. 2009]. We will also extend the variants to the case of imperfect information. After that, we will point out that the "forgetting" phenomenon applies also to $\mathsf{ATL}^\star$ with persistent strategies, and we will look at the formal consequences of the fact. In what follows, we only focus on the most popular semantics for agents with perfect recall (i.e., perfect information and subjective ability under uncertainty). We conjecture that analogous results can be obtained for the other kinds of ability under imperfect information, but we leave the detailed analysis out of this work.

7.1. ATL* with Long-Term Commitment

7.1.1. Irrevocable $\mathsf{ATL}^\star$. The simplest interpretation of $\mathsf{ATL}^\star$ formulae, that assumes "irrevocability" or long-term commitment of agents' strategies, was proposed in [Ågotnes et al. 2007]. The logic takes the syntax of $\mathsf{ATL}^\star$ but changes the semantics to ensure persistence of strategies.[12] In the original version, this is done by unfolding the CGS to a tree and then pruning all the transitions that are inconsistent with the strategies selected by agents. Here, we give an equivalent semantics based on the idea of "strategy contexts", cf. [Brihaye et al. 2009; Jamroga et al. 2005; Walther et al. 2007] or Section 7.1.2 for more details. Formally, the semantics is given in terms of the semantic relation $\models_{x,c}$ that interprets a formula, given an iCGS, a path in it, and a *strategy context* which "stores" the strategies selected so far by the agents. Parameter $x \in \{I, i\}$ indicates whether we assume agents to have perfect or imperfect information. We first recall the definition of *strategy update*.

Definition 7.1 (*Strategy update [Brihaye et al. 2009]*). Given collective strategies s_A and s_B of group A and B, respectively, we define the strategy $s_B \dagger s_A$ as follows: $s_B \dagger s_A|_i(h) = s_A|_i(h)$ if $i \in A$ and $s_B \dagger s_A|_i(h) = s_B|_i(h)$ if $i \in B \setminus A$.

Thus, $s_B \dagger s_A$ combines the strategies s_A and s_B, and the actions specified by s_A will *override* those specified by s_B for agents in $A \cap B$. The semantic relation $\models_{x,c}$ is defined as follows:

$M, \lambda, s \models_{x,c} p$ iff $\lambda[0] \in \pi(p)$ (where $p \in \Pi$);
$M, \lambda, s \models_{x,c} \neg\varphi$ iff $M, \lambda, s \not\models_{x,c} \varphi$;
$M, \lambda, s \models_{x,c} \varphi_1 \wedge \varphi_2$ iff $M, \lambda, s \models_{x,c} \varphi_1$ and $M, \lambda, s \models_{x,c} \varphi_2$;
$M, \lambda, s \models_{x,c} \langle\!\langle A \rangle\!\rangle\varphi$ iff there is a collective x-strategy s_A such that for all $\lambda' \in \mathit{plays}^x(\lambda[0], s_A \dagger s)$, we have $M, \lambda', s_A \dagger s \models_{x,c} \varphi$;
$M, \lambda, s \models_{x,c} \bigcirc\varphi$ iff $M, \lambda[1, \infty], s \models_{x,c} \varphi$;
$M, \lambda, s \models_{x,c} \varphi_1 \mathcal{U} \varphi_2$ iff there is $i \in \mathbb{N}_0$ such that $M, \lambda[i, \infty], s \models_{x,c} \varphi_2$ and for all $0 \leq j < i$, we have $M, \lambda[j, \infty], s \models_{x,c} \varphi_1$.

[12]The logic is called $\mathsf{MIATL}^\star$ for "$\mathsf{ATL}^\star$ with irrevocable strategies and memory" in [Ågotnes et al. 2007]. Here, we rather use the acronym $\mathsf{ATL}^\star_c$ to indicate that we deal with the same formulae as in $\mathsf{ATL}^\star$, but strategy commitment is required in the semantics.

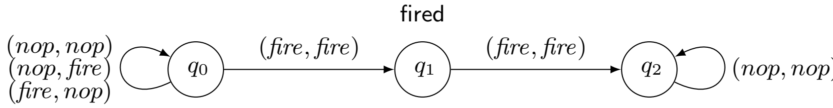


Fig. 9. The cake dilemma with multi-player coordination: CGS M_9

Note that the above clauses define two different logics: $\mathsf{ATL}^\star_{\mathsf{I,c}}$ ($\mathsf{ATL}^\star$ with perfect information and strategy commitment) which corresponds to $\mathsf{MIATL}^\star$ from [Ågotnes et al. 2007], and $\mathsf{ATL}^\star_{\mathsf{i,c}}$ ($\mathsf{ATL}^\star$ with imperfect information and strategy commitment) which, to the best of our knowledge has not been considered anywhere yet. We also observe that in $\mathsf{ATL}^\star_{\mathsf{I,c}}$ and $\mathsf{ATL}^\star_{\mathsf{i,c}}$ strategies are indeed irrevocable. The first strategy selected by an agent is never overridden by a subsequent strategy. This is reflected in the order of strategy updates: the oldest updates are applied last.

Given a state q and a state formula φ, we define $M, q \models_{x,c} \varphi$ iff $M, \lambda, s_\emptyset \models_{x,c} \varphi$ for all $\lambda \in \Lambda_M(q)$, where $s_\emptyset$ refers to the only possible strategy of the empty coalition. Moreover, φ is *valid* in $\mathsf{ATL}^\star_{\mathsf{x,c}}$ iff $M, q \models_{x,c} \varphi$ for every iCGS M and state q in it.

Example 7.2. Consider CGS M_8 in Figure 8. Unlike in the standard semantics of $\mathsf{ATL}^\star$, we have $M_8, q_0 \not\models_{x,c} \langle\!\langle a \rangle\!\rangle \Box \langle\!\langle a \rangle\!\rangle \bigcirc \mathsf{eat}$. Suppose that the formula holds in M_8, q_0. On one hand, if Alice selects the "eat" strategy eat then the formula $\varphi \equiv \langle\!\langle a \rangle\!\rangle \Box \langle\!\langle a \rangle\!\rangle \bigcirc \mathsf{eat}$ can be only satisfied if $M_8, q_0 q_1 (q_2)^\omega, eat \models_{x,c} \Box \langle\!\langle a \rangle\!\rangle \bigcirc \mathsf{eat}$, but this is not possible since $M_8, q_1 \not\models_{x,c} \langle\!\langle a \rangle\!\rangle \bigcirc \mathsf{eat}$. On the other hand, if Alice selects the "do nothing" strategy nop then the formula φ can be only satisfied if $M_8, (q_0)^\omega, nop \models_{x,c} \Box \langle\!\langle a \rangle\!\rangle \bigcirc \neg\mathsf{eat}$, which is not possible since $M_8, q_0, nop \not\models_{x,c} \langle\!\langle a \rangle\!\rangle \bigcirc \neg\mathsf{cake}$: Alice is already bound to her "do nothing" strategy.

7.1.2. ATL with Strategy Contexts.* "$\mathsf{ATL}^\star$ with strategy contexts" [Brihaye et al. 2009] offers a more elaborate framework for reasoning about strategy commitment. Compared to $\mathsf{ATL}^\star_{\mathsf{I,c}}$ and $\mathsf{ATL}^\star_{\mathsf{i,c}}$, it allows agents to commit, override and revoke their strategies. The syntax extends the language of $\mathsf{ATL}^\star$ with a *strategic release operator* $\rangle A \langle$, and the semantics $\models_{x,sc}$ differs from the one presented in Section 7.1.1 by the following clauses:

$M, \lambda, s \models_{x,sc} \langle\!\langle A \rangle\!\rangle \varphi$ iff there is a collective x-strategy s_A such that for each $\lambda' \in plays^x(\lambda[0], s \dagger s_A)$, $M, \lambda', s \dagger s_A \models_{x,sc} \varphi$;
$M, \lambda, s \models_{x,sc} \rangle A \langle\, \varphi$ iff $M, \lambda, s_{\setminus A} \models_{x,sc} \varphi$;

where $s_{\setminus A}$ denotes the strategy context with strategies of agents from A being removed. We emphasize the swap in the update operator in comparison to $\mathsf{ATL}^\star_{\mathsf{x,c}}$: $s \dagger s_A$ versus $s_A \dagger s$. Thus, according to relation $\models_{x,sc}$, newer strategies override older ones.

Again, we define $M, q \models_{x,sc} \varphi$ iff $M, \lambda, s_\emptyset \models_{x,sc} \varphi$ and $M, q, s \models_{x,sc} \varphi$ iff $M, \lambda, s \models_{x,sc} \varphi$ for all $\lambda \in \Lambda_M(q)$. Moreover, φ is *valid* iff $M, q \models_{x,sc} \varphi$ for every M, q.

Example 7.3. Figure 9 depicts a two-player variant of the "cake dilemma." This time, two agents – Alice and Bob – are needed to fire a missile; if they do not coordinate, the game stays in the initial state q_0. Now, we have for example that $M_9, q_0 \models_{x,sc} \langle\!\langle a \rangle\!\rangle \mathcal{N} \langle\!\langle b \rangle\!\rangle \bigcirc \mathsf{fired}$.[13] To achieve this, we select the strategy "*fire the missile whenever the system is in state* q_0*, and do nop elsewhere*" when evaluating the modality $\langle\!\langle a \rangle\!\rangle$ for Alice, and the same strategy later on, when evaluating the modality $\langle\!\langle b \rangle\!\rangle$ for Bob. However, $M_9, q_0 \not\models_x \langle\!\langle a \rangle\!\rangle \mathcal{N} \langle\!\langle b \rangle\!\rangle \bigcirc \mathsf{fired}$: when evaluating $\langle\!\langle b \rangle\!\rangle \bigcirc \mathsf{fired}$ in the standard $\mathsf{ATL}^\star$ semantics, Bob cannot assume that Alice's strategy will persist, and therefore every strategy of his may result in the path q_0^ω. This illustrates that the semantics $\models_{x,sc}$ allows agents to base their decisions on the strategies previously selected by other players, which was not possible in the original semantics of $\mathsf{ATL}^\star$.

[13]Recall that $\mathcal{N}\varphi \equiv \varphi\, \mathcal{U}\, \varphi$ stands for "φ holds in the current moment."

Also, $M_9, q_0 \models_{x,sc} \langle\langle a,b\rangle\rangle\Box(\neg\mathsf{fired} \wedge \rangle a,b\langle\ \langle\langle a,b\rangle\rangle\bigcirc\,\mathsf{fired})$: the right-hand-side of the conjunction expresses that Alice and Bob release their strategy and then are free to choose a new one. Therefore, "$\mathsf{ATL}^\star$ with strategy contexts" allows agents to explicitly *decommit* previously selected strategies.

Finally, notice that $M_9, q_0 \models_{x,sc} \langle\langle a,b\rangle\rangle\Box\langle\langle a,b\rangle\rangle\bigcirc\,\mathsf{fired}$ because Alice and Bob override their former strategy when trying to satisfy the subformula $\langle\langle a,b\rangle\rangle\bigcirc\,\mathsf{fired}$. Therefore, unlike $\mathsf{ATL}^\star_{\mathsf{x,c}}$, strategies in "$\mathsf{ATL}^\star$ with strategy contexts" are *not* irrevocable.

Remark 7.4. It was shown in [Brihaye et al. 2009, Proposition 3] that the strategic release operator $\rangle A\langle$ adds no expressive power to "$\mathsf{ATL}^\star$ with strategy contexts." For instance, $\rangle A\langle\ \langle\langle B\rangle\rangle\gamma$ can be equivalently rewritten as $\langle\langle B\rangle\rangle\neg\langle\langle A\setminus B\rangle\rangle\neg\gamma$ (B has a strategy such that agents outside A cannot prevent γ). Thus, we can omit strategic release from the syntax without losing generality of our results. In the rest of the paper, we will only use formulae defined by the standard syntax of $\mathsf{ATL}^\star$, exactly like for the other logics studied here.

Assuming that we use only the standard syntax of $\mathsf{ATL}^\star$, the semantic relations $\models_{x,sc}$ define, again, two logical systems: $\mathsf{ATL}^\star_{\mathsf{I,sc}}$ ($\mathsf{ATL}^\star$ with perfect information and strategy contexts), which corresponds to $\mathsf{ATL}^\star_{\mathsf{sc},\infty}$ from [Brihaye et al. 2009], and $\mathsf{ATL}^\star_{\mathsf{i,sc}}$ ($\mathsf{ATL}^\star$ with imperfect information and strategy contexts) which is a new variant of alternating-time temporal logic to our best knowledge.

7.1.3. Comparing Logics of Strategy Commitment. In Sections 7.1.1 and 7.1.2, we have presented four variants of alternating-time logic for reasoning about persistent strategies: $\mathsf{ATL}^\star_{\mathsf{I,c}}$, $\mathsf{ATL}^\star_{\mathsf{i,c}}$, $\mathsf{ATL}^\star_{\mathsf{I,sc}}$, and $\mathsf{ATL}^\star_{\mathsf{i,sc}}$. How do the variants relate? We show that $\mathsf{ATL}^\star_{\mathsf{x,c}}$ can be in fact embedded in $\mathsf{ATL}^\star_{\mathsf{x,sc}}$ for $x \in \{I,i\}$. To this end, we define a translation tr_A such that, for all models M, path λ, strategy context s_A and formula φ of $\mathsf{ATL}^\star$, we have that $M,\lambda,s_A \models_{x,c} \varphi$ iff $M,\lambda,s_A \models_{x,sc} tr_A(\varphi)$:

$$
\begin{aligned}
tr_A(p) &= p\\
tr_A(\neg\varphi) &= \neg tr_A(\varphi)\\
tr_A(\varphi_1\wedge\varphi_2) &= tr_A(\varphi_1)\wedge tr_A(\varphi_2)\\
tr_A(\langle\langle B\rangle\rangle\varphi) &= \langle\langle B\setminus A\rangle\rangle tr_{B\cup A}(\varphi)\\
tr_A(\bigcirc\varphi) &= \bigcirc\, tr_A(\varphi)\\
tr_A(\varphi_1\,\mathcal{U}\,\varphi_2) &= tr_A(\varphi_1)\,\mathcal{U}\,tr_A(\varphi_2)
\end{aligned}
$$

PROPOSITION 7.5. *For each iCGS M, path $\lambda \in \Lambda_M$, collective strategy s_A where $A \subseteq \mathbb{A}\mathrm{gt}$, and $\mathsf{ATL}^\star_{\mathsf{c}}$ formula φ we have that $M,\lambda,s_A \models_{x,c} \varphi$ iff $M,\lambda,s_A \models_{x,sc} tr_A(\varphi)$ for $x \in \{i,I\}$.*

PROOF. The proof is done by induction over the formula structure of φ.

Base cases: Case $\underline{\varphi = p}$ is straightforward. Case $\underline{\varphi = \langle\langle B\rangle\rangle\gamma}$ where γ contains no cooperation modalities. First, we observe ($\star$) $s_B \dagger s_A = s_{B\setminus A} \dagger s_A = s_A \dagger s_{B\setminus A}$: all individual strategies of agents in $A\cap B$ are overridden by those of s_A in $s_B \dagger s_A$ thus if two coalitions are disjoint the update order for their collective strategies is irrelevant; and ($\star\star$) $tr_X(\gamma) = \gamma$ if γ contains no cooperation modalities.

Next, $M,\lambda,s_A \models_{x,c} \langle\langle B\rangle\rangle\gamma$ iff there exists s_B such that $\forall\lambda' \in plays^x(\lambda[0], s_B \dagger s_A)$ we have $M,\lambda',s_B \dagger s_A \models_{x,c} \gamma$. By ($\star$),($\star\star$) $M,\lambda',s_B\dagger s_A \models_{x,c} \gamma$ iff $M,\lambda',s_A \dagger s_{B\setminus A} \models_{x,c} tr_{B\cup A}(\gamma)$. Since $tr_{B\cup A}(\gamma)$ contains no cooperation modalities, the semantics $\models_{x,c}$ and $\models_{x,sc}$ coincide for $tr_{B\cup A}(\gamma)$: $M,\lambda',s_A\dagger s_{B\setminus A} \models_{x,c} tr_{B\cup A}(\gamma)$ iff $M,\lambda',s_A\dagger s_{B\setminus A} \models_{x,sc} tr_{B\cup A}(\gamma)$ iff $M,\lambda,s_A \models_{x,c} \langle\langle B\setminus A\rangle\rangle tr_{B\cup A}(\gamma)$ iff $M,\lambda,s_A \models_{x,c} tr_A(\langle\langle B\rangle\rangle\gamma)$.

Induction step: Cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi'\wedge\varphi''}$ are straightforward. Case $\underline{\varphi = \langle\langle B\rangle\rangle\gamma}$ where γ contains cooperation modalities. Let ξ be an arbitrary outermost occurrence of a formula $\langle\langle B'\rangle\rangle\gamma'$ in γ. We label each state $\lambda'[0]$ such that $M,\lambda',s_B \dagger s_A \models_{x,c} \xi$ (for some $\lambda' \in \Lambda_M$ and each

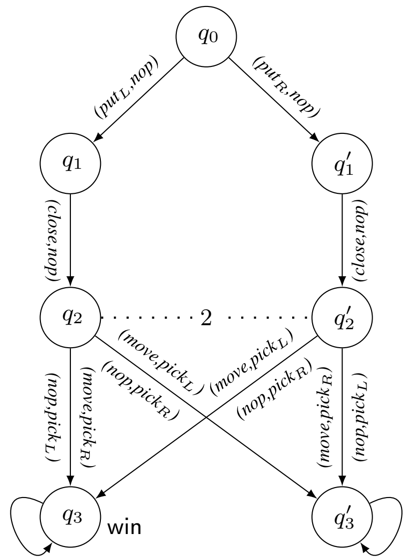


Fig. 10. M_{10}: the shell game with final shuffling

collective strategy $s_B \dagger s_A$), with a new proposition p_ξ which does not occur in Π. We note that each strategy $s_{A\cup B}$ may be written as $s_B \dagger s_A$. By induction hypothesis, $M, \lambda', s_B \dagger s_A \models_{x,c} \xi$ iff $M, \lambda', s_B \dagger s_A \models_{x,sc} tr_{B\cup A}(\xi)$. We replace each occurrence of a subformula ξ by p_ξ in γ. The resulting formula contains no cooperation modalities. We proceed exactly as in the second base case. $\square$

Note that, in particular, $M, q \models_{x,c} \varphi$ iff $M, q \models_{x,sc} tr_\emptyset(\varphi)$. Thus, we obtain the following as an immediate consequence of Proposition 7.5.

COROLLARY 7.6. $\mathsf{ATL}^\star_{\mathsf{x,sc}}$ *is at least as expressive and as distinguishable as* $\mathsf{ATL}^\star_{\mathsf{x,c}}$ *for* $x \in \{i, I\}$. *That is,* $\mathsf{ATL}^\star_{\mathsf{x,c}} \preceq_d \mathsf{ATL}^\star_{\mathsf{x,sc}}$ *and* $\mathsf{ATL}^\star_{\mathsf{x,c}} \preceq_e \mathsf{ATL}^\star_{\mathsf{x,sc}}$.

PROPOSITION 7.7. $\mathsf{ATL}^\star_{\mathsf{x,c}}$ *is not as distinguishable as* $\mathsf{ATL}^\star_{\mathsf{x,sc}}$, *i.e.,* $\mathsf{ATL}^\star_{\mathsf{x,sc}} \not\preceq_d \mathsf{ATL}^\star_{\mathsf{x,c}}$.

The proof is given in Appendix A.3.
Corollary 7.6 and Proposition 7.7 imply the following.

THEOREM 7.8. $\mathsf{ATL}^\star_{\mathsf{x,sc}}$ *is strictly more distinguishing and expressive than* $\mathsf{ATL}^\star_{\mathsf{x,c}}$. *That is,* $\mathsf{ATL}^\star_{\mathsf{x,sc}} \prec_d \mathsf{ATL}^\star_{\mathsf{x,c}}$ *and* $\mathsf{ATL}^\star_{\mathsf{x,sc}} \prec_e \mathsf{ATL}^\star_{\mathsf{x,c}}$.

7.2. Commitment and Truly Perfect Recall

In the previous section, we have presented two variants of how strategic commitment can be added to the standard semantics of $\mathsf{ATL}^\star$. How doe it change the picture? First, we show that persistent strategies *per se* do not rule out counterintuitive effects.

Example 7.9 (*Shell game with final shuffling*). Consider the iCGS M_{10} from Figure 10. It depicts a version of the shell game in which the shuffler can switch the shells quickly in the very last moment (after the ball has been enclosed). This is modeled by the action $move$ available to the shuffler at states q_2 and q_2'. First, we observe that – in contrast to Example 2.4 – the guesser no longer has a strategy to eventually win in q_0, that is, $M_{10}, q_0 \not\models_i \langle\!\langle 2\rangle\!\rangle\Diamond\mathsf{win}$. On the other hand, if the shuffler is committed to some (arbitrary) strategy s_1, the guesser can secure a win: $M_{10}, q_0, s_1 \models_{i,sc} \langle\!\langle 2\rangle\!\rangle\Diamond\mathsf{win}$. Thus, $M_{10}, q_0 \models_{i,sc} \neg\langle\!\langle 1\rangle\!\rangle\neg\langle\!\langle 2\rangle\!\rangle\Diamond\mathsf{win}$: there is no strategy of player 1 that would prevent player 2 from winning, provided that 1 knows the strategy of 2 in advance.[14] So, there are situations where it makes sense to consider strategy commitment.

On the other hand, we still have $M_{10}, q_2, s_1 \models_{i,sc} \neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$ regardless of the strategy s_1 of the shuffler. Thus, again, $M_{10}, q_0 \models_{i,sc} \langle\langle 1\rangle\rangle\Diamond\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$, which is counterintuitive.

Example 7.9 shows that, when reasoning about persistent strategies of agents, we also need to carefully define the semantics in order to avoid the "forgetting" phenomenon. This leads to the following semantics of strategic ability with commitment.

Definition 7.10 ($\mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$). The semantics $\models^{nf}_{x,sc}$ of $\mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$, with $x \in \{i, I\}$ is defined by changing the clause for $\langle\langle A\rangle\rangle\gamma$ in the following way:

$M, \lambda, k, s \models^{nf}_{x,sc} \langle\langle A\rangle\rangle\varphi$ iff there is a collective x-strategy s_A such that, for each $\lambda' \in$ *plays*$^x(\lambda[0,k], s \dagger s_A)$, we have $M, \lambda', s \dagger s_A \models^{nf}_{x,sc} \varphi$;

The other semantic clauses are defined analogously to Definition 3.2.

Definition 7.11 ($\mathsf{ATL}^\star_{\mathsf{nf,x,c}}$). The semantics of $\mathsf{ATL}^\star_{\mathsf{nf,x,c}}$ can be defined in two alternative (equivalent) ways: either we directly update the semantic clauses from Section 7.1.1, or we apply the translation from Section 7.1.3 and use the semantics of $\mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$ from Definition 7.10. To simplify the presentation, we chose the latter option. Thus, for an $\mathsf{ATL}^\star$ formula φ, we define $M, \lambda, k, s_A \models^{nf}_{x,c} \varphi$ iff $M, \lambda, k, s_A \models^{nf}_{x,sc} tr_A(\varphi)$.

Remark 7.12. The way in which we defined the semantics of $\mathsf{ATL}^\star_{\mathsf{nf,x,c}}$ immediately implies that $\mathsf{ATL}^\star_{\mathsf{nf,x,c}} \preceq_e \mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$, and hence also $\mathsf{ATL}^\star_{\mathsf{nf,x,c}} \preceq_d \mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$.

The notions of truth of a state formula in a pointed model and validity of a formula are defined as in the previous sections.

Example 7.13 (*Commitment and truly perfect recall*). Consider the iCGS M_{10} again. Similarly to Example 3.4, we now have that $M_{10}, q_0 \models^{nf}_{i,sc} \neg\langle\langle 1\rangle\rangle\Diamond\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$. If the shuffler commits to doing nothing (action nop) in q_2 (resp. q_2'), the guesser uses the history-based strategy from Example 2.4. If the shuffler intends late shuffling (action $move$), the guesser uses the "swap" strategy of picking, i.e. in comparison to the previous strategy selects $pick_L$ instead of $pick_R$, and $pick_R$ for $pick_L$.

Remark 7.14. Another variant of alternating-time temporal logic with imperfect information, truly perfect recall, and strategy contexts has been already proposed and studied in [Guelev and Dima 2012]. The differences to our work are as follows. First, the variant of $\mathsf{ATL}^\star$ in [Guelev and Dima 2012] features very ornate syntax, including past tense operators, collective knowledge operators, and strategic modalities that indicate which members of the coalition are allowed to revise their strategies, and which are not. This makes comparative analysis rather difficult to conduct. Secondly, their semantics differs from the standard approach by assuming runs to be interleaved sequences of states and action profiles, which affects indistinguishability relations. Thirdly, the interaction between the strategic and the epistemic aspects is further complicated by extending epistemic relations to indistinguishability over strategies. All of this is for a reason: the focus of [Guelev and Dima 2012] is on providing a rich logical framework where all aspects of persistent play under imperfect information can be modeled, described, and studied. In contrast, we start with a simple update of the standard semantics of persistent play from [Ågotnes et al. 2007; Brihaye et al. 2009], and focus on a comparison between different semantics of perfect recall.

[14] Note that $\neg\langle\langle 1\rangle\rangle\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win}$ is a formula of $\mathsf{ATL}^\star$ but not ATL. The same property can be equivalently expressed with the ATL formula $\neg\langle\langle 1\rangle\rangle\mathcal{N}(\neg\langle\langle 2\rangle\rangle\Diamond\mathsf{win})$.

7.3. Expressivity of ATL* with Commitment and Truly Perfect Recall

In this section we study the relation between the "forgetting" and "no forgetting" variants of $\mathsf{ATL}^\star$ with strategy commitment, i.e., $\mathsf{ATL}^\star_{\mathsf{x,c}}$ vs. $\mathsf{ATL}^\star_{\mathsf{nf,x,c}}$ and $\mathsf{ATL}^\star_{\mathsf{x,sc}}$ vs. $\mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$.

7.3.1. Perfect Information. Similarly to Section 4.1 we have the following results for the perfect information case. The proof is done analogously to Proposition 3.6 (see Appendix A.3 for details).

PROPOSITION 7.15. *For all M, λ, s and every $\mathsf{ATL}^\star$ formula φ, we have that $M, \lambda, 0, s \models^{nf}_{I,sc} \varphi$ iff $M, \lambda, s \models_{I,sc} \varphi$.*

Thus, similarly to the setting with standard non-persistent strategies, the truly perfect recall and standard perfect recall semantics are equivalent under perfect information. We obtain the following as an immediate corollary.

COROLLARY 7.16. *$\mathsf{ATL}^\star_{\mathsf{I,sc}}$ and $\mathsf{ATL}^\star_{\mathsf{nf,I,sc}}$ are equally expressive and have the same sets of validities. By Proposition 7.5, the same holds for $\mathsf{ATL}^\star_{\mathsf{I,c}}$ vs. $\mathsf{ATL}^\star_{\mathsf{nf,I,c}}$.*

7.3.2. Imperfect Information. The next theorem shows that, under imperfect information, the truly perfect recall semantics for persistent strategies differs from the standard one.

PROPOSITION 7.17. *There is a pointed iCGS (M, q) and an $\mathsf{ATL}^\star$ formula φ such that $M, q \models_{i,sc} \varphi$ and $M, q \not\models^{nf}_{i,sc} \varphi$.*

PROOF. The result follows from Examples 7.9 and 7.13 for $M = M_{10}$, $q = q_0$ and $\varphi \equiv \langle\!\langle 1 \rangle\!\rangle \Diamond \neg \langle\!\langle 2 \rangle\!\rangle \Diamond \mathsf{win}$. $\square$

The following propositions compare the distinguishing power of the truly perfect recall vs. "forgetting" semantics for $\mathsf{ATL}^\star_{\mathsf{i,sc}}$, and $\mathsf{ATL}^\star_{\mathsf{i,c}}$. The proofs are given in Appendix A.3.

PROPOSITION 7.18. *There are iCGSs which satisfy the same formulae of $\mathsf{ATL}^\star_{\mathsf{i,sc}}$, but can be distinguished in $\mathsf{ATL}^\star_{\mathsf{nf,i,c}}$. That is, $\mathsf{ATL}^\star_{\mathsf{nf,i,c}} \not\preceq_d \mathsf{ATL}^\star_{\mathsf{i,sc}}$.*

PROPOSITION 7.19. *There are iCGSs which satisfy the same formulae of $\mathsf{ATL}^\star_{\mathsf{nf,i,sc}}$, but can be distinguished in $\mathsf{ATL}^\star_{\mathsf{i,c}}$. That is, $\mathsf{ATL}^\star_{\mathsf{i,c}} \not\preceq_d \mathsf{scATL}^\star_{\mathsf{nf,i}}$.*

Combining Propositions 7.18 and 7.19 with Proposition 7.5 and Remark 7.12, we get that:

THEOREM 7.20. *For all $\mathsf{L}_1 \in \{\mathsf{ATL}^\star_{\mathsf{i,sc}}, \mathsf{ATL}^\star_{\mathsf{i,c}}\}$ and $\mathsf{L}_2 \in \{\mathsf{ATL}^\star_{\mathsf{nf,i,sc}}, \mathsf{ATL}^\star_{\mathsf{nf,i,c}}\}$, the logics L_1 and L_2 are incomparable with respect to distinguishing and expressive power.*

7.4. Comparing Validity Sets

Finally, we compare the sets of valid sentences of the "no forgetting" logics $\mathsf{ATL}^\star_{\mathsf{nf,x,c}}$ and $\mathsf{ATL}^\star_{\mathsf{nf,x,sc}}$ with those of their "forgetting" counterparts.

7.4.1. Perfect Information. The following is an immediate consequence of Propositions 7.15 and 7.5.

PROPOSITION 7.21. *$\mathsf{ATL}^\star_{\mathsf{I,sc}}$ and $\mathsf{ATL}^\star_{\mathsf{nf,I,sc}}$ have the same sets of validities. The same holds for $\mathsf{ATL}^\star_{\mathsf{I,c}}$ vs. $\mathsf{ATL}^\star_{\mathsf{nf,I,c}}$.*

7.4.2. Imperfect Information. For imperfect information, we can obtain the result below analogously to Proposition 5.6, i.e., by using epistemic tree unfoldings.

PROPOSITION 7.22. *$Val(\mathsf{scATL}^\star_{\mathsf{i}}) \subseteq Val(\mathsf{scATL}^\star_{\mathsf{nf,i}})$. Thus, also $Val(\mathsf{ATL}^\star_{\mathsf{i,c}}) \subseteq Val(\mathsf{ATL}^\star_{\mathsf{nf,i,c}})$.*

A detailed proof is presented in Appendix A.3.

PROPOSITION 7.23. $\mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{nf,i,c}}) \not\subseteq \mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{i,c}})$. *Thus, also* $\mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{nf,i,sc}}) \not\subseteq \mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{i,sc}})$.

PROOF. We can reuse formula φ from the proof of Proposition 5.7 because it does not contain nested modalities—apart from $\emptyset$—and over such formulae the commitment and no-commitment semantics coincide. □

Similarly to Theorem 5.8, we immediately obtain the following characterization:

THEOREM 7.24. $\mathit{Val}(\mathsf{scATL}^{\star}_{\mathsf{i}}) \subsetneq \mathit{Val}(\mathsf{scATL}^{\star}_{\mathsf{nf,i}})$ *and* $\mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{i,c}}) \subsetneq \mathit{Val}(\mathsf{ATL}^{\star}_{\mathsf{nf,i,c}})$

Thus, games of truly perfect recall can be seen as a special subclass of games with "standard" perfect recall also in the case of persistent strategies, as captured by the semantics of strategic ability proposed in [Ågotnes et al. 2007; Brihaye et al. 2009].

8. CONCLUSIONS

In this paper, we formally study a semantics of strategic ability which propagates agents' observations to nested strategic modalities. Thus, unlike the standard semantics of alternating-time logic $\mathsf{ATL}^{\star}$, it models agents who *never* forget their past observations. Formally, this is done by keeping not only the future, but also the past in the path that serves as the reference to the semantic relation. Most importantly, we investigate the relationship between the two approaches, encoded by the "forgetting" and the "truly perfect recall" semantics. Both approaches turn out to be equivalent for agents with perfect information of the global state of the system. In the more interesting case of incomplete information, however, the two kinds of semantics are significantly different. In particular, they yield logical systems that are incomparable with respect to their expressive as well as distinguishing power. Equally interesting is the comparison of general properties of games induced by the different semantics. Formally, this means to compare the sets of validities generated by the alternative semantics. We show that the validities according to the truly perfect recall semantics form a strict superset of the "forgetting" validities. Thus, they capture a more specific class of games than the standard $\mathsf{ATL}^{\star}_{\mathsf{i}}$.

The same pattern of results carries over to the setting where agents are assumed to persist with their strategies by some kind of (irrevocable or revocable) strategic commitment.

It is usually assumed that agents A can enforce φ if they have a strategy which succeeds on all executions from states that are epistemically possible for someone in A. That is, the semantics looks at the execution paths starting from the "everybody knows" neighborhood. On the other hand, there is no particular conceptual reason for using this type of collective indistinguishability. An interesting question arises: are our results specific to this particular semantics, or do they hold in other variants of $\mathsf{ATL}^{\star}$ for imperfect information? Along this line, we show that the same properties are obtained as long as the set of execution paths is based on an established notion of collective knowledge (common knowledge, distributed knowledge). Furthermore, rather surprisingly, we prove that the pattern is significantly different for so called *objective ability*: in that case, the true perfect recall is already provided by the standard future-based semantics.

From the computational point of view, reasoning about agents with perfect recall is always complex, but assuming truly perfect recall for subjective abilities makes it even harder, as exemplified by the nonelementary complexity of model checking even for abilities of individual agents. That seems to suggest that, in practice, using the truly perfect recall semantics puts one at a disadvantage. To remain within **EXPTIME**, one should stick to the future-based semantics, or switch to objective ability for which the two kinds of semantics coincide. We point out, however, that the most important thing in verification is to *verify the right property in the right model*. In particular, the input model must match the relevant aspects of the system, and the formula together with its semantic interpretation must capture the property that we want to verify. The same applies to any other kind of reasoning and analysis. We believe that, when reasoning about agents who are supposed to memorize their observations, the truly perfect recall semantics of ability is the right one. Our technical results show that it cannot be replaced by the standard, more compositional semantics of $\mathsf{ATL}^\star_\mathsf{i}$, despite the latter offering somewhat lower complexity of related computational tasks. This is because the properties definable in both frameworks are essentially different, and because $\mathsf{ATL}^\star_\mathsf{i}$ allows for paradoxical specifications that should not be satisfiable for agents with real perfect memory.

Acknowledgements. Wojciech Jamroga acknowledges the support of the National Research Fund (FNR) Luxembourg under the project GALOT (INTER/DFG/12/06), the 7th Framework Programme of the European Union under the Marie Curie IEF project ReVINK (PIEF-GA-2012-626398), and the National Centre for Research and Development (NCBR), Poland, under the PolLux project VoteVerif (POLLUX-IV/1/2016).

REFERENCES

T. Ågotnes. 2006. Action and Knowledge in Alternating-time Temporal Logic. *Synthese* 149, 2 (2006), 377–409.

T. Ågotnes, V. Goranko, and W. Jamroga. 2007. Alternating-time Temporal Logics with Irrevocable Strategies. In *Proceedings of TARK XI*. 15–24.

T. Ågotnes, V. Goranko, W. Jamroga, and M. Wooldridge. 2015. Knowledge and Ability. In *Handbook of Epistemic Logic*, H.P. van Ditmarsch, J.Y. Halpern, W. van der Hoek, and B.P. Kooi (Eds.). College Publications, 543–589.

T. Ågotnes and D. Walther. 2009. A Logic of Strategic Ability Under Bounded Memory. *Journal of Logic, Language and Information* 18, 1 (2009), 55–77.

N. Alechina, B. Logan, N.H. Nga, and A. Rakib. 2009. Verifying Properties of Coalitional Ability under Resource Bounds. In *Proceedings of the Logics for Agents and Mobility (LAM)*.

N. Alechina, B. Logan, H.N. Nguyen, and F. Raimondi. 2017. Model-checking for Resource-Bounded ATL with production and consumption of resources. *J. Comput. System Sci.* 88 (2017), 126–144. DOI:http://dx.doi.org/10.1016/j.jcss.2017.03.008

N. Alechina, B. Logan, H.N. Nguyen, and A. Rakib. 2010. Resource-Bounded Alternating-Time Temporal Logic. In *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 481–488.

N. Alechina, B. Logan, H.N. Nguyen, and A. Rakib. 2011. Logic for Coalitions with Bounded Resources. *Journal of Logic and Computation* 21, 6 (2011), 907–937.

N. Alechina, B. Logan, and M. Whitsey. 2004. A Complete and Decidable Logic for Resource-Bounded Agents. In *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IEEE Computer Society, 606–613.

R. Alur, L. de Alfaro, R. Grossu, T.A. Henzinger, M. Kang, C.M. Kirsch, R. Majumdar, F.Y.C. Mang, and B.-Y. Wang. 2001. jMocha: A Model-Checking Tool that Exploits Design Structure. In *Proceedings of International Conference on Software Engineering (ICSE)*. IEEE Computer Society Press, 835–836.

R. Alur, L. de Alfaro, T. A. Henzinger, S.C. Krishnan, F.Y.C. Mang, S. Qadeer, S.K. Rajamani, and S. Tasiran. 2000. *MOCHA: Modularity in Model Checking*. Technical Report. University of Berkeley.

R. Alur, T. A. Henzinger, and O. Kupferman. 1997. Alternating-Time Temporal Logic. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society Press, 100–109.

R. Alur, T. A. Henzinger, and O. Kupferman. 2002. Alternating-Time Temporal Logic. *J. ACM* 49 (2002), 672–713. DOI:http://dx.doi.org/10.1145/585265.585270

Francesco Belardinelli. 2014. Reasoning about Knowledge and Strategies: Epistemic Strategy Logic. In *Proceedings 2nd International Workshop on Strategic Reasoning, SR 2014, Grenoble, France, April 5-6, 2014*. 27–33. DOI:http://dx.doi.org/10.4204/EPTCS.146.4

F. Belardinelli, A. Lomuscio, A. Murano, and S. Rubin. 2017. Verification of Multi-agent Systems with Imperfect Information and Public Actions. In *Proceedings of AAMAS*. 1268–1276.

N. Belnap and M. Perloff. 1988. Seeing to It That: a Canonical Form for Agentives. *Theoria* 54 (1988), 175–199.

R. Berthon, B. Maubert, A. Murano, S. Rubin, and M. Y. Vardi. 2017. Strategy logic with imperfect information. In *Proceedings of LICS*. 1–12. DOI:http://dx.doi.org/10.1109/LICS.2017.8005136

D. Berwanger and L. Kaiser. 2010. Information Tracking in Games on Graphs. *Journal of Logic, Language and Information* 19, 4 (2010), 395–412. DOI:http://dx.doi.org/10.1007/s10849-009-9115-8

D. Berwanger, L. Kaiser, and B. Puchala. 2011. A Perfect-Information Construction for Coordination in Games. In *Proceedings of FSTTCS*. 387–398.

D. Berwanger and A.B. Mathew. 2014. Infinite games with finite knowledge gaps. *CoRR* abs/1411.5820 (2014). http://arxiv.org/abs/1411.5820

D. Berwanger, A.B. Mathew, and M. van den Bogaard. 2015. Hierarchical Information Patterns and Distributed Strategy Synthesis. In *Proceedings of ATVA (Lecture Notes in Computer Science)*, Vol. 9364. Springer, 378–393. DOI:http://dx.doi.org/10.1007/978-3-319-24953-7_28

T. Brihaye, A. Da Costa Lopes, F. Laroussinie, and N. Markey. 2009. ATL with Strategy Contexts and Bounded Memory. In *Proceedings of LFCS (Lecture Notes in Computer Science)*, Vol. 5407. Springer, 92–106.

N. Bulling and J. Dix. 2010. Modelling and Verifying Coalitions using Argumentation and ATL. *Inteligencia Artificial* 14, 46 (March 2010), 45–73.

N. Bulling and B. Farwer. 2010a. Expressing Properties of Resource-Bounded Systems: The Logics RTL* and RTL. In *Proceedings of Computational Logic in Multi-Agent Systems (CLIMA) (Lecture Notes in Computer Science)*, Vol. 6214. 22–45.

N. Bulling and B. Farwer. 2010b. On the (Un-)Decidability of Model Checking Resource-Bounded Agents. In *Proceedings of ECAI (Frontiers in Artificial Intelligence and Applications)*, Vol. 215. IOS Press, 567–572.

N. Bulling and W. Jamroga. 2014. Comparing Variants of Strategic Ability: How Uncertainty and Memory Influence General Properties of Games. *Journal of Autonomous Agents and Multi-Agent Systems* 28, 3 (2014), 474–518.

N. Bulling, W. Jamroga, and J. Dix. 2008. Reasoning about Temporal Properties of Rational Play. *Annals of Mathematics and Artificial Intelligence* 53, 1-4 (2008), 51–114.

N. Bulling, W. Jamroga, and M. Popovici. 2013. Agents with truly perfect recall in alternating-time temporal logic (extended abstract). In *Proceedings of AAMAS*. 1561–1562.

N. Bulling, W. Jamroga, and M. Popovici. 2014. Agents With Truly Perfect Recall: Expressivity and Validities. In *Proceedings of ECAI*. 177–182.

S. Busard. 2017. *Symbolic Model Checking of Multi-Modal Logics: Uniform Strategies and Rich Explanations*. Ph.D. Dissertation. Universite Catholique de Louvain.

S. Busard, C. Pecheur, H. Qu, and F. Raimondi. 2014. Improving the Model Checking of Strategies under Partial Observability and Fairness Constraints. In *Formal Methods and Software Engineering*. Lecture Notes in Computer Science, Vol. 8829. Springer, 27–42. DOI:http://dx.doi.org/10.1007/978-3-319-11737-9_3

S. Busard, C. Pecheur, H. Qu, and F. Raimondi. 2015. Reasoning about memoryless strategies under partial observability and unconditional fairness constraints. *Information and Computation* 242 (2015), 128–156. DOI:http://dx.doi.org/10.1016/j.ic.2015.03.014

K. Chatterjee, L. Doyen, T.A. Henzinger, and J.-F. Raskin. 2007. Algorithms for Omega-regular Games of Incomplete Information. *Logical Methods in Computer Science* 3, 3 (2007).

K. Chatterjee, T.A. Henzinger, and N. Piterman. 2010. Strategy Logic. *Information and Computation* 208, 6 (2010), 677–693.

B. F. Chellas. 1969. *The Logical Form of Imperatives*. Ph.D. Dissertation. Stanford University.

T. Chen, V. Forejt, M. Kwiatkowska, D. Parker, and A. Simaitis. 2013. PRISM-games: A Model Checker for Stochastic Multi-Player Games. In *Proceedings of Tools and Algorithms for Construction and Analysis of Systems (TACAS) (Lecture Notes in Computer Science)*, Vol. 7795. Springer, 185–191.

E.M. Clarke and B.-H. Schlingloff. 2001. Model checking. In *Handbook of Automated Reasoning*, A. Robinson and A. Voronkov (Eds.). Elsevier, 1635–1790.

R. Diaconu and C. Dima. 2012. Model-Checking Alternating-Time Temporal Logic with Strategies Based on Common Knowledge is Undecidable. *Applied Artificial Intelligence* 26, 4 (2012), 331–348. DOI:http://dx.doi.org/10.1080/08839514.2012.652905

C. Dima, C. Enea, and D.P. Guelev. 2010. Model-Checking an Alternating-time Temporal Logic with Knowledge, Imperfect Information, Perfect Recall and Communicating Coalitions. In *Proceedings of Games, Automata, Logics and Formal Verification (GandALF)*. 103–117.

C. Dima and F.L. Tiplea. 2011. Model-checking ATL under Imperfect Information and Perfect Recall Semantics is Undecidable. *CoRR* abs/1102.4225 (2011).

E.A. Emerson and J.Y. Halpern. 1986. "Sometimes" and "Not Never" Revisited: On Branching versus Linear Time Temporal Logic. *J. ACM* 33, 1 (1986), 151–178.

R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. 1995. *Reasoning about Knowledge*. MIT Press.

D.P. Guelev and C. Dima. 2008. Model-Checking Strategic Ability and Knowledge of the Past of Communicating Coalitions. In *Proceedings of DALT*. 75–90.

D.P. Guelev and C. Dima. 2012. Epistemic ATL with Perfect Recall, Past and Strategy Contexts. In *Proceedings of Computational Logic in Multi-Agent Systems (CLIMA) (Lecture Notes in Computer Science)*, Vol. 7486. Springer, 77–93. DOI:http://dx.doi.org/10.1007/978-3-642-32897-8_7

D.P. Guelev, C. Dima, and C. Enea. 2011. An alternating-time temporal logic with knowledge, perfect recall and past: axiomatisation and model-checking. *Journal of Applied Non-Classical Logics* 21, 1 (2011), 93–131.

P. Hawke. 2010. Coordination, Almost Perfect Information and Strategic Ability. In *Proceedings of LAMAS*.

X. Huang and R. van der Meyden. 2014. Symbolic Model Checking Epistemic Strategy Logic. In *Proceedings of AAAI Conference on Artificial Intelligence*. 1426–1432.

W. Jamroga and T. Ågotnes. 2007. Constructive Knowledge: What Agents Can Achieve under Incomplete Information. *Journal of Applied Non-Classical Logics* 17, 4 (2007), 423–475.

W. Jamroga, M. Knapik, and D. Kurpiewski. 2017. Fixpoint Approximation of Strategic Abilities under Imperfect Information. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, 1241–1249.

W. Jamroga, W. Penczek, P. Dembiński, and A. Mazurkiewicz. 2018. Towards Partial Order Reductions for Strategic Ability. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS. To appear.

W. Jamroga and W. van der Hoek. 2004. Agents that Know how to Play. *Fundamenta Informaticae* 63, 2–3 (2004), 185–219.

W. Jamroga, W. van der Hoek, and M. Wooldridge. 2005. Intentions and Strategies in Game-Like Scenarios. In *Progress in Artificial Intelligence: Proceedings of EPIA 2005*, Carlos Bento, Amílcar Cardoso, and Gaël Dias (Eds.). Lecture Notes in Artificial Intelligence, Vol. 3808. Springer, 512–523.

M. Kacprzak and W. Penczek. 2004. Unbounded Model Checking for Alternating-Time Temporal Logic. In *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IEEE Computer Society, 646–653. DOI: http://dx.doi.org/10.1109/AAMAS.2004.10089

P. Kaźmierczak, T. Ågotnes, and W. Jamroga. 2014. Multi-Agency Is Coordination And (Limited) Communication. In *Proceedings of PRIMA (Lecture Notes in Computer Science)*, Vol. 8861. Springer, 91–106.

A. Lomuscio, H. Qu, and F. Raimondi. 2015. MCMAS: An Open-Source Model Checker for the Verification of Multi-Agent Systems. *International Journal on Software Tools for Technology Transfer* (2015). DOI: http://dx.doi.org/10.1007/s10009-015-0378-x Availabe online.

A. Lomuscio and F. Raimondi. 2006. MCMAS : A Model Checker for Multi-Agent Systems. In *Proceedings of Tools and Algorithms for Construction and Analysis of Systems (TACAS) (Lecture Notes in Computer Science)*, Vol. 4314. Springer, 450–454.

J. McCarthy and P.J. Hayes. 1969. Some Philosophical Problems From the Standpoint of Artificial Intelligence. In *Machine Intelligence 4*, B. Meltzer and D. Michie (Eds.). Edinburgh University Press, 463–502.

F. Mogavero, A. Murano, G. Perelli, and M.Y. Vardi. 2014. Reasoning About Strategies: On the Model-Checking Problem. *ACM Transactions on Computational Logic* 15, 4 (2014), 1–42.

F. Mogavero, A. Murano, and M.Y. Vardi. 2010a. Reasoning About Strategies. In *Proceedings of FSTTCS*. 133–144.

F. Mogavero, A. Murano, and M.Y. Vardi. 2010b. Relentful Strategic Reasoning in Alternating-time Temporal Logic. In *Proceedings of LPAR*. Springer, 371–386.

R.C. Moore. 1985. A Formal Theory of Knowledge and Action. In *Formal Theories of the Commonsense World*, J. Hobbs and R.C. Moore (Eds.). Ablex Publishing Corp.

R. C. Moore. 1977. Reasoning About Knowledge and Action. In *Proceedings International Joint Conference on Artificial Intelligence (IJCAI)*. William Kaufmann, 223–227.

G. Peterson and J. Reif. 1979. Multiple-person alternation. In *Proceedings of the 20th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society Press, 348–363.

G. Peterson, J. Reif, and S. Azhar. 2001. Lower Bounds for Multiplayer Noncooperative Games of Incomplete Information. *Computers and Mathematics with Applications* 41, 7 (2001), 957–992.

J. Pilecki, M.A. Bednarczyk, and W. Jamroga. 2014. Synthesis and Verification of Uniform Strategies for Multi-Agent Systems. In *Proceedings of CLIMA XV (Lecture Notes in Computer Science)*, Vol. 8624. Springer, 166–182.

A. Pnueli and R. Rosner. 1990. Distributed reactive systems are hard to synthesize. In *Proceedings of the 31th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society Press, 746–757.

G. Ryle. 1949. *The Concept of Mind*. Chicago University Press.

P. Y. Schobbens. 2004. Alternating-Time Logic with Imperfect Recall. *Electronic Notes in Theoretical Computer Science* 85, 2 (2004), 82–93.

N. V. Shilov, N. O. Garanina, and K.-M. Choe. 2006. Update and Abstraction in Model Checking of Knowledge and Branching Time. *Fundamenta Informaticae* 72, 1-3 (2006), 347–361.

N. V. Shilov, N. O. Garanina, and N. A. Kalinina. 2004. Model Checking Knowledge, Actions and Fixpoints. In *Proceedings of CS&P 2004*. 351–357.

W. van der Hoek, W. Jamroga, and M. Wooldridge. 2005. A Logic for Strategic Reasoning. In *Proceedings of AAMAS'05*. 157–164.

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

R. van der Meyden and N.V. Shilov. 1999. Model Checking Knowledge and Time in Systems with Perfect Recall (Extended Abstract). In *Proceedings of Foundations of Software Technology and Theoretical Computer Science (Lecture Notes in Computer Science)*, Vol. 1738. Springer, 432–445.

H. van Ditmarsch and S. Knight. 2014. Partial Information and Uniform Strategies. In *Proceedings of Computational Logic in Multi-Agent Systems (CLIMA) (Lecture Notes in Computer Science)*. Springer, 183–198. DOI: http://dx.doi.org/10.1007/978-3-319-09764-0_12

S. Vester. 2013. Alternating-time temporal logic with finite-memory strategies. In *Proceedings of GandALF (EPTCS)*. 194–207. DOI: http://dx.doi.org/10.4204/EPTCS.119.17

D. Walther, W. van der Hoek, and M. Wooldridge. 2007. Alternating-time Temporal Logic with Explicit Strategies. In *Proceedings TARK XI*. Presses Universitaires de Louvain, 269–278.

A. PROOFS

A.1. Proofs of Section 4

PROPOSITION 4.4. *There are pointed iCGSs which satisfy the same* $\mathsf{ATL}^\star_\mathsf{i}$*-formulae, but can be distinguished in* $\mathsf{ATL}^\star_\mathsf{nf,i}$*. Thus,* $\mathsf{ATL}^\star_\mathsf{nf,i} \not\preceq_d \mathsf{ATL}^\star_\mathsf{i}$*, and hence also* $\mathsf{ATL}^\star_\mathsf{nf,i} \not\preceq_e \mathsf{ATL}^\star_\mathsf{i}$*.*

PROOF. Let M_2 and M_2' be the iCGSs shown in Figure 4 and φ be some $\mathsf{ATL}^\star$-formula. We define $M_2^\uparrow$ (resp. $M_2^\downarrow$) as the sub-model of M_2 obtained by keeping only the states $x \in \{a_0, b_0, a_1', b_1', a_2', b_2'\}$ (resp. $x \in \{a_0, b_0, a_1, b_1, a_2, b_2\}$) and removing all other states and subsequent transitions. The model $M_2'^\uparrow$ and $M_2'^\downarrow$ are defined analogously as sub-models of M_2'.

First, we observe that $M_2^\uparrow$ and $M_2^\downarrow$ (resp. $M_2'^\uparrow$, $M_2'^\downarrow$) are bisimilar. As a consequence, we have:

$$M_2, x_j \models_i \varphi \text{ iff } M_2, x_j' \models_i \varphi \text{ for } x \in \{a, b\} \text{ and } j \in \{1, 2\} \tag{1}$$

and analogously for M_2'. Moreover, we have

$$M_2, x_j \models_i \varphi \text{ iff } M_2', x_j \models_i \varphi \text{ for } x \in \{a, b, a', b'\} \text{ and } j \in \{1, 2\} \tag{2}$$

We prove:

$$(\star)\ M_2, a_0 \models_i \varphi \text{ iff } M_2', a_0 \models_i \varphi$$

by induction over the formula structure of φ.

Base cases: Case $\varphi = \mathsf{p}$. Straightforward. Case $\varphi = \langle\!\langle A \rangle\!\rangle\gamma$ where γ contains no cooperation modalities. For $A \in \{\emptyset, \{1\}\}$ the proposition follows immediately as each strategy of A generates the same set $\mathit{plays}^i_M(a_0, s_A)$ in both models.

Suppose $A = \{2\}$. Direction "⇐": as $\mathit{plays}^i_{M_2'}(a_0, s_A) = \mathit{out}_{M_2'}(a_0, s_A) \cup \mathit{out}_{M_2'}(b_0, s_A)$ and $\mathit{plays}^i_{M_2}(a_0, s_A) = \mathit{out}_{M_2}(a_0, s_A)$, we have $\mathit{plays}^i_{M_2}(a_0, s_A) \subseteq \mathit{plays}^i_{M_2'}(a_0, s_A)$. $(\star)$ follows immediately.

Direction "⇒": let s_2 be an arbitrary uniform strategy of player 2 in M_2. We investigate $\mathit{plays}^i_{M_2}(a_0, s_2)$. First: (i) we observe that player 2 cannot prevent any of a_1', b_1' from being possible next-states of the game. Second: (ii) for each action in $\{\alpha, \beta, \mu\}$ which player 2 may play in both $a_0 a_1'$ and $a_0 b_1'$ (the same action must be planned for both histories, since they are indistinguishable to 2 and s_2 is uniform), $\mathit{plays}^i_{M_2}(a_0, s_2)$ contains a path on which win eventually holds (e.g. $a_0 b_1'(b_2')^\omega$ if α is played) and one on which win never holds ($a_0 a_1'(a_2')^\omega$ if α is played). The statements (i),(ii) also hold in M_2'. Player 2 can neither prevent win nor ensure win from $a_0 a_1'$ and $a_0 b_1'$, in either model. This concludes case $A = \{2\}$.

Suppose $A = \{1, 2\}$. Direction "⇐ ": we again note $\mathit{plays}^i_{M_2}(a_0, s_A) \subseteq \mathit{plays}^i_{M_2'}(a_0, s_A)$.

Direction "⇒": Let $s_A = (s_1, s_2)$ be an arbitrary uniform strategy for A. We first observe that $\mathit{plays}^i_{M_2}(a_0, s_A)$ contains a unique path, however $\mathit{plays}^i_{M_2'}(a_0, s_A)$ contains two paths. Starting from s_A we build $s_A' = (s_1', s_2')$ such that both paths from $\mathit{plays}^i_{M_2'}(a_0, s_A')$ are propositionally equivalent to that from $\mathit{plays}^i_{M_2}(a_0, s_A)$: win is either maintained false or eventually (and thereafter always) fulfilled. The construction is as follows: $s_j'(a_0) = s_j(a_0)$ for $j \in \{1, 2\}$ (s_A' replicates s_A in the initial state);

for histories $h \in \{a_0 a_1, a_0 a_1'\}$:
if $s_2(h) = \alpha$ then $s_1'(h) = \epsilon$ and $s_2'(h) = \mu$;
if $s_2(h) = \beta$ then $s_1'(h) = \mu$ and $s_2'(h) = \mu$;
otherwise $s_1'(h) = s_1(h)$ and $s_2'(h) = s_2(h)$;
for history $h = a_0 b_1'$:
if $s_2(h) = \alpha$ then $s_1'(h) = \mu$ and $s_2'(h) = \mu$;
if $s_2(h) = \beta$ then $s_1'(h) = \epsilon$ and $s_2'(h) = \mu$;
otherwise $s_1'(h) = s_1(h)$ and $s_2'(h) = s_2(h)$;

For all other histories, the assigned actions are unimportant. We note that s_2 is uniform, and also that in the absence of actions μ and ϵ we would, e.g, have that $M_2, a_0 \models_i \langle\langle 1,2\rangle\rangle\Diamond\mathsf{win}$ but $M_2', a_0 \not\models_i \langle\langle 1,2\rangle\rangle\Diamond\mathsf{win}$.

Induction step: The cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\langle A\rangle\rangle\gamma$ where γ contains cooperation modalities. Let ξ be an arbitrary occurrence of an outermost formula $\langle\langle A'\rangle\rangle\gamma'$ in γ. We note that $M_2, x \models_i \xi$ iff $M_2', x \models_i \xi$ by induction hypothesis if $x = a_0$ and by (1-2), otherwise. We label each state x of M_2 and M_2' where ξ holds by a new proposition p_ξ. The resulting models retain properties (1-2). We replace each ξ by p_ξ in γ and obtain a formula without cooperation modalities. We proceed as in the second base case. This concludes the proof of $(\star)$.

In Ex. 4.3 we have shown that both pointed models can be distinguished in $\mathsf{ATL}^\star_{\mathsf{nf,i}}$. For every $\mathsf{ATL}^\star_{\mathsf{i}}$-formula φ we have $a_0 \in [\![\varphi, M_2]\!]_{\models_i}$ iff $a_0 \in [\![\varphi, M_2']\!]_{\models_i}$ but $a_0 \in [\![\varphi', M_2]\!]_{\models_i^{nf}}$ and $a_0 \notin [\![\varphi', M_2']\!]_{\models_i^{nf}}$ for some φ'. Thus, we have that $\mathsf{ATL}^\star_{\mathsf{nf,i}} \not\preceq_d \mathsf{ATL}^\star_{\mathsf{i}}$. □

A.2. Proofs of Section 5

PROPOSITION 5.5. *$M, q \models_i^{nf} \varphi$ iff $T^{nf}(M,q), q \models_i^{nf} \varphi$ iff $T^{nf}(M,q), q \models_i \varphi$, for all* $\mathsf{ATL}^\star$-*formulae, iCGSs M and states q.*

PROOF. To simplify the notations, we write T^{nf} instead of $T^{nf}(M,q)$. We observe the following, from the construction of T^{nf}:

(1) $\lambda \in \Lambda_M$ iff $\hat\lambda = h_0 h_1 \ldots h_i \ldots \in \Lambda_{T^{nf}}$ where $h_i = \lambda[0,i]$ and $i \in \mathbb{N}$;
$h \in \Lambda_M^{fin}$ iff $\hat h = h_0 h_1 \ldots h_{k-1} \in \Lambda_{T^{nf}}^{fin}$ where $h_i = h[0,i]$ for $i < k$ and $k = |h|$.
Moreover, λ and $\hat\lambda$ are propositionally equivalent.
(2) $h \approx_a^M h'$ iff $\hat h \approx_a^{T^{nf}} \hat h'$ for all $a \in \mathbb{A}\mathrm{gt}$.
(3) for each collective strategy s_A in M there exists a collective strategy $\hat s_A$ in T^{nf} such that $\lambda \in out(h, s_A)$ iff $\hat\lambda \in out(\hat h, \hat s_A)$ and vice-versa.

We assume that it is clear from context how the different histories are concatenated, e.g. $(q_0), (q_0 q_1), (q_0 q_1 q_2), \ldots$. To increase the readability, we omit parentheses.

We prove both equivalences separately.

(i) $M, q \models_i^{nf} \varphi$ iff $T^{nf}, q \models_i^{nf} \varphi$. We prove the stronger statement: $M, h \circ \lambda, k \models_i^{nf} \varphi$ iff $T^{nf}, \hat h \circ \hat\lambda, k \models_i^{nf} \varphi$, where $k = |h| - 1$. Our claim follows for $k = 0$. The proof is by induction over the formula structure of φ.

Base cases: Case $\varphi = p$. It is sufficient to note that $(h \circ \lambda)[k] \in \pi^M(p)$ iff $(\hat h \circ \hat\lambda)[k] \in \pi^{T^{nf}}(p)$.
Case $\varphi = \langle\langle A\rangle\rangle\gamma$ where γ contains no cooperation modalities. From (2-3) it follows that $\lambda \in plays_M^i(h, s_A)$ iff $\hat\lambda \in plays_{T^{nf}}^i(\hat h, \hat s_A)$. Since λ and $\hat\lambda$ are propositionally equivalent (c.f. (1)): $M, h \circ \lambda \models_i^{nf} \langle\langle A\rangle\rangle\gamma$ iff $T^{nf}, \hat h \circ \hat\lambda \models_i^{nf} \langle\langle A\rangle\rangle\gamma$.
Induction step: Cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\langle A\rangle\rangle\gamma$ where γ contains cooperation modalities. Let $\{\langle\langle B_i\rangle\rangle\psi_i \mid i = 1, \ldots, k\}$ be the set of outermost (positive) $\mathsf{ATL}^\star$-subformulae in γ. We define the following two sets:

$H_i = \{(h \circ \lambda, k) \mid M, h \circ \lambda, k \models_i^{nf} \langle\langle B_i\rangle\rangle\psi_i\}$

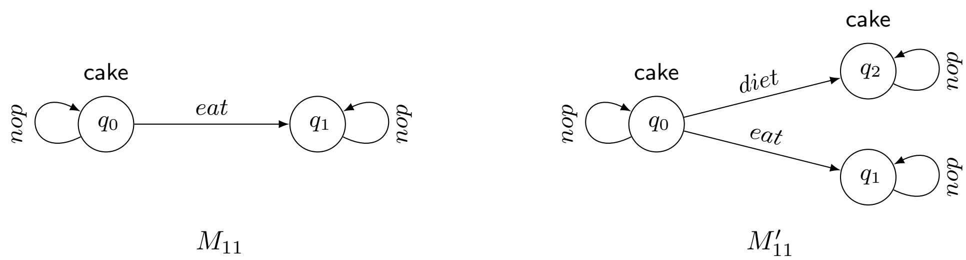


Fig. 11. Simplified cake dilemma: with and without diet

$\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid T^{nf}, \hat{h} \circ \hat{\lambda}, k \models_i^{nf} \langle\langle B_i \rangle\rangle \psi_i\}$

By induction hypothesis, we have that $\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid (h \circ \lambda, k) \in H_i\}$ as all formulae have a simpler structure that φ $(\star)$.

We now consider the two directions of (i): "$\Rightarrow$": Let s_A be a witnessing strategy for $\langle\langle A \rangle\rangle \gamma$ from history $h \circ \lambda$ and index k, i.e. $\forall \lambda \in \mathit{plays}_M^i(h, s_A) : M, \lambda, k \models_i^{nf} \gamma$. By (2) and (3), there exists a strategy $\hat{s}_A$ in T^{nf} such that:

$$\{\hat{\lambda} \mid \hat{\lambda} \in \mathit{plays}_M^i(h, s_A)\} = \{\lambda \mid \lambda \in \mathit{plays}_{T^{nf}}^i(\hat{h}, \hat{s}_A)\}$$

We note that for all h it holds that $\hat{s}_A(\hat{h}) = s_A(h) = s_A(\mathit{last}(\hat{h}))$. The claim now follows by $(\star)$. The direction "$\Leftarrow$" follows exactly the same argument.

(ii) $T^{nf}, q \models_i^{nf} \varphi$ iff $T^{nf}, q \models_i \varphi$. As before, we prove the stronger claim $T^{nf}, \hat{h} \circ \hat{\lambda}, k \models_i^{nf} \varphi$ iff $T^{nf}, \mathit{last}(\hat{h}) \circ \hat{\lambda} \models_i \varphi$ where $k = |\hat{h}| - 1$, by induction over the formula structure of φ. Then, the claim follows for $k = 0$.

Base cases: Case $\varphi = p$ is straightforward. Case $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ contains no cooperation modalities. From (1-3) we observe that for each uniform collective strategy s_A and history $\hat{h} \in \Lambda_{T^{nf}}^{fin}$, there exists a uniform collective strategy s'_A such that $\mathit{plays}_{T^{nf}}^i(\hat{h}, s_A) = \{\hat{h} \circ \hat{\lambda} \mid \mathit{last}(\hat{h}) \circ \hat{\lambda} \in \mathit{plays}_{T^{nf}}^i(\mathit{last}(\hat{h}), s'_A)\}$. Informally, each state $\mathit{last}(\hat{h})$ encodes a history. Hence, if coalition A has a perfect-recall strategy w.r.t. $\hat{h}$, then cf. (2-3) it can implement a strategy with the same effects from $\mathit{last}(\hat{h})$ alone. Similarly, for each collective strategy s'_A and $\hat{h} \in \Lambda_{T^{nf}}^{fin}$ there exists a collective strategy s_A such that $\mathit{plays}_{T^{nf}}^i(\mathit{last}(\hat{h}), s'_A) = \{\mathit{last}(\hat{h}) \circ \hat{\lambda} \mid \hat{h} \circ \hat{\lambda} \in \mathit{plays}_{T^{nf}}^i(\hat{h}, s_A)\}$. Thus $T^{nf}, \hat{h} \circ \hat{\lambda}, k \models_i^{nf} \langle\langle A \rangle\rangle \gamma$ iff $T^{nf}, \mathit{last}(\hat{h}) \circ \hat{\lambda} \models_i \langle\langle A \rangle\rangle \gamma$.

Induction step: Cases $\varphi = \neg \varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ contains cooperation modalities.

Let $\{\langle\langle B_i \rangle\rangle \psi_i \mid i = 1, \ldots, k\}$ be the set of outermost (positive) $\mathsf{ATL}^\star$-subformulae in γ. We also define the following sets:

$\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid T^{nf}, \hat{h} \circ \hat{\lambda}, k \models_i^{nf} \langle\langle B_i \rangle\rangle \psi_i\}$ and

$\hat{L}_i = \{\hat{\lambda} \mid T^{nf}, \hat{\lambda} \models_i \langle\langle B_i \rangle\rangle \psi_i\}$.

By induction hypothesis: $\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid \mathit{last}(\hat{h}) \circ \hat{\lambda} \in \hat{L}_i\}$ $(\star)$. We now consider the two directions of (ii): "$\Rightarrow$": Let s_A be a witnessing strategy for $\langle\langle A \rangle\rangle \gamma$ from $\hat{h} \circ \hat{\lambda}$ and index k, i.e. $\forall \hat{\lambda}' \in \mathit{plays}_{T^{nf}}^i(s_A, \hat{h})$, we have $T^{nf}, \hat{\lambda}', k \models_i^{nf} \gamma$. By (2) and (3) there exists a s'_A such that:

$$\mathit{plays}_{T^{nf}}^i(s_A, \hat{h}) = \{\hat{h} \circ \hat{\lambda} \mid \mathit{last}(\hat{h}) \circ \hat{\lambda} \in \mathit{plays}_{T^{nf}}^i(s'_A, \mathit{last}(\hat{h}))\}$$

The claim now follows by $(\star)$. Direction "$\Leftarrow$" follows exactly the same argument.

□

A.3. Proofs of Section 7

PROPOSITION 7.7. $\mathsf{ATL}^\star_{\mathsf{x,c}}$ *is not as distinguishable as* $\mathsf{ATL}^\star_{\mathsf{x,sc}}$*, i.e.,* $\mathsf{ATL}^\star_{\mathsf{x,sc}} \not\preceq_d \mathsf{ATL}^\star_{\mathsf{x,c}}$.

PROOF. We show that $M_{11}, q_0 \models_{x,c} \varphi$ iff $M'_{11}, q_0 \models_{x,c} \varphi$ for all $\varphi \in \mathsf{ATL}^\star_{\mathsf{x,c}}$, for models M_{11} and M'_{11} shown in Figure 11. We first observe that:

$$M_{11}, q_1, s \models_{x,c} \varphi \text{ iff } M'_{11}, q_1, s \models_{x,c} \varphi \text{ for all } \varphi \in \mathsf{ATL}^\star_{\mathsf{x,c}} \text{ and all strategies } s \neq s_\emptyset \tag{1}$$

$$\mathit{plays}^x_{M_{11}}(q, s) = \mathit{out}_{M_{11}}(q, s) \text{ for all states } q \text{ of } M_{11} \text{ and all strategies } s. \tag{2}$$

We prove: $(\star)$ $M_{11}, q_0, s \models_{x,c} \varphi \implies M'_{11}, q_0, s \models_{x,c} \varphi$ for all strategies s in M_{11} and all $\varphi \in \mathsf{ATL}^\star_{\mathsf{x,c}}$.

Base cases: Case $\underline{\varphi = p}$ is straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ does not contain cooperation modalities. Since $\mathbb{A}\mathrm{gt} = \{1\}$ in M_{11}, we have only two possible coalitions: $A = \emptyset$ or $A = \{1\}$. We consider each possiblity in turn. First, suppose $A = \{1\}$, $M_{11}, q_0, s \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$ and let s_1 be a witnessing strategy i.e. $M_{11}, \lambda[0], s_1 \dagger s \models_{x,c} \gamma$ for all $\lambda \in \mathit{out}_{M_{11}}(q_0, s_1 \dagger s)$. Since $s_1 \dagger s$ is a $\{1\}$-strategy (if $s = s_\emptyset$ then $s_1 \dagger s = s_1$ otherwise: $s_1 \dagger s = s$), it follows that $\mathit{out}_{M_{11}}(q_0, s_1 \dagger s) = \mathit{out}_{M'_{11}}(q_0, s_1 \dagger s)$ by construction of the models. Then $M'_{11}, \lambda[0], s_1 \dagger s \models_{x,c} \gamma$ and therefore $M'_{11}, q_0, s \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$.

Second, suppose $A = \emptyset$. If $s \neq s_\emptyset$ the argument is exactly as above. If $s = s_\emptyset$ we observe that $\mathit{plays}^x_{M_{11}}(q_0, s_\emptyset) = \{q_0^+ q_1^\omega, q_0^\omega\}$ and that $\mathit{plays}^x_{M'_{11}}(q_0, s_\emptyset) = \{q_0^+ q_2^\omega, q_0^+ q_1^\omega, q_0^\omega\}$. It is sufficient to note that each path in $\mathit{plays}^x_{M_{11}}(q_0, s_\emptyset)$ is propositionally equivalent to one in $\mathit{plays}^x_{M'_{11}}(q_0, s_\emptyset)$ and vice-versa.

Induction step: Cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi' \wedge \varphi''}$ are straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ contains cooperation modalities and $A = \{1\}$. Suppose $M_{11}, q_0, s \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$ and let s_1 be a witnessing strategy. For each outermost $\mathsf{ATL}^\star$-subformula $\langle\!\langle B_i \rangle\!\rangle \psi_i$ in γ (with $i = 1, \ldots, k$), strategy s^* and state q such that $M_{11}, q, s^* \models_{x,c} \langle\!\langle B_i \rangle\!\rangle \psi_i$, we observe that s^* can only be of the form $s' \dagger s_1$. Hence $s^* = s_1$ since s_1 is irrevocable. Then $M_{11}, q, s^* \models_{x,c} \langle\!\langle B_i \rangle\!\rangle \psi_i$ iff $M_{11}, q, s_1 \models_{x,c} \psi_i$. Therefore we can treat φ exactly as in the base case. The same holds for $A = \emptyset$.

We prove $(\star\star)$ $M'_{11}, q_0, s' \models_{x,c} \varphi \implies M_{11}, q_0, t(s') \models_{x,c} \varphi$ where $t(s')$ is the strategy s such that $s(h) = \mathit{nop}$ if $s'(h) = \mathit{diet}$ and $s(h) = s'(h)$, otherwise. Also, t leaves the empty strategy unchanged: $t(s_\emptyset) = s_\emptyset$.

Base cases: Case $\underline{\varphi = p}$ is straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ does not contain cooperation modalities. As before, we first consider $A = \{1\}$. Suppose $M'_{11}, q_0, s' \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$, $s = t(s')$ and s_1 be a witnessing strategy of the former, i.e. $M'_{11}, \lambda[0], s_1 \dagger s' \models_{x,c} \gamma$ for all $\lambda \in \mathit{out}_{M_{11}}(q_0, s_1 \dagger s')$. We note that $s_1 \dagger s'$ is a $\{1\}$-strategy and that $\mathit{out}_{M'_{11}}(q_0, s')$ and $\mathit{out}_{M_{11}}(q_0, s)$ contain unique paths which are propositionally equivalent. For instance, if s' executes *diet* for some history yielding path $q_0^+ q_2^\omega$ then s will execute *nop* for that same history resulting in the path q_0^ω. Therefore we have $M_{11}, \lambda[0], s_1 \dagger s \models_{x,c} \gamma$ thus $M_{11}, q_0, s \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$. For $A = \emptyset$ we follow exactly the same argument.

Induction step: Cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi' \wedge \varphi''}$ are straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ contains cooperation modalities and $A = \{1\}$. Suppose $M'_{11}, q_0, s' \models_{x,c} \langle\!\langle 1 \rangle\!\rangle \gamma$ and let s_1 be a witnessing strategy. For each outermost $\mathsf{ATL}^\star$-subformula $\langle\!\langle B_i \rangle\!\rangle \psi_i$ in γ (with $i = 1, \ldots, k$), strategy s^* and state q such that $M_{11}, q, s^* \models_{x,c} \langle\!\langle B_i \rangle\!\rangle \psi_i$, we observe that s^* can only be of the form $s' \dagger s_1$. Hence $s^* = s_1$ since s_1 is irrevocable. Then $M'_{11}, q, s^* \models_{x,c} \langle\!\langle B_i \rangle\!\rangle \psi_i$ iff $M'_{11}, q, s_1 \models_{x,c} \psi_i$. Therefore we can treat φ exactly as in the base case. The same holds for $A = \emptyset$.

From $(\star), (\star\star)$ it follows that: $M_{11}, q_0, s_\emptyset \models_{x,c} \varphi$ iff $M'_{11}, q_0, s_\emptyset \models_{x,sc} \varphi$, which concludes this part of the proof.

Let $\varphi = \langle\!\langle 1 \rangle\!\rangle \Box (\mathsf{cake} \wedge \neg \langle\!\langle 1 \rangle\!\rangle \bigcirc \neg \mathsf{cake})$. The formula expresses that player 1 has a strategy to maintain cake, however given that strategy it is not possible for him/her to achieve $\neg$cake in the next state. We have $M_{11}, q_0 \not\models_{x,sc} \varphi$ but $M'_{11}, q_0 \models_{x,sc} \varphi$, which concludes the proof. $\Box$

PROPOSITION 7.15. *For all* M, λ, s *and every* ATL* *formula* φ*, we have that* $M, \lambda, 0, s \models^{nf}_{I,sc} \varphi$ *iff* $M, \lambda, s \models_{I,sc} \varphi$.

PROOF. For each history $\eta \in \Lambda^{fin}_M$, and strategy s, we define:

$$s^{+\eta}(h) = \begin{cases} s(last(\eta) \circ h') & \text{iff } h = \eta \circ h' \\ s(last(\eta)) & \text{iff } h = \eta \\ \text{undefined} & \text{otherwise} \end{cases} \qquad s^{-\eta}(h) = \begin{cases} s(\eta) & \text{iff } h = last(\eta) \\ s(\eta \circ h') & \text{iff } h = last(\eta) \circ h' \\ \text{undefined} & \text{otherwise} \end{cases}$$

Informally, $s^{+\eta}$ reproduces s starting from history $\eta[0, |\eta|-1]$, while $s^{-\eta}$ reproduces s as if history $\eta[0, |\eta|-1]$ did not occur. We recall from the proof of Proposition 3.6, that:

$$plays^I_M(h, s_B \dagger s_A) = out_M(h, s_B \dagger s_A) \tag{1}$$

for all $h \in \Lambda_M$ and all collective-$A \cup B$ strategies $s_B \dagger s_A$. Also for all collective strategies s_A and histories $h \in \Lambda_M$, there is a collective strategy s'_A such that:

$$out_M(h, s \dagger s_A) = \{h \circ \lambda \mid last(h) \circ \lambda \in out_M(last(h), s^{-h} \dagger s'_A)\} \tag{2}$$

and also, for each collective strategy s'_A there is a collective strategy s_A such that:

$$out_M(last(h), s \dagger s'_A) = \{last(h) \circ \lambda \mid h \circ \lambda \in out_M(h, s^{+h} \dagger s_A)\} \tag{3}$$

We show $(\star) M, h \circ \lambda, k, s \models^{nf}_{I,sc} \varphi$ iff $M, last(h) \circ \lambda, s^{-h} \models_{I,sc} \varphi$, for all paths $h \circ \lambda \in \Lambda_M$, such that $k = |h| - 1$, $|h| \geq 1$ and all scATL*-formulae φ. The proof is by induction over the formula structure of φ. The proposition follows from $(\star)$ for $k = 0$.

Base cases: The case $\underline{\varphi = p}$ is straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ does not contain cooperation modalities.

$M, h \circ \lambda, k, s \models^{nf}_{I,sc} \langle\!\langle A \rangle\!\rangle \gamma$ iff

— $\exists s_A$ such that $\forall h \circ \lambda' \in plays^I_M(h, s \dagger s_A)$ we have $M, h \circ \lambda', k, s \dagger s_A \models^{nf}_{I,sc} \gamma$ iff (by (1-3))
— $\exists s'_A$ such that $\forall last(h) \circ \lambda' \in plays^I_M(last(h), s^{-h} \dagger s'_A)$ we have $last(h) \circ \lambda', s^{-h} \dagger s_A \models_{I,sc} \gamma$ iff
— $M, last(h), s^{-h} \circ \lambda \models_{I,sc} \langle\!\langle A \rangle\!\rangle \gamma$.

Induction step: The cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi' \wedge \varphi''}$ are straightforward. Case $\underline{\varphi = \langle\!\langle A \rangle\!\rangle \gamma}$ where γ contains cooperation modalities.

We consider the two directions of $(\star)$:"$\Rightarrow$": Suppose $M, h \circ \lambda, k, s \models^{nf}_{I,sc} \langle\!\langle A \rangle\!\rangle \gamma$ and let s_A be a witnessing strategy, i.e. $M, h \circ \lambda', k, s \dagger s_A \models^{nf}_{I,sc} \gamma$ for all $h \circ \lambda' \in out_M(h, s \dagger s_A)$. Also, let $\langle\!\langle B_i \rangle\!\rangle \psi_i$with $i = 1, \ldots, k$ be an outermost ATL*-subformula in γ. As before, we define:

$$H_i = \{(h \circ \lambda, k) \mid M, h \circ \lambda, k, s \dagger s_A \models^{nf}_{I,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$$
$$L_i = \{\lambda \mid M, \lambda, s \models_{I,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$$

By induction hypothesis, $H_i = \{(h \circ \lambda, k) \mid last(h) \circ \lambda \in L_i\}$ (4). By (4) and (2) our claim follows immediately. Direction "$\Leftarrow$" follows exactly the same argument. $\Box$

PROPOSITION 7.18. *There are iCGSs which satisfy the same formulae of* $\mathsf{ATL}^\star_{\mathsf{i,sc}}$*, but can be distinguished in* $\mathsf{ATL}^\star_{\mathsf{nf,i,c}}$*. That is,* $\mathsf{ATL}^\star_{\mathsf{nf,i,c}} \not\preceq_d \mathsf{ATL}^\star_{\mathsf{i,sc}}$.

PROOF. The proof is similar to that of Proposition 4.4. First, we use the models M_2 and M_2' from Figure 4 to show that $M_2, a_0 \models_{i,sc} \varphi$ iff $M_2', a_0 \models_{i,sc} \varphi$. We start by observing that a strategy is uniform in M_2 iff is also uniform in M_2'. This holds since player 2 has a unique available action in a_0 (and b_0) respectively.

We define the strategy $t(s_A)$ with respect to s_A as follows:

$$t(s_A) = \begin{cases} s'_A & \text{if } A = \{1, 2\} \\ s_A & \text{otherwise} \end{cases}$$

where s'_A denotes the collective strategy whose construction was illustrated in the proof of Proposition 4.4. We observe that, s and $t(s)$ differ only in those actions assigned to histories prefixed by a_0. Thus:

$$M_2, x_j, s \models_{i,sc} \varphi \text{ iff } M_2, x'_j, t(s) \models_{i,sc} \varphi \tag{4}$$

for $x \in \{a, b\}$, $j \in \{1, 2\}$ and all strategies s (the same is the case for M_2')

$$M_2, x_j, s \models_{i,sc} \varphi \text{ iff } M_2', x_j, t(s) \models_{i,sc} \varphi \tag{5}$$

for $x \in \{a, b, a', b'\}$, $j \in \{1, 2\}$ and each strategy s.

We prove:

$(\star)$ $M_2, a_0, s_B \models_{i,sc} \varphi$ iff $M_2', a_0, t(s_B) \models_{i,sc} \varphi$, for all $B \in \{\emptyset, 1, 2, \{1, 2\}\}$

by induction over the formula structure of φ. Note that our claim follows for $B = \emptyset$, since $t(s_\emptyset) = s_\emptyset$. Base cases:The case $\varphi = p$ is straightforward. Case $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ contains no strategic modalities.

Suppose $A \cup B \in \{\emptyset, \{1\}\}$. The proposition follows immediately as $\mathit{plays}^i_{M_2}(a_0, s_B \dagger s_A) = \mathit{plays}^i_{M_2'}(a_0, s_B \dagger s_A)$.

Suppose $A \cup B = \{2\}$. We consider each direction of $(\star)$ in turn: "$\Leftarrow$": Since $\mathit{plays}^i_{M_2'}(a_0, s_B \dagger s_A) = \mathit{out}_{M_2'}(a_0, s_B \dagger s_A) \cup \mathit{out}_{M_2'}(b_0, s_B \dagger s_A)$ and $\mathit{plays}^i_{M_2}(a_0, s_B \dagger s_A) = \mathit{out}_{M_2}(a_0, s_B \dagger s_A)$, we have $\mathit{plays}^i_{M_2}(a_0, s_B \dagger s_A) \subseteq \mathit{plays}^i_{M_2'}(a_0, s_B \dagger s_A)$. $(\star)$ follows immediately.

Direction "$\Rightarrow$" follows the very same reasoning as that from the proof of Proposition 4.4, where s_2 is replaced by $s_{A \cup B}$. Note that $s_{A\cup B} = t(s_{A\cup B})$, since $A \cup B = \{2\}$.

Suppose $A \cup B = \{1, 2\}$. Direction "$\Leftarrow$ ": we note $\mathit{plays}^i_{M_2}(a_0, s_B \dagger s_A) \subseteq \mathit{plays}^i_{M_2'}(a_0, s_B \dagger s_A)$.

Direction "$\Rightarrow$": the set $\mathit{plays}^i_{M_2}(a_0, s_B \dagger s_A)$ contains a unique path, while $\mathit{plays}^i_{M_2}(a_0, t(s_B \dagger s_A))$ contains two paths. By the construction of t, the latter two paths are propositionally equivalent to the former.

Induction step: The cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\langle A \rangle\rangle \gamma$ where γ contains cooperation modalities. Let ξ be an arbitrary occurrence of an outermost formula $\langle\langle A' \rangle\rangle \gamma'$ in γ. We note that $M_2, x, s \models_{i,sc} \xi$ iff $M_2', x, t(s) \models_{i,sc} \xi$ by induction hypothesis if $x = a_0$ and by (1-2), otherwise. We label each state x of M_2 and M_2' where ξ holds by a new proposition p_ξ. The resulting models retain properties (1-2). We replace each ξ by p_ξ in γ and obtain a formula without cooperation modalities. We proceed as in the second base case. This concludes the proof of $(\star)$.

Second, we show that the models from Figure 4 can be distinguished by an $\mathsf{ATL}^\star_{\mathsf{nf,i,c}}$-formula. Recall that $\mathsf{E}\varphi \equiv \neg\langle\langle\emptyset\rangle\rangle\neg\varphi$. Then, we have $M_2, a_0 \models^{nf}_{i,sc} \mathsf{E}\bigcirc\langle\langle 2 \rangle\rangle\bigcirc\mathsf{win}$ but $M_2', a_0' \not\models^{nf}_{i,sc} \mathsf{E}\bigcirc\langle\langle 2 \rangle\rangle\bigcirc\mathsf{win}$: in M_2, there is a path on which player 2 can ensure win by itself in the next state. The latter does not hold in M_2'. $\square$

PROPOSITION 7.19. *There are iCGSs which satisfy the same formulae of* $\mathsf{ATL}^\star_{\mathsf{nf,i,sc}}$*, but can be distinguished in* $\mathsf{ATL}^\star_{\mathsf{i,c}}$*. That is,* $\mathsf{ATL}^\star_{\mathsf{i,c}} \not\preceq_d \mathsf{scATL}^\star_{\mathsf{nf,i}}$.

PROOF. The proof is similar to that of Proposition 4.6. First, we use models M_3 and M'_3 from Figure 5 to show ($\star$) $M_3, h, s \models^{nf}_{i,sc} \varphi$ iff $M'_3, h, t(s) \models^{nf}_{i,sc} \varphi$ for all histories $h \in \Lambda^{fin}_{M_3}(a_0)$ and all $\mathsf{ATL}^\star$-formulae φ. We start by defining the strategy $t(s_A)$ with respect to s_A, as follows:

$$t(s_A) = \begin{cases} s'_A & \text{if } A = \{2\} \\ s_A & \text{otherwise} \end{cases}$$

where $s'_A(b_1) = s_A(a_1)$ and $s'_A(h) = s_A(h)$ for all $h \in \Lambda^{fin}_{M_3}$. Note that, for all strategies s, $t(s)$ is a uniform strategy in M'_3.

The proof is by induction over the formula structure of φ.

Base cases: Case $\underline{\varphi = p}$ is straightforward. Case $\underline{\varphi = \langle\langle A\rangle\rangle\gamma}$ where γ contains no strategic modalities. It is sufficient to observe:

$$\mathit{plays}^i_{M_3}(h, s \dagger s_A) = \mathit{plays}^i_{M'_3}(h, t(s) \dagger t(s_A)) \tag{1}$$

for all collective strategies s_A and $h \in \Lambda^{fin}_{M_3}(a_0)$. We note that s and $t(s)$ produce the same effects when the initial state is a_0, for all uniform strategies s. However, we use $t(s)$ instead of s since the latter may not be uniform in M'_3 — the transformation t enforces uniformity by ensuring that agents play the same action in indistinguishable states a_1 and b_1 of M'_3.

Cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi' \wedge \varphi''}$ are straightforward. Case $\underline{\varphi = \langle\langle A\rangle\rangle\gamma}$ where γ contains cooperation modalities. Suppose $M_3, h, s \models^{nf}_{i,sc} \langle\langle A\rangle\rangle\gamma$ and let s_A be a witnessing strategy. Let $\langle\langle B_i\rangle\rangle\psi_i$ (with $i = 1 \dots k$) be an outermost $\mathsf{ATL}^\star$-subformula in γ. By induction hypothesis, we have: $M_3, h, s \dagger s_A \models^{nf}_{i,sc} \langle\langle B_i\rangle\rangle\psi_i$ iff $M'_3, h, t(s \dagger s_A), \models^{nf}_{i,sc} \langle\langle B_i\rangle\rangle\psi_i$. Moreover, $t(s \dagger s_A) = t(s) \dagger t(s_A)$ — applying the uniformity constraint on $s \dagger s_A$ is equivalent to applying it on s and s_A individually.

It follows by (1) that s_A is a witnessing strategy for $M_3, h \models^{nf}_i \langle\langle A\rangle\rangle\gamma$ iff $t(s_A)$ is a witnessing strategy for $M'_3, h \models^{nf}_i \langle\langle A\rangle\rangle\gamma$.

Finally, to show that M_3 and M'_3 can be distinguished by an $\mathsf{ATL}^\star_{\mathsf{i,c}}$-formula we consider the formula $\varphi \equiv \mathsf{E}\bigcirc\langle\langle 2\rangle\rangle\bigcirc\,\mathsf{win}$ from Proposition 7.18. We have that $M_3, a_0 \models \varphi$ and $M'_3, a'_0 \not\models \varphi$. $\square$

PROPOSITION 7.22. *$Val(\mathsf{scATL}^\star_{\mathsf{i}}) \subseteq Val(\mathsf{scATL}^\star_{\mathsf{nf,i}})$. Thus, also $Val(\mathsf{ATL}^\star_{\mathsf{i,c}}) \subseteq Val(\mathsf{ATL}^\star_{\mathsf{nf,i,c}})$.*

PROOF. We proceed similarly to the proof of Proposition 5.5. First, we show:

$$M, q, s \models^{nf}_{i,sc} \varphi \text{ iff } T^{nf}(M, q), q, t(s) \models^{nf}_{i,sc} \varphi \text{ iff } T^{nf}(M, q), q, t(s) \models_{i,sc} \varphi$$

for all $\mathsf{ATL}^\star$-formulae, iCGSs M, states q, strategies s which are uniform w.r.t. $\mathbb{A}\mathrm{gt}$ and where $t(s)$ is the strategy s' in $T^{nf}(M, q)$, such that $s'(\hat{h}) = s(h)$, for all $h \in \Lambda^{fin}_M(q)$ and $t(s_\emptyset) = s_\emptyset$. We first observe that $t(s \dagger s') = t(s) \dagger t(s')$. The proof also relies on the observations (1-3) from the proof of Proposition 5.5. We prove both equivalences separately.

(i). $M, q, s \models^{nf}_i \varphi$ iff $T^{nf}, q, t(s) \models^{nf}_i \varphi$. We prove the stronger statement: $M, h \circ \lambda, k, s \models^{nf}_i \varphi$ iff $T^{nf}, \hat{h} \circ \hat{\lambda}, k, t(s) \models^{nf}_i \varphi$, where $k = |h| - 1$. Our claim follows for $k = 0$. The proof is by induction over the formula structure of φ.

Base cases: Case $\underline{\varphi = p}$. It is sufficient to note that $(h \circ \lambda)[k] \in \pi^M(p)$ iff $(\hat{h} \circ \hat{\lambda})[k] \in \pi^{T^{nf}}(p)$. Case $\underline{\varphi = \langle\langle A\rangle\rangle\gamma}$ where γ contains no cooperation modalities. From (2-3) it follows that $\lambda \in \mathit{plays}^i_M(h, s \dagger s_A)$ iff $\hat{\lambda} \in \mathit{plays}^i_{T^{nf}}(\hat{h}, t(s) \dagger t(s_A))$. Since λ and $\hat{\lambda}$ are propositionally equivalent cf. (1), we have: $M, h \circ \lambda, s \models^{nf}_i \langle\langle A\rangle\rangle\gamma$ iff $T^{nf}, \hat{h} \circ \hat{\lambda}, t(s) \models^{nf}_i \langle\langle A\rangle\rangle\gamma$.

Induction step: Cases $\underline{\varphi = \neg\varphi'}$ and $\underline{\varphi = \varphi' \wedge \varphi''}$ are straightforward. Case $\underline{\varphi = \langle\langle A\rangle\rangle\gamma}$ where γ contains cooperation modalities. We consider the two directions of (i) in turn: "$\Rightarrow$": Suppose

$M, h \circ \lambda, k, s \models^{nf}_{i,sc} \langle\!\langle A \rangle\!\rangle \gamma$ and let s_A be a witnessing strategy i.e. $M, h \circ \lambda', k, s \dagger s_A \models^{nf}_{i,sc} \gamma$ for all paths $h \circ \lambda' \in plays^i_M(h, s \dagger s_A)$. Also, let $\langle\!\langle B_i \rangle\!\rangle \psi_i$ with $i = 1, \ldots, k$ be an outermost $\mathsf{ATL}^\star$-subformula in γ and define the sets:

$H_i = \{(h \circ \lambda, k) \mid M, h \circ \lambda, k, s \dagger s_A \models^{nf}_{i,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$

$\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid T^{nf}, \hat{h} \circ \hat{\lambda}, k, t(s) \dagger t(s_A) \models^{nf}_{i,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$

By induction hypothesis, $\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid (h \circ \lambda, k) \in H_i\}$ (4). Also, from (2-3) we have:

$$\{\hat{\lambda} \mid \hat{\lambda} \in plays^i_M(h, s \dagger s_A)\} = \{\lambda \mid \lambda \in plays^i_{T^{nf}}(\hat{h}, t(s) \dagger t(s_A))\}$$

The claim now follows by (4). Direction ”$\Leftarrow$“ follows exactly the same argument.

(ii). $T^{nf}, q, s \models^{nf}_i \varphi$ iff $T^{nf}, q, s \models_i \varphi$. As before, we prove the stronger claim $(\star)$ $T^{nf}, \hat{h} \circ \hat{\lambda}, k, s \models^{nf}_i \varphi$ iff $T^{nf}, last(\hat{h}) \circ \hat{\lambda}, s \models_i \varphi$ where $k = |\hat{h}| - 1$, by induction over the formula structure of φ. Then (ii) follows for $k = 0$.

Base cases: Case $\varphi = p$ is straightforward. Case $\varphi = \langle\!\langle A \rangle\!\rangle \gamma$ where γ contains no cooperation modalities. From (1-3) we observe that for each uniform collective strategy s_A and history $\hat{h} \in \Lambda^{fin}_{T^{nf}}$, there exists a uniform collective strategy s'_A such that:

$$plays^i_{T^{nf}}(\hat{h}, s \dagger s_A) = \{\hat{h} \circ \hat{\lambda} \mid last(\hat{h}) \circ \hat{\lambda} \in plays^i_{T^{nf}}(last(\hat{h}), s \dagger s'_A)\}$$

Informally, each state $last(\hat{h})$ encodes a history. Hence, the action of coalition A assigned by a perfect-recall strategy to $\hat{h}$ can be executed with the same effects from $last(\hat{h})$ alone. Similarly, for each collective strategy s'_A and history $\hat{h} \in \Lambda^{fin}_{T^{nf}}$ there exists a collective strategy s_A such that:

$$plays^i_{T^{nf}}(last(\hat{h}), s \dagger s'_A) = \{last(\hat{h}) \circ \hat{\lambda} \mid \hat{h} \circ \hat{\lambda} \in plays^i_{T^{nf}}(\hat{h}, s \dagger s_A)\}$$

We conclude that $T^{nf}, \hat{h} \circ \hat{\lambda}, k, s \models^{nf}_i \langle\!\langle A \rangle\!\rangle \gamma$ iff $T^{nf}, last(\hat{h}) \circ \hat{\lambda}, s \models_i \langle\!\langle A \rangle\!\rangle \gamma$.

Induction step: Cases $\varphi = \neg\varphi'$ and $\varphi = \varphi' \wedge \varphi''$ are straightforward. Case $\varphi = \langle\!\langle A \rangle\!\rangle \gamma$ where γ contains cooperation modalities.

Direction “$\Rightarrow$”: Supose $T^{nf}, \hat{h} \circ \hat{\lambda}, k, s \models^{nf}_{i,sc} \langle\!\langle A \rangle\!\rangle \gamma$ and let s_A be a witnessing strategy i.e. $T^{nf}, \hat{h} \circ \hat{\lambda}', k, s \dagger s_A \models^{nf}_{i,sc} \gamma$ for all paths $\hat{h} \circ \hat{\lambda}' \in plays^i_{T^{nf}}(\hat{h}, s \dagger s_A)$. Also, let $\langle\!\langle B_i \rangle\!\rangle \psi_i$ with $i = 1, \ldots, k$ be an outermost $\mathsf{ATL}^\star$-subformula in γ and define:

$\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid T^{nf}, \hat{h} \circ \hat{\lambda}, k, s \dagger s_A \models^{nf}_{i,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$

$\hat{L}_i = \{\hat{\lambda} \mid T^{nf}, \hat{\lambda}, s \dagger s_A \models_{i,sc} \langle\!\langle B_i \rangle\!\rangle \psi_i\}$

By induction hypothesis: $\hat{H}_i = \{(\hat{h} \circ \hat{\lambda}, k) \mid last(\hat{h}) \circ \hat{\lambda} \in \hat{L}_i\}$ (5). Finally, by (2,3) we have:

$$plays^i_{T^{nf}}(s \dagger s_A, \hat{h}) = \{\hat{h} \circ \hat{\lambda} \mid last(\hat{h}) \circ \hat{\lambda} \in plays^i_{T^{nf}}(t(s \dagger s_A), last(\hat{h}))\}$$

The claim now follows by (5). Direction “$\Leftarrow$” follows exactly the same argument.

We show $Sat(\mathsf{scATL}^\star_{\mathsf{nf,i}}) \subseteq Sat(\mathsf{ATL}^\star_{\mathsf{i,sc}})$. Suppose $\varphi \in Sat(\mathsf{scATL}^\star_{\mathsf{nf,i}})$. Thus there exists an iCGS M and state q such that $M, q, s_\emptyset \models^{nf}_{i,sc} \varphi$. By (i, ii) and since $t(s_\emptyset) = s_\emptyset$, $T^{nf}(M, q), s_\emptyset \models_i \varphi$. Hence $\varphi \in Sat(\mathsf{ATL}^\star_{\mathsf{i,sc}})$. $\square$