Making Agents' Abilities Explicit

Abstract

Alternating-time temporal logics (ATL/ATL*) represent a family of modal logics for reasoning about agents' strategic abilities in multiagent systems. The interpretations of ATL/ATL* over the semantic model Concurrent Game Structures (CGS) usually vary depending on the agents' abilities, for instance, perfect vs. imperfect information, perfect vs. imperfect recall, resulting in a variety of variants which have been studied extensively in literature. However, they are defined at the semantic level, which may limit modeling flexibilities and may give counter-intuitive interpretations. To mitigate these issues, in this work, we propose to extend CGS with agents' abilities and study the new semantics of ATL/ATL* under this model. We give PSPACE/2EXPTIME model-checking algorithms for ATL/ATL* and implement them as a prototype tool. Experiment results show the practical feasibility of the approach.

1 Introduction

Multiagent systems (MAS) comprising multiple autonomous agents have become a widely adopted paradigm of intelligent systems. Game-based models and associated logics, as the foundation of MAS, have received tremendous attentions in recent years. The seminar work [Alur *et al.*, 2002] proposed *concurrent game structures* (CGS) as the model of MAS and alternating-time temporal logics (typically ATL and ATL*) as specification languages for expressing temporal goals. In a nutshell, a CGS consists of multiple players which are used to represent autonomous agents, components and the environment. The model describes how the MAS evolves according to the collective behavior of agents. ATL/ATL*, an extension of the Computational Tree Logic (CTL/CTL*), features coalition modality $\langle\langle A\rangle\rangle$. The formula $\langle\langle A\rangle\rangle\varphi$ expresses the property that the coalition (i.e. the agent group) A has a collective strategy to achieve a certain goal specified by φ.

A series of extensions of ATL-like logics have been studied which take different agents' abilities into account. These abilities typically include whether agents can identify the current state of the system completely or only partially (perfect vs. imperfect information), and whether agents can memorize the whole history of observations or simply part of them (perfect vs. imperfect recall). Different abilities usually induce distinct semantics, which are indeed necessary because of the versatility of problem domains. These semantic variants and their model-checking problems comprise subjects of active research for almost two decades, to cite a few [Schobbens, 2004; Jamroga and van der Hoek, 2004; Ågotnes *et al.*, 2007; Dima and Tiplea, 2011; Bulling and Jamroga, 2011; Laroussinie and Markey, 2015].

While agents' abilities play a prominent role [Bulling and Jamroga, 2014], the semantics of ATL-like logics only refers to them *implicitly*. In other words, the logic per se does not specify what ability an agent has; instead one could infer the ability an agent requires by examining the goal specified in the logic. This approach, being elegant and valuable to understand the relationship between different abilities, suffers from a few shortcomings: (1) From the modeling perspective, it is common in practice that agents in an MAS vary in their abilities (for instance, agents modeling sensors may not identify the complete state of system so can only use strategies with imperfect information). When constructing a model, these abilities ought to be encoded explicitly. Such modeling flexibility is not supported by the existing formalisms. (2) From the semantic perspective, ATL-like logics may exhibit some counter-intuitive semantics. The core modality of ATL, $\langle\langle A\rangle\rangle\varphi$, is interpreted as that the coalition A has a collective strategy to achieve the goal φ "no matter what the other agents do" rather than "no matter which strategies the other agents choose". The delicate difference suggests that the (multi-player) game nature in the evolution of MAS is not fully captured by ATL. For instance, in the imperfect information/recall setting, only agents in A are assumed to use imperfect information/recall strategies, while the agents *not* in A may still use perfect information and perfect recall strategies. Even worse, if the coalition modalities are nested, the same agent may have different abilities to fulfill the objectives specified in different sub-formulae, resulting in inconsistency in the strategies it uses. This phenomenon has also been mentioned in [Mogavero *et al.*, 2014; Cermák *et al.*, 2014] which proposed a strategic logic making explicit references to strategies of all agents (including those not in A), though all agents should have same abilities therein.

To summarize, it occurs to us that the current approach in which the temporal formulae are with implicit agents' abilities at the semantic level impedes necessary modeling flexibilities and often yields unpleasant (even weird) semantics. Instead, we argue that coupling agents' abilities at the syntactic level of system models would deliver a potentially better approach to overcome the aforementioned limitations. Bearing the rationale in mind, we propose a new MAS model, i.e., *Agents' Abilities Augmented Concurrent Game Structures* (ACGS), which encompass agents' abilities explicitly.

We investigate ATL/ATL* over ACGS. We show that in general the new semantics of ATL/ATL* over ACGS is incomparable with others even if the underlying CGS models are the same. We also study the model-checking problem of ATL/ATL* over ACGS. We show that this problem is generally undecidable. However, we manage to show that the model-checking problem for ATL* (resp. ATL) on ACGS is 2EXPTIME-complete (resp. in PSPACE) when the imperfect information and perfect recall strategies are disallowed. We implement our algorithms in a prototype tool and conduct experiments on some standard applications from the literature. The results confirm the feasibility of our approach.

Some further experiments, proofs and comparison of ATL/ATL* semantics between CGS and ACGS are given in the appendix.

2 Concurrent Game Structures

Given an infinite word $\rho = s_0 s_1 \ldots$, we denote the symbol s_j by ρ_j, the prefix $s_0 \ldots s_j$ by $\rho_{[0..j]}$, and the suffix $s_j s_{j+1} \ldots$ by $\rho_{[j..\infty]}$. For a finite word $\rho = s_0 s_1 \cdots s_m$, we further denote the (last) symbol s_m by $\mathtt{lst}(\rho)$.

Let AP denote a finite set of atomic propositions. A *concurrent game structure* (CGS) is a tuple

$$\mathcal{G} \triangleq (S, S_0, Ag, (Ac_i)_{i\in Ag}, (\sim_i)_{i\in Ag}, (P_i)_{i\in Ag}, \Delta, \lambda),$$

where S is a finite set of *states*; $S_0 \subseteq S$ is a set of *initial states*; $Ag = \{1, ..., n\}$ is a finite set of *agents*; Ac_i is a finite set of *local actions* of agent i; $\sim_i \subseteq S \times S$ is an *epistemic accessibility relation* (an equivalence relation); $P_i : S \to 2^{Ac_i}$ is a *protocol function* such that $P_i(s) = P_i(s')$ for every $s \sim_i s'$; $\Delta : S \times Ac \to S$ is a transition function with $Ac = \prod_{i\in Ag} Ac_i$ being a set of joint actions; and $\lambda : S \to 2^{AP}$ is a *labeling function* which assigns each state with a set of atomic propositions. Given a joint action $\vec{a} = \langle a_1, ..., a_n\rangle \in Ac$, we use $\vec{a}(i)$ to denote the local action of agent i in $\vec{a}$.

A *path* is an infinite sequence of states $\rho = s_0 s_1 \ldots$ such that for every $j \geq 0$, $s_{j+1} = \Delta(s_j, \vec{a}_j)$ for some $\vec{a}_j \in \prod_{i\in Ag} P_i(s_j)$. Two sequences $\rho = s_0 \ldots s_m \in S^+$ and $\rho' = s'_0 \ldots s'_m \in S^+$ are *indistinguishable* for agent i, denoted by $\rho \sim_i \rho'$, if for every $j : 0 \leq j \leq m$, $s_j \sim_i s'_j$.

Strategies. Typical agents' abilities are captured by the following types of strategies [Schobbens, 2004]. For $i \in Ag$,

Ir-strategy $\theta_i : S \to Ac$ with $\forall s \in S$, $\theta_i(s) \in P_i(s)$;

IR-strategy $\theta_i : S^+ \to Ac$ with $\forall \rho \in S^+$, $\theta_i(\rho) \in P_i(\mathtt{lst}(\rho))$;

ir-strategy $\theta_i : S \to Ac$, the same as the Ir-strategy but with the additional constraint $s \sim_i s' \Rightarrow \theta_i(s) = \theta_i(s')$;

iR-strategy $\theta_i : S^+ \to Ac$, the same as the IR-strategy but with the additional constraint $\rho \sim_i \rho' \Rightarrow \theta_i(\rho) = \theta_i(\rho')$.

Intuitively, $\mathtt{i}$ (resp. $\mathtt{I}$) signifies that agents can only observe partial information characterized via epistemic accessibility relations (resp. complete information with all epistemic accessibility relations being the identity relation), while $\mathtt{r}$ (resp. $\mathtt{R}$) signifies that agents can make decision based on the current observation (resp. the whole history of observations). We will, by slightly abusing notation, extend Ir-strategies and ir-strategies θ_i to the domain S^+ such that for all $\rho \in S^+$, $\theta_i(\rho) = \theta_i(\mathtt{lst}(\rho))$. We denote by Θ_i^σ for $\sigma \in \{\mathtt{Ir}, \mathtt{IR}, \mathtt{ir}, \mathtt{iR}\}$ the set of σ-strategies for agent i.

Outcomes. A *collective σ-strategy* of a set of agents A is a function υ_A^σ assigning each agent $i \in A$ with a σ-strategy $\upsilon_A^\sigma(i) \in \Theta_i^\sigma$. For $i \in A$ and $\rho \in S^+$, we denote the local action $\upsilon_A^\sigma(i)(\rho)$ of agent i by $\upsilon_A^\sigma(i,\rho)$, and the set $Ag \setminus A$ by $\overline{A}$.

Given a state s, two collective σ/σ'-strategies υ_A^σ and $\upsilon_{\overline{A}}^{\sigma'}$ yield a path ρ, denoted by $\mathtt{play}(s, \upsilon_A^\sigma, \upsilon_{\overline{A}}^{\sigma'})$, where $\rho_0 = s$ and for every $j \geq 0$, $\rho_{j+1} = \Delta(\rho_j, \vec{a}_j)$, $\vec{a}_j(i) = \upsilon_A^\sigma(i, \rho_{[0..j]})$ for $i \in A$ and $\vec{a}_j(i) = \upsilon_{\overline{A}}^{\sigma'}(i, \rho_{[0..j]})$ for $i \in \overline{A}$.

For every state $s \in S$ and collective σ-strategy υ_A^σ of A, the *outcome* function is defined as follows:

$$O_{\mathcal{G}}^\sigma(s, \upsilon_A^\sigma) \triangleq \left\{\mathtt{play}(s, \upsilon_A^\sigma, \upsilon_{\overline{A}}^{\mathtt{IR}}) \mid \forall i \in \overline{A}, \upsilon_{\overline{A}}^{\mathtt{IR}}(i) \in \Theta_i^{\mathtt{IR}}\right\},$$

i.e., the set of all possible plays that may occur when each agent $i \in A$ enforces its σ-strategy $\upsilon_A^\sigma(i)$ from the state s. The subscript $\mathcal{G}$ is dropped in $O_{\mathcal{G}}^\sigma$ when it is clear from the context.

3 Alternating-Time Temporal Logics

The alternating-time temporal logic ATL/ATL* is an extension of the branching-time logic CTL/CTL* by replacing the existential path quantifiers $\mathbf{E}$ with coalition modalities $\langle\langle A\rangle\rangle$ [Alur *et al.*, 2002]. Intuitively, the formula $\langle\langle A\rangle\rangle\phi$ expresses that the set of agents A has a collective strategy to achieve the goal ϕ no matter which strategies the agents in $\overline{A}$ choose. Formally, ATL* is defined by the following grammar:

$$\varphi ::= q \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle\phi$$
$$\phi ::= \varphi \mid \neg\phi \mid \phi \wedge \phi \mid \mathbf{X}\,\phi \mid \phi\,\mathbf{U}\,\phi$$

where φ (resp. ϕ) denotes state (resp. path) formulae, $q \in AP$ and $A \subseteq Ag$.

The derived operators are defined as usual: $\phi_1 \to \phi_2 \triangleq \phi_2 \vee \neg\phi_1$, $\mathbf{F}\,\phi \triangleq \mathtt{true}\,\mathbf{U}\,\phi$, $\mathbf{G}\,\phi \triangleq \neg\mathbf{F}\,\neg\phi$, $\phi_1\,\mathbf{R}\,\phi_2 \triangleq \mathbf{G}\phi_2 \vee \phi_2\mathbf{U}(\phi_1 \wedge \phi_2)$, and $[[A]]\phi \triangleq \neg\langle\langle A\rangle\rangle\neg\phi$. An LTL formula is an ATL* path formula by restricting φ to atomic propositions.

The semantics of ATL* is traditionally defined over CGS. When agents's abilities are considered, it is often parameterized with a strategy type $\sigma \in \mathrm{T}_{\mathtt{str}}$, denoted by ATL_σ^* [Bulling and Jamroga, 2014]. Formally, let $\mathcal{G}$ be a CGS and s be a state of $\mathcal{G}$, the semantics of ATL_σ^* (i.e. the satisfaction relation) is defined inductively as follows: (where ϱ is a state s or path ρ)

- $\mathcal{G}, s \models_\sigma q$ iff $q \in \lambda(s)$;
- $\mathcal{G}, s \models_\sigma \langle\langle A\rangle\rangle\phi$ iff there exists a collective σ-strategy υ_A^σ of agents A such that $\forall\rho \in O^\sigma(s, \upsilon_A^\sigma)$, $\mathcal{G}, \rho \models_\sigma \phi$;
- $\mathcal{G}, \rho \models_\sigma \varphi$ iff $\mathcal{G}, \rho_0 \models_\sigma \varphi$;
- $\mathcal{G}, \rho \models_\sigma \mathbf{X}\phi$ iff $\mathcal{G}, \rho_{[1..\infty]} \models_\sigma \phi$;

- $\mathcal{G}, \rho \models_\sigma \phi_1 \mathbf{U} \phi_2$ iff $\exists k \geq 0$ such that $\mathcal{G}, \rho_{[k..\infty]} \models \phi_2$ and $\forall j : 0 \leq j < k, \mathcal{G}, \rho_{[j..\infty]} \models_\sigma \phi_1$;
- $\mathcal{G}, \varrho \models_\sigma \phi_1 \wedge \phi_2$ iff $\mathcal{G}, \varrho \models_\sigma \phi_1$ and $\mathcal{G}, \varrho \models_\sigma \phi_2$;
- $\mathcal{G}, \varrho \models_\sigma \neg\phi$ iff $\mathcal{G}, \varrho \not\models_\sigma \phi$.

Vanilla ATL. ATL is a sublogic of ATL* where each occurrence of the coalition modality $\langle\langle A \rangle\rangle$ is immediately followed by a temporal operator. Formally, ATL is defined by the following grammar:
$\varphi ::= q \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A \rangle\rangle \mathbf{X} \varphi \mid \langle\langle A \rangle\rangle [\varphi \mathbf{R} \varphi] \mid \langle\langle A \rangle\rangle [\varphi \mathbf{U} \varphi]$ where $q \in \mathcal{A}$ and $A \subseteq Ag$.

Remark that the operator $\mathbf{R}$ cannot be defined using other operators in ATL with imperfect information [Laroussinie *et al.*, 2008], so is included for completeness.

Given an ATL* formula φ, a CGS $\mathcal{G}$ and a strategy type $\sigma \in \mathrm{T}_{\mathtt{str}}$, the *model-checking problem* is to determine whether $\mathcal{G}, s \models_\sigma \varphi$ or not, for each initial state s of the CGS $\mathcal{G}$.

Some Semantic Issues. We observe that the semantics of ATL/ATL* refers to the agents' abilities in an implicit manner. For the formula $\langle\langle A \rangle\rangle\varphi$, the specified σ-strategies only apply to agents in A while the agents in $\overline{A}$ could still choose beyond σ-strategies (e.g. $\mathtt{IR}$-strategies). In other words, A has a collective σ-strategy to achieve φ no matter what the other agents do. When σ is $\mathtt{IR}$ as in the original work by [Alur *et al.*, 2002], this interpretation of $\langle\langle A \rangle\rangle\varphi$ is plausible, as "no matter what the other agents do" is effectively the same as "no matter which strategies the other agents choose". However, when σ is set to be more restricted than $\mathtt{IR}$, agents not in A are still allowed to use $\mathtt{IR}$-strategies.

As mentioned in the introduction, this results in a few shortcomings. From a modeling perspective, arguably agents' abilities should be decided by the practical scenario. Namely, they should be fixed when the model is built, and all agents stick to their respective abilities independent of logic formulae. More concretely, from the semantic perspective, the existing semantics does not take into account the abilities of agents who are not in A, and neglects the (multi-player) game nature in the evolution of MAS. As a result, it may exhibit some counter-intuitive semantics. For instance, consider two formulae $\langle\langle A \rangle\rangle\phi$ and $\langle\langle A' \rangle\rangle\phi'$ such that agent $i \in A \setminus A'$, i may have different abilities to achieve ϕ and ϕ'. Let us consider an autonomous road vehicle scenario to see why this is not ideal. There are several autonomous cars which can only observe partial information and have bounded memory. An MAS model $\mathcal{G}$ consists of agents in A modeling autonomous cars, and an additional environment agent e. We can reasonably assume that all the car agents use $\mathtt{ir}$-strategies, while e uses $\mathtt{IR}$-strategies. The property $\langle\langle A' \rangle\rangle\phi$ expresses that autonomous cars in A' can cooperatively achieve the goal ϕ no matter what strategies the other cars and the environment choose. Verifying that $\mathcal{G}$ satisfies $\langle\langle A' \rangle\rangle\phi$ under the existing semantics would allow car agents in $A \setminus A'$ to use $\mathtt{IR}$-strategies. If $\mathcal{G}$ satisfies $\langle\langle A' \rangle\rangle\phi$, then the result is conclusive, i.e., $\langle\langle A' \rangle\rangle\phi$ holds for the system. However, if $\mathcal{G}$ invalidates $\langle\langle A' \rangle\rangle\phi$, we *cannot* deduce that $\langle\langle A' \rangle\rangle\phi$ fails because we overestimate the abilities of agents in $A \setminus A'$ when evaluating $\langle\langle A' \rangle\rangle\phi$. In other words, for the formula $\langle\langle A \rangle\rangle\varphi$ under $\models_\sigma$ where $\sigma \neq \mathtt{IR}$, it seems to be inappropriate to render the agents not in A extra powers of $\mathtt{IR}$ to potentially defeat agents from A when their abilities are actually much weaker.

4 Agents' Abilities Augmented CGS

In this section, we introduce *agents' abilities augmented concurrent game structures* (ACGS in short), which explicitly equip each agent with a strategy type from $\mathrm{T}_{\mathtt{str}}$. As such, agents have fixed abilities throughout their lives for a given CGS. Formally, an ACGS is a pair $\mathcal{M} \triangleq (\mathcal{G}, \pi)$, where $\mathcal{G}$ is a CGS and $\pi : Ag \to \mathrm{T}_{\mathtt{str}}$ is a function that assigns a strategy type $\pi(i)$ to the agent i. We assume that, for each agent $i \in Ag$ with $\pi(i) \in \{\mathtt{IR}, \mathtt{Ir}\}$, $\sim_i$ is an identity relation, as agents with perfect information should be able to distinguish two distinct states. Paths of $\mathcal{M}$ are defined the same as for the CGS $\mathcal{G}$, but strategies and outcomes of $\mathcal{M}$ have to be redefined as follows.

Strategies and Outcomes. Let A be a set of agents. A *collective strategy* of A in $\mathcal{M}$ is a function ξ_A that assigns each agent $i \in A$ with a $\pi(i)$-strategy $\xi_A(i) \in \Theta_i^{\pi(i)}$. Given a state $s \in S$ and a collective strategy ξ_A of the set A of agents, the outcome $O_{\mathcal{M}}(s, \xi_A)$ of $\mathcal{M}$ is the set of all possible paths that may occur when each agent $i \in A$ enforces its $\pi(i)$-strategy $\xi_A(i)$ from state s, and other agents $i \in \overline{A}$ can only choose $\pi(i)$-strategies instead of general $\mathtt{IR}$-strategies. Formally, $O_{\mathcal{M}}(s, \xi_A)$ is defined as

$$O_{\mathcal{M}}(s, \xi_A) \triangleq \{\mathtt{play}(s', \xi_A, \xi_{\overline{A}}) \mid \forall i \in \overline{A}, \xi_{\overline{A}}(i) \in \Theta_i^{\pi(i)}\}$$

We will omit $\mathcal{M}$ from $O_{\mathcal{M}}(s, \xi_A)$ when it is clear from context.

Semantics of ATL/ATL*. The difference of outcomes between ACGS and CGS induces a distinct semantics of ARL/ATL* on ACGS than CGS. Let $\mathcal{M}$ be an ACGS and s be a state in $\mathcal{M}$, the semantics of ATL/ATL* on $\mathcal{M}$ is defined similar to the one on CGS, except that

$\mathcal{M}, s \models \langle\langle A \rangle\rangle\phi$ iff there exists a collective strategy ξ_A of A such that $\mathcal{M}, \rho \models \phi$ for all $\rho \in O(s, \xi_A)$.

Remark that this semantics takes into account whether the agents from $\overline{A}$ have perfect or imperfect information/recall.

Given an ACGS $\mathcal{M}$ and an ATL/ATL* formula φ, the *model-checking problem* is to determine whether $\mathcal{M}, s \models \varphi$ holds, for every initial state s of G. Given a state formula φ, let $[\![\varphi]\!]_{\mathcal{M}}$ denote the set of all the states of $\mathcal{M}$ that satisfies φ.

The semantics of ATL/ATL* defined on ACGS is different from the one defined on CGS, hence they are incomparable.

Proposition 1 *There are an ACGS $\mathcal{M} = (\mathcal{G}, \pi)$, an ATL/ATL* formula $\langle\langle A \rangle\rangle\phi$, and a type $\sigma \in \mathrm{T}_{\mathtt{str}}$ such that $\pi(i) = \sigma$ for all $i \in A$ and $\mathcal{M}, s \models \langle\langle A \rangle\rangle\phi$ holds, but $\mathcal{G}, s \not\models_\sigma \langle\langle A \rangle\rangle\phi$.*

Proof. Let us consider the CGS shown in Figure 1. There are two agents $\{1, 2\}$, four states $\{s_0, s_1, s_2, s_3\}$ (s_0 is the initial state), $\lambda(s_0) = \lambda(s_1) = \lambda(s_2) = \{q\}$ and $\lambda(s_3) = \emptyset$, $\sim_1$ is the identity relation, $s \sim_2 s'$ for every $s, s' \in \{s_0, s_1, s_2\}$ and $s_3 \sim_2 s_3$. Consider the function π such that $\pi(1) = \mathtt{IR}$ and $\pi(2) = \mathtt{ir}$, it is easy to see that $\mathcal{M}, s_0 \models \langle\langle \{1\} \rangle\rangle \mathbf{G} q$, while $\mathcal{G}, s_0 \not\models_{\mathtt{IR}} \langle\langle \{1\} \rangle\rangle \mathbf{G} q$. □

5 Model-Checking Algorithms

It has been shown that the Turing Halting problem can be reduced to the model-checking problem of CGS against the

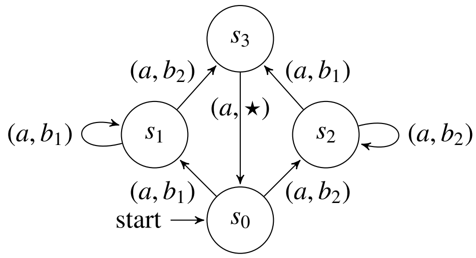

Figure 1: An illustrating example, where $\star \in \{b_1, b_2\}$.

ATL formula $\varphi = \langle\langle\{1,2\}\rangle\rangle\mathbf{G}\ ok$ under the $\mathtt{iR}$-setting [Dima and Tiplea, 2011], where ok is an atomic proposition. By adapting the proof, we get that:

Theorem 1 *The model-checking problem for ACGS $\mathcal{M} = (\mathcal{G}, \pi)$ against the ATL/ATL* formula $\langle\langle\{1,2\}\rangle\rangle\mathbf{G}\ ok$ is undecidable, where $\{1,2,3\} \subseteq Ag$, $\pi(1) = \pi(2) = \mathtt{iR}$ and $\pi(i) = \mathtt{IR}$ for $i \in Ag \setminus \{1,2\}$.*

By Theorem 1, we focus on the model-checking problem of ACGS by restricting the function π to $Ag \to \mathrm{T}_{\mathtt{str}} \setminus \{\mathtt{iR}\}$. In general, we propose model-checking algorithms which iteratively compute the set of states satisfying state formulae from the innermost subformulae. The main challenge is to compute $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$. To this end, we first show how to compute $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$ for a simple formula $\langle\langle A\rangle\rangle\phi$, and then present the more general algorithm. An ATL/ATL* formula $\langle\langle A\rangle\rangle\phi$ is called *simple* if ϕ is an LTL formula.

Let us fix a simple formula $\langle\langle A\rangle\rangle\phi$ and an ACGS $\mathcal{M} = (\mathcal{G}, \pi)$ with $\mathcal{G} = (S, S_0, Ag, (Ac_i)_{i\in Ag}, (\sim_i)_{i\in Ag}, (P_i)_{i\in Ag}, \Delta, \lambda)$. Given a set of agents A' and a strategy type $\sigma \in \mathrm{T}_{\mathtt{str}}$, we denote by A'_σ the set $\{i \in A' \mid \pi(i) = \sigma\}$.

5.1 Model-Checking Simple ATL Formulae

For the simple ATL formula $\langle\langle A\rangle\rangle\phi$, it is easy to see that whether agents in A have perfect call or not does not matter if these agents have perfect information abilities.

Proposition 2 *Given an ACGS $(\mathcal{G}, \pi)$ with $\pi : Ag \to \mathrm{T}_{\mathtt{str}} \setminus \{\mathtt{iR}\}$, and a simple ATL formula $\langle\langle A\rangle\rangle\varphi$, let π' be a function such that for every $i \in Ag$, $\pi'(i) = \mathtt{Ir}$ if $\pi(i) = \mathtt{IR}$ and $i \in A$, otherwise $\pi'(i) = \pi(i)$. For every state s in $\mathcal{M}$,*

$$(\mathcal{G}, \pi), s \models \langle\langle A\rangle\rangle\varphi \text{ iff } (\mathcal{G}, \pi'), s \models \langle\langle A\rangle\rangle\varphi.$$

By Proposition 2, all the agents in A with $\mathtt{IR}$-strategies can be seen as agents with $\mathtt{Ir}$-strategies. All the agents in Ag with $\mathtt{Ir}$-strategies can be seen as agents with $\mathtt{ir}$-strategies (i.e., all the epistemic accessibility relations of them are the identity relation). Therefore, we can assume that $\pi(i) = \mathtt{ir}$ for all $i \in A$, and $\pi(i) \in \{\mathtt{ir}, \mathtt{IR}\}$ for all $i \in \overline{A}$. For two collective strategies ξ_A and $\xi_{\overline{A}_{\mathtt{ir}}}$, let $\mathcal{M}(\xi_A, \xi_{\overline{A}_{\mathtt{ir}}}) = (\mathcal{G}', \pi)$ be the ACGS obtained from $(\mathcal{G}, \pi)$ by enforcing strategies ξ_A and $\xi_{\overline{A}_{\mathtt{ir}}}$, namely, by removing transitions whose actions of agents in $A \cup \overline{A}_{\mathtt{ir}}$ do not conform to ξ_A and $\xi_{\overline{A}_{\mathtt{ir}}}$. We have that

$$[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}} \equiv \bigcup_{\xi_A} \bigcap_{\xi_{\overline{A}_{\mathtt{ir}}}} [\![\langle\langle\emptyset\rangle\rangle\phi]\!]_{\mathcal{M}(\xi_A, \xi_{\overline{A}_{\mathtt{ir}}})}.$$

Computing $[\![\langle\langle\emptyset\rangle\rangle\phi]\!]_{\mathcal{M}(\xi_A, \xi_{\overline{A}_{\mathtt{ir}}})}$ amounts to CTL model-checking, which can be done in polynomial time (and thus in polynomial space) in the size of $\mathcal{M}(\xi_A, \xi_{\overline{A}_{\mathtt{ir}}})$ and $\langle\langle\emptyset\rangle\rangle\phi$ [Clarke *et al.*, 1983]. Since the number of strategies ξ_A and $\xi_{\overline{A}_{\mathtt{ir}}}$ is finite, we get that:

Lemma 1 *For the simple ATL formula $\langle\langle A\rangle\rangle\phi$, $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$ can be computed in PSPACE.*

5.2 Model-Checking Simple ATL* Formulae

We compute $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$ by a reduction to the problem of computing the winning region of a turn-based two-player parity game. We first introduce some basic concepts which will be used in our reduction.

A *deterministic parity automaton* (DPA) is a tuple $\mathcal{A} = (P, \Sigma, \delta, p_0, R)$, where P is a finite set of states, Σ is the input alphabet, $\delta : P \times \Sigma \to P$ is a transition function, $p_0 \in P$ is the initial state and $R : P \to \{0, ..., k\}$ is a *rank* function. A *run* ρ of $\mathcal{A}$ over an ω-word $\alpha_0\alpha_1... \in \Sigma^\omega$ is an infinite sequence of states $\rho = p_0 p_1...$ such that for every $i \geq 0$, $p_{i+1} = \delta(p_i, \alpha_i)$. Let $\mathsf{inf}(\rho)$ be the set of states visited infinitely often in ρ. An infinite word is *recognized* by $\mathcal{A}$ if $\mathcal{A}$ has a run ρ over this word such that $\min_{p\in\mathsf{inf}(\rho)} R(p)$ is *even*. For the LTL formula ϕ, one can construct a DPA $\mathcal{A}_\phi = (P, 2^{AP}, \delta, p_0, R)$ with $2^{2^{O(|\phi|)}}$ states and rank $k = 2^{O(|\phi|)}$ such that $\mathcal{A}_\phi$ recognizes all the ω-words satisfying ϕ [Piterman, 2006].

A *(turned-based, two-player) parity game* $\mathcal{P}$ is a tuple $(V = V_0 \uplus V_1, E, \Xi)$, where V_i for $i \in \{0, 1\}$ is a finite set of vertices controlled by Player-i, $E \subseteq V \times V$ is a finite set of edges, and $\Xi : V \to \{0, ..., k\}$ is a *rank* function. A *play* ρ starting from v_0 is an infinite sequence of vertices $v_0v_1...$ such that for every $i \geq 0$, $(v_i, v_{i+1}) \in E$. A *strategy* of Player-i is a function $\theta : V^*V_i \to V$ such that for every $\rho \in V^*$ and $v \in V_i$, $(v, \theta(\rho \cdot v)) \in E$. Given a strategy θ_0 for Player-0 and a strategy θ_1 for Player-1, let G_{θ_0,θ_1} be the play where Player-0 and Player-1 enforce their strategies θ_0 and θ_1. θ_0 is a winning strategy for Player-0 if $\min_{s\in\mathsf{inf}(G_{\theta_0,\theta_1})} \Xi(s)$ is *even* for every strategy θ_1 of Player-1. The winning region of Player-0, denoted by WR_0, is the set of vertices from which Player-0 has a winning strategy.

A *partial strategy* of A' is a partial function $f : A' \times S \to \bigcup_{i\in A'} Ac_i$ such that for each $i \in A'$ and $s \in S$, if $f(i, s)$ is defined, then for all $s' \in S$ with $s \sim_i s'$, $f(i, s) = f(i, s') \in P_i(s)$. We denote by $\mathsf{dom}(f)$ the domain of f. Let $F_{\mathtt{ir}}$ (resp. $G_{\mathtt{ir}}$) be the set of partial strategies of $A_{\mathtt{ir}}$ (resp. $\overline{A}_{\mathtt{ir}}$). Let $F^\top_{\mathtt{ir}}$ denote the set $\{f \in F_{\mathtt{ir}} \mid \mathsf{dom}(f) = A_{\mathtt{ir}} \times S\}$, and $g_\perp \in G_{\mathtt{ir}}$ denote the partial strategy such that $\mathsf{dom}(g_\perp) = \emptyset$. Given a state s, let $F^s_{\mathtt{IR}}$ be the set of functions $f : A_{\mathtt{IR}} \to \bigcup_{i\in A_{\mathtt{IR}}} Ac_i$ such that for every $i \in A_{\mathtt{IR}}$, $f(i) \in P_i(s)$. Let $F_{\mathtt{IR}} := \bigcup_{s\in S} F^s_{\mathtt{IR}}$ and $\Pi_{\mathtt{ir}} := \{G \subseteq G_{\mathtt{ir}} \mid \forall g, g' \in G.\ \mathsf{dom}(g) = \mathsf{dom}(g')\}$.

We define a parity game $\mathcal{P}_\phi \triangleq (V = V_0 \uplus V_1, E, \Xi)$, where $V_0 = S \cup (S \times P \times F^\top_{\mathtt{ir}} \times \Pi_{\mathtt{ir}})$, $V_1 = (S \times F^\top_{\mathtt{ir}}) \cup (S \times P \times F^\top_{\mathtt{ir}} \times F_{\mathtt{IR}} \times \Pi_{\mathtt{ir}})$, $\Xi : V \to \{0, \cdots, k\}$ is a function such that for every $s \in S$,

- $\Xi(s) = \Xi(s, f_\top) = 0$, $\forall f_\top \in F^\top_{\mathtt{ir}}$,
- $\Xi(s, p, f_\top, G) = \Xi(s, p, f_\top, f, G) = R(p)$, $\forall p \in P$, $\forall f_\top \in F^\top_{\mathtt{ir}}$, $\forall f \in F_{\mathtt{IR}}$ and $\forall G \in \Pi_{\mathtt{ir}}$,

E is defined as follows:

- $(s, (s, f_\top)) \in E$ for $s \in S$ and $f_\top \in F^\top_{\mathtt{ir}}$;
- $((s, f_\top), (s, p_0, f_\top, \{g_\perp\})) \in E$ for $s \in S$ and $f_\top \in F^\top_{\mathtt{ir}}$;

- $((s,p,f_\top,G),(s,p,f_\top,f,G)) \in E$ for $(s,p,f_\top,G) \in V_0$ and $f \in F^s_{\mathtt{IR}}$;
- $((s,p,f_\top,f,G),(s',\delta(p,\lambda(s)),f_\top,G')) \in E$ for every $(s,p,f_\top,f,G) \in V_1$ and $s' \in S$, where $G' \subseteq \Pi_{\mathtt{ir}}$ is the largest set such that the follows hold: for every $g' \in G'$,
  1. there exists $g \in G$ such that $\mathtt{dom}(g') = \mathtt{dom}(g) \cup \{(i,s'') \in \overline{A}_{\mathtt{ir}} \times S \mid s \sim_i s''\}$ and for every $(i,s'') \in \mathtt{dom}(g)$, $g'(i,s'') = g(i,s'')$;
  2. there exists $\vec{a} \in Ac$ such that $s' = \Delta(s,\vec{a})$, and $\forall i \in Ag, ((i,s) \in \mathtt{dom}(f_\top) \Rightarrow \vec{a}_i = f_\top(i,s)) \wedge (i \in \mathtt{dom}(f) \Rightarrow \vec{a}_i = f(i)) \wedge ((i,s) \in \mathtt{dom}(g') \Rightarrow \vec{a}_i = g'(i,s))$.

In this reduction, $f_\top$ encodes a collective $\mathtt{ir}$-strategy of agents in $A_{\mathtt{ir}}$, the collection of f's in one play of $\mathcal{P}_\phi$ encodes a collective $\mathtt{IR}$-strategy of agents in $A_{\mathtt{IR}}$, and each $g \in G$ encodes a collective $\mathtt{ir}$-strategy of agents in $\overline{A}_{\mathtt{ir}}$. The imperfect information abilities of agents are ensured by the definitions of partial strategies.

To check whether $s \in [\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$, $\mathcal{P}_\phi$ starts with the vertex s. At first step, Player-0 chooses a function $f_\top \in F^\top_{\mathtt{ir}}$ meaning that all agents in $A_{\mathtt{ir}}$ choose an $\mathtt{ir}$-strategy. Next, $\mathcal{P}_\phi$ moves from $(s,f_\top)$ to $(s,p_0,f_\top,\{g_\perp\})$ which let the DPA $\mathcal{A}_\phi$ start with p_0 (note that Player-1 has only one choice at this step). At a vertex $(s,p,f_\top,G)$ controlled by Player-0, Player-0 chooses actions for agents in $A_{\mathtt{IR}}$ by choosing one function $f \in F^s_{\mathtt{IR}}$. Then Player-1 chooses actions for agents in $\overline{A}$ with respect to the chosen actions of agents in $\overline{A}_{\mathtt{ir}}$ tracked by G. These selections of actions together with $f_\top$ and G determine a joint action $\vec{a}$, based on which $\mathcal{P}_\phi$ moves to $(s',\delta(p,\lambda(s)),f_\top,G')$ such that s' is the next state of s under $\vec{a}$, $\delta(p,\lambda(s))$ is the next state of p in $\mathcal{A}_\phi$ which allows to mimics the run of $\mathcal{A}_\phi$ over the ω-word induced by the play of $\mathcal{M}$. During this step, f is dropped from the vertex of $\mathcal{P}_\phi$, as f corresponds to actions of agents in $A_{\mathtt{IR}}$ and needs not to track. The actions of agents in $\overline{A}_{\mathtt{ir}}$ are tracked by computing G' from G. This ensures imperfect recall abilities of agents in $\overline{A}_{\mathtt{ir}}$. We then can get that:

Lemma 2 $WR_0 \cap S = [\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$.

The winning region of Player-0 in $\mathcal{P}_\phi$ can be computed in polynomial time of $|V| \cdot |E| \cdot 2^k$ [Jurdzinski, 2000]. In this reduction, each G contributes at most $O(|S|)$ sets of G'. Therefore, $|V| \cdot |E|$ is exponential in $|\mathcal{G}| \cdot 2^{|\phi|}$. Recall that $k = 2^{O(|\phi|)}$. Consequently, we have

Lemma 3 *For the simple ATL* formula $\langle\langle A\rangle\rangle\phi$, $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$ can be computed in 2EXPTIME.*

5.3 The Overall Algorithm

We now present the overall procedure, which computes $[\![\varphi]\!]_{\mathcal{M}}$ from the innermost subformulae. Algorithm 1 shows the pseudo code, which takes an ACGS $\mathcal{M} = (\mathcal{G},\pi)$ and an ATL/ATL* formula φ as inputs, and outputs $[\![\varphi]\!]_{\mathcal{M}}$. Then $\mathcal{M}$ satisfies φ iff the set of initial states of $\mathcal{M}$ is a subset of $[\![\varphi]\!]_{\mathcal{M}}$. We also incorporate epistemic modalities $\mathbf{K}_i\varphi, \mathbf{E}_A\varphi, \mathbf{D}_A\varphi, \mathbf{C}_A\varphi$ from [van der Hoek and Wooldridge, 2003; Cermák *et al.*, 2014] into our algorithm with the following semantics:

Algorithm 1: Function $\mathtt{MC}(\mathcal{M},\varphi)$ outputs $[\![\varphi]\!]_{\mathcal{M}}$

switch φ **:**
- **case** q **: return** $\{s \in S \mid q \in \lambda(s)\}$;
- **case** $\neg\varphi'$ **: return** $S \setminus \mathtt{MC}(\mathcal{M},\varphi')$;
- **case** $\varphi_1 \wedge \varphi_2$ **: return** $\mathtt{MC}(\mathcal{M},\varphi_1) \cap \mathtt{MC}(\mathcal{M},\varphi_2)$;
- **case** $\mathbf{K}_i\varphi'$ **: return** $\{s \in S \mid [s]^{\sim_i} \subseteq \mathtt{MC}(\mathcal{M},\varphi')\}$;
- **case** $\mathbf{E}_A\varphi'$ **: return** $\{s \in S \mid [s]^{\sim^E_A} \subseteq \mathtt{MC}(\mathcal{M},\varphi')\}$;
- **case** $\mathbf{D}_A\varphi'$ **: return** $\{s \in S \mid [s]^{\sim^D_A} \subseteq \mathtt{MC}(\mathcal{M},\varphi')\}$;
- **case** $\mathbf{C}_A\varphi'$ **: return** $\{s \in S \mid [s]^{\sim^C_A} \subseteq \mathtt{MC}(\mathcal{M},\varphi')\}$;
- **case** $\langle\langle A\rangle\rangle\phi$ **:**
  - **foreach** *sub-state-formula* φ' **in** ϕ **do**
    - Replace φ' by a fresh atomic proposition $q_{\varphi'}$ in φ, and let $\lambda(q_{\varphi'}) := \mathtt{MC}(\mathcal{M},\varphi')$;
  - Compute $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$ by Lemma 1 or 3;
  - **return** $[\![\langle\langle A\rangle\rangle\phi]\!]_{\mathcal{M}}$;

- $\mathcal{G}, s \models_\sigma \mathbf{K}_i\varphi$ iff $\forall s' \in S,\ s \sim_i s' \Longrightarrow \mathcal{G}, s' \models_\sigma \varphi$;
- $\mathcal{G}, s \models_\sigma \mathbf{E}_A\varphi$ iff $\forall s' \in S,\ s \sim^E_A s' \Longrightarrow \mathcal{G}, s' \models_\sigma \varphi$;
- $\mathcal{G}, s \models_\sigma \mathbf{D}_A\varphi$ iff $\forall s' \in S,\ s \sim^D_A s' \Longrightarrow \mathcal{G}, s' \models_\sigma \varphi$;
- $\mathcal{G}, s \models_\sigma \mathbf{C}_A\varphi$ iff $\forall s' \in S,\ s \sim^C_A s' \Longrightarrow \mathcal{G}, s' \models_\sigma \varphi$;

where φ is a state formula, $\sim^E_A = \bigcup_{i\in A} \sim_i$, $\sim^D_A = \bigcap_{i\in A} \sim_i$, $\sim^C_A = (\sim^E_A)^+$ (note that $(\sim^E_A)^+$ is the transitive closure of $\sim^E_A$). $\mathbf{K}_i\varphi, \mathbf{E}_A\varphi, \mathbf{D}_A\varphi$ and $\mathbf{C}_A\varphi$ respectively denote that i knows, every agent in A knows, agents in A have common knowledge and agents in A have distributed knowledge, on the fact φ. The ATL (resp. ATL*) logic extended with these epistemic modalities is called ATLK (resp. ATLK*) logic. Given a state $s \in S$ and a binary relation $\asymp \subseteq S \times S$, we denote by $[s]^{\asymp}$ the set $\{s' \in S \mid s \asymp s'\}$.

By Lemma 1 and Lemma 3, the model-checking problem for ATLK and ATLK* on ACGS can be solved in PSPACE and 2EXPTIME respectively. As the model-checking problem of ATL* on CGS under $\mathtt{IR}$-setting is 2EXPTIME-complete [Alur *et al.*, 2002], we have that

Theorem 2 *The model-checking problem for ATLK* (resp. ATLK) on ACGS is 2EXPTIME-complete (resp. in PSPACE).*

6 Implementation and Experiments

We have implemented the ATLK/ATLK* model-checking algorithms in MCMAS [Lomuscio *et al.*, 2017]. All experiments were conducted on a computer with 1.70GHz Intel Core E5-2603 CPU and 32GB of memory.

We conducted experiments on the castle game (CG, [Pilecki *et al.*, 2014]) where there are several agents modelling workers and an environment agent. Each worker works for the benefit of a castle, and the environment keeps track of the Health Points (HP) of castles. Each castle preserves an HP valued up to 3, and 0 means it's defeated. Workers are able to attack a castle which they don't work for, or defend the castle which they work for, or do nothing. The castle gets damaged if the number of attackers is greater than the number of defenders, and the differences influence its HP. In this model, the number of states is $8000 \times 4n$, the environment

Table 1: Results of the castle game

π	φ_1			φ_2		
	Lem. 1	Lem. 3	SAT	Lem. 1	Lem. 3	SAT
(IR, IR, IR, IR)	N/A	20.295	Y	N/A	18.178	Y
(IR, IR, IR, ir)	N/A	7523.67	Y	N/A	7377.44	Y
(IR, IR, ir, IR)	N/A	31.904	Y	N/A	30.578	N
(IR, ir, IR, IR)	N/A	32.446	Y	N/A	31.259	N
(IR, IR, ir, ir)	N/A	3402.56	Y	N/A	3451.59	N
(IR, ir, IR, ir)	N/A	3294.51	Y	N/A	3366.71	N
(IR, ir, ir, IR)	5.822	24.254	Y	77.514	23.37	N
(IR, ir, ir, ir)	13.791	113.493	Y	45.679	113.647	N

agent has 1 local action, and each worker agent has 4 local actions, where n denotes the number of workers.

In this experiment, we consider ACGS consisting of three worker agents w_1, w_2, w_3 and an environment agent e, where w_i works for the castle c_i. For the properties, we consider the following two:

- $\varphi_1 \equiv \langle\langle\{w_1, w_2\}\rangle\rangle\mathbf{F}(castle3Defeated)$, which expresses that workers w_1 and w_2 can make castle c_3 defeated, no matter which strategies the other agents use.
- $\varphi_2 \equiv \langle\langle\{w_1, w_2\}\rangle\rangle\mathbf{F}(allDefeated)$, which expresses that workers w_1 and w_2 can make all the castles defeated, no matter which strategies the other agents use.

The results are shown in Table 1, where $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ in each row denotes the strategy types of agents e, w_1, w_2, w_3, N/A denotes timeout (2.5 hours), Y (resp. N) denotes that the model satisfies (resp. invalidates) the formula, and columns 2–4 (resp. 5–7) show total time (in seconds) and result of verifying φ_1 (resp. φ_2) using Lemma 1 and Lemma 3 respectively.

We observe that: (1) the strategy types of agents do affect the performance and results. In particular, the time significantly increases when w_3 is ir-typed while w_1 or w_2 is IR-typed; (2) Lemma 1 is more efficient when both w_1 and w_2 are ir-typed; otherwise Lemma 3 is more efficient. This is because the number of possible strategies of w_1 and w_2 is small (using Lemma 1) if both w_1 and w_2 are ir-typed.

Further experiments (on the Book Store Application and Dining Cryptographers Protocol [Lomuscio *et al.*, 2017]) can be found in the appendix.

7 Related Work

The family of alternating-time temporal logics ATL, ATL* and alternating μ-calculus (AMC) [Alur *et al.*, 2002] for reasoning about games was introduced with motivations partially from MAS. Model-checking algorithms were also given with IR-strategies. [van der Hoek and Wooldridge, 2003] extended ATL with knowledge operators and proposed corresponding model-checking algorithms. In their work, epistemic accessibility relations are considered in the interpretation of knowledge operators, but not for the strategies and outcomes. This means that agents still use IR strategies for coalition modalities $\langle\langle A\rangle\rangle\phi$. This issue was discussed in [Jamroga, 2003] which proposed an idea of iR-strategies. [Schobbens, 2004] introduced the notion of imperfect recall into ATL/ATL*, and investigated the model-checking problem for ATL/ATL* under four different strategic types. Importantly, with iR-strategies the model-checking problem becomes undecidable [Dima and Tiplea, 2011]. [Bulling and Jamroga, 2011] studied the semantics of AMC and proposed a model-checking algorithm for the alternation-free fragment under the imperfect information setting. [Bulling and Jamroga, 2014] further conducted a comprehensive comparison of variants of ATL/ATL* with different strategic types. This study corroborates that the agents' abilities play a prominent role in logic semantics.

In the previous work, strategies are revocable, i.e., when it comes to achieve a goal in the (nested) sub-formula, previously selected strategies are deleted. [Ågotnes *et al.*, 2007] introduced a variant of ATL with *irrevocable* strategies under the imperfect recall setting. It was generalized into ATL/ATL* with strategy contexts [Laroussinie and Markey, 2015], which allows agents to drop or inherit previously selected strategies.

Two strategic logics were introduced by [Chatterjee *et al.*, 2010] and [Mogavero *et al.*, 2014] and the model-checking problem was investigated therein under the IR-setting. Strategic logics extend LTL with first-order quantifications over strategies which naturally captures the multi-player game nature in the evolution of MAS. [Cermák *et al.*, 2014] introduced knowledge operators in the strategic logic of [Mogavero *et al.*, 2014] and proposed a model-checking algorithm with ir-strategies. Here all agents must take ir-strategies (so the potential inconsistency can be ruled out), but no other abilities are considered. To gain decidability under iR-setting, specific restrictions on the abilities of the agents were proposed [Berthon *et al.*, 2017a; Berthon *et al.*, 2017b; Belardinelli *et al.*, 2017a; Belardinelli *et al.*, 2017b; Belardinelli *et al.*, 2018].

Our work is orthogonal to the existing work which defines the agents' abilities at the semantics level, but takes a more syntactic level by strengthening the model.

8 Conclusion and Future Work

In this paper, we discussed the problem of existing semantics of ATL/ATL*, and advocated the approach to make agents' abilities explicit in modeling. For this purpose, we introduced an extension of standard CGS model, i.e. ACGS, which define agents' abilities at the syntactic level of the system model. We explored the effects of strategy types in the semantics, in particular model-checking, of ATL/ ATL* over ACGS, and provided model-checking algorithms with identified complexity. The algorithms are implemented in a tool MCMAS-ACGS, which has been applied to several applications to demonstrate the feasibility of the approach. This work represents the first systematic study towards different agents' abilities at the syntactic level, which is in contrast to the previous approaches at the semantic level.

Currently we use ATL/ATL* as the specification, but the methodology can be extended to other logics such as Strategy Logic, and other agents' abilities such as strategy contexts. Several questions are left open such as axiomatization and satisfiability problem. We leave them for future work.

References

[Ågotnes *et al.*, 2007] Thomas Ågotnes, Valentin Goranko, and Wojciech Jamroga. Alternating-time temporal logics with irrevocable strategies. In *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 15–24, 2007.

[Alur *et al.*, 2002] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, 2002.

[Belardinelli *et al.*, 2017a] Francesco Belardinelli, Alessio Lomuscio, Aniello Murano, and Sasha Rubin. Verification of broadcasting multi-agent systems against an epistemic strategy logic. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 91–97, 2017.

[Belardinelli *et al.*, 2017b] Francesco Belardinelli, Alessio Lomuscio, Aniello Murano, and Sasha Rubin. Verification of multi-agent systems with imperfect information and public actions. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1268–1276, 2017.

[Belardinelli *et al.*, 2018] Francesco Belardinelli, Alessio Lomuscio, Aniello Murano, and Sasha Rubin. Decidable verification of multi-agent systems with bounded private actions. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1865–1867, 2018.

[Berthon *et al.*, 2017a] Raphaël Berthon, Bastien Maubert, and Aniello Murano. Decidability results for ATL* with imperfect information and perfect recall. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1250–1258, 2017.

[Berthon *et al.*, 2017b] Raphaël Berthon, Bastien Maubert, Aniello Murano, Sasha Rubin, and Moshe Y. Vardi. Strategy logic with imperfect information. In *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–12, 2017.

[Bulling and Jamroga, 2011] Nils Bulling and Wojciech Jamroga. Alternating epistemic mu-calculus. In *Proceedings of the 22rd International Joint Conference on Artificial Intelligence*, pages 109–114, 2011.

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

[Cermák *et al.*, 2014] Petr Cermák, Alessio Lomuscio, Fabio Mogavero, and Aniello Murano. MCMAS-SLK: A model checker for the verification of strategy logic specifications. In *Proceedings of the 26th International Conference on Computer Aided Verification (CAV)*, pages 525–532, 2014.

[Chatterjee *et al.*, 2010] K. Chatterjee, T. A. Henzinger, and N. Piterman. Strategy logic. *Information and Computation*, 208(6):677–693, 2010.

[Clarke *et al.*, 1983] Edmund M. Clarke, E. Allen Emerson, and A. Prasad Sistla. Automatic verification of finite state concurrent systems using temporal logic specifications: A practical approach. In *POPL*, pages 117–126, 1983.

[Dima and Tiplea, 2011] Catalin Dima and Ferucio Laurentiu Tiplea. Model-checking ATL under imperfect information and perfect recall semantics is undecidable. *CoRR*, abs/1102.4225, 2011.

[Jamroga and van der Hoek, 2004] Wojciech Jamroga and Wiebe van der Hoek. Agents that know how to play. *Fundam. Inform.*, 63(2-3):185–219, 2004.

[Jamroga, 2003] Wojciech Jamroga. Some remarks on alternating temporal epistemic logic. *Proceedings of Formal Approaches to Multi-Agent Systems (FAMAS)*, pages 133–140, 2003.

[Jurdzinski, 2000] Marcin Jurdzinski. Small progress measures for solving parity games. In *Proceedings of the 17th Symposium on Theoretical Aspects of Computer Science*, pages 290–301, 2000.

[Laroussinie and Markey, 2015] François Laroussinie and Nicolas Markey. Augmenting ATL with strategy contexts. *Inf. Comput.*, 245:98–123, 2015.

[Laroussinie *et al.*, 2008] François Laroussinie, Nicolas Markey, and Ghassan Oreiby. On the expressiveness and complexity of ATL. *Logical Methods in Computer Science*, 4(2), 2008.

[Lomuscio *et al.*, 2017] Alessio Lomuscio, Hongyang Qu, and Franco Raimondi. MCMAS: an open-source model checker for the verification of multi-agent systems. *STTT*, 19(1):9–30, 2017.

[Mogavero *et al.*, 2014] Fabio Mogavero, Aniello Murano, Giuseppe Perelli, and Moshe Y. Vardi. Reasoning about strategies: On the model-checking problem. *ACM Transations on Computational Logic*, 15(4):34:1–34:47, 2014.

[Pilecki *et al.*, 2014] Jerzy Pilecki, Marek A. Bednarczyk, and Wojciech Jamroga. Model checking properties of multi-agent systems with imperfect information and imperfect recall. In *Proceedings of the 7th International Conference on Intelligent Systems (IS)*, pages 415–426, 2014.

[Piterman, 2006] N. Piterman. From nondeterministic Büchi and Streett automata to deterministic parity automata. In *LICS 2006*, pages 255–264, 2006.

[Schobbens, 2004] Pierre-Yves Schobbens. Alternating-time logic with imperfect recall. *Electr. Notes Theor. Comput. Sci.*, 85(2):82–93, 2004.

[van der Hoek and Wooldridge, 2003] Wiebe van der Hoek and Michael Wooldridge. Cooperation, knowledge, and time: Alternating-time temporal epistemic logic and its applications. *Studia Logica*, 75(1):125–157, 2003.

A Proof of Theorem 1

Given a Turing machine M, let $\mathcal{G} = (S, \{s_0\}, Ag, (Ac_i)_{i\in Ag}, (\sim_i)_{i\in Ag}, (P_i)_{i\in Ag}, \Delta, \lambda)$ be the CGS constructed as in [Dima and Tiplea, 2011] and s_0 be an initial state in $\mathcal{G}$ such that $\mathcal{G}, s_0 \models_{\mathtt{iR}} \langle\langle\{1,2\}\rangle\rangle\mathbf{G}\ ok$ iff M does not halt on the empty word. Let $\mathcal{M} = (\mathcal{G}, \pi)$ be an ACGS such that for every agent i in $\mathcal{M}$, $\pi(i) = \mathtt{iR}$ if $i \in \{1,2\}$, otherwise $\pi(i) = \mathtt{IR}$. Therefore, $\mathcal{G}, s_0 \models_{\mathtt{iR}} \langle\langle\{1,2\}\rangle\rangle\mathbf{G}\ ok$ iff $\mathcal{M}, s_0 \models \langle\langle\{1,2\}\rangle\rangle\mathbf{G}\ ok$. The result immediately follows.

B Proof of Proposition 2

Without loss of generality, we assume that for every $i \in Ag$, $\pi(i) \neq \mathtt{Ir}$, as $\mathtt{Ir}$ can be seen as a special case of $\mathtt{ir}$. We first construct the tree-unfolding of $\mathcal{M}$ from the state s.

The tree-unfolding of $\mathcal{M}$ from s is an ACGS $\mathcal{M}_s^* = (\mathcal{G}^*, \pi^*)$ such that $\mathcal{G}^* = (S^+, S_0, Ag, (Ac_i)_{i\in Ag}, (\sim_i^*)_{i\in Ag}, (P_i^*)_{i\in Ag}, \Delta^*, \lambda^*)$, where

- for every $i \in Ag$, $\pi^*(i) = \mathtt{Ir}$ if $\pi(i) = \mathtt{IR}$ and $i \in A$, otherwise $\pi^*(i) = \pi(i)$;
- for every $i \in Ag$ and $\rho_1, \rho_2 \in S^+$, $\rho_1 \sim_i^* \rho_2$, if either $\mathtt{lst}(\rho_1) \sim_i \mathtt{lst}(\rho_2)$ and $\pi(i) \neq \mathtt{IR}$, or $\rho_1 = \rho_2$ and $\pi(i) = \mathtt{IR}$,
- $\lambda^*(\rho) = \lambda(\mathtt{lst}(\rho))$ for every $\rho \in S^+$;
- $P_i^*(\rho) = P_i(\mathtt{lst}(\rho))$ for every $i \in Ag$ and $\rho \in S^+$;
- $\Delta^*(\rho, \vec{d}) = \rho \cdot \Delta(\mathtt{lst}(\rho), \vec{d})$ for every $\rho \in S^+, \vec{d} \in Ac$.

From the above definition, we can immediately get that the tree-unfolding $\mathcal{M}_q^*$ is a tree-like ACGS, namely, every state can be reached by a unique finite path from the root. $\mathtt{IR}$-strategies of A from s in $\mathcal{M}$ correspond exactly to $\mathtt{Ir}$-strategies of A in the tree unfolding $\mathcal{M}_q^*$ from s, while the types of other agents are same under π and π^*. Thus, $\mathcal{M}, s \models \langle\langle A\rangle\rangle\varphi$ iff $\mathcal{M}_s^*, s \models \langle\langle A\rangle\rangle\varphi$. Note that this result does not hold if φ is a general LTL formula.

Next, we will show that $\mathcal{M}', s \models \langle\langle A\rangle\rangle\varphi$ iff $\mathcal{M}_s^*, s \models \langle\langle A\rangle\rangle\varphi$.

($\Rightarrow$) Suppose $\mathcal{M}', s \models \langle\langle A\rangle\rangle\varphi$, let ξ_A be the collective strategy such that for every path $\rho \in O_{\mathcal{M}'}(s, \xi_A)$, $\mathcal{M}', \rho \models \varphi$. Let ξ_A^* be the function such that for every $i \in A$ and $\rho \in S^+$, $\xi_A^*(i)(\rho) = \xi_A(i)(\mathtt{lst}(\rho))$.

First, we show that ξ_A^* is a collective strategy of A in $\mathcal{M}_s^*$. Consider an agent $i \in A$ and two states $\rho_1, \rho_2 \in S^+$ such that $\rho_1 \sim_i^* \rho_2$, if $\pi(i) \neq \mathtt{IR}$, then $\mathtt{lst}(\rho_1) \sim_i \mathtt{lst}(\rho_2)$ which implies that $\xi_A(i)(\mathtt{lst}(\rho_1)) = \xi_A(i)(\mathtt{lst}(\rho_2))$, hence $\xi_A^*(i)(\rho_1) = \xi_A^*(i)(\rho_2)$. Otherwise $\rho_1 = \rho_2$ and $\pi(i) = \mathtt{IR}$. This implies that $\xi_A(i)(\mathtt{lst}(\rho_1)) = \xi_A(i)(\mathtt{lst}(\rho_2))$, hence $\xi_A^*(i)(\rho_1) = \xi_A^*(i)(\rho_2)$ as well. Therefore, ξ_A^* is a collective strategy of A in $\mathcal{M}_s^*$.

Next, we show that for every collective strategy $\xi_{\overline{A}}^*$ of $\overline{A}$ in $\mathcal{M}_s^*$, $\mathtt{play}(s, \xi_A^*, \xi_{\overline{A}}^*) \models \varphi$. Suppose $\mathtt{play}(s, \xi_A^*, \xi_{\overline{A}}^*) = \rho_0\rho_1\cdots$. Let $\xi_{\overline{A}}$ be the function such that for every $i \in \overline{A}$ and $j \geq 0$, $\xi_A(i)(\mathtt{lst}(\rho_j)) = \xi_A^*(i)(\rho_j)$ if $\pi(i) \neq \mathtt{IR}$, otherwise $\xi_A(i)(\rho_j) = \xi_A^*(i)(\rho_0\cdots\rho_j)$. Consider $j, k \geq 0$ such that $\mathtt{lst}(\rho_j) \sim_i \mathtt{lst}(\rho_k)$ for some $i \in \overline{A}$, if $\pi(i) \neq \mathtt{IR}$, then $\rho_j \sim_i^* \rho_k$, which implies that $\xi_A^*(i)(\rho_j) = \xi_A^*(i)(\rho_k)$, hence $\xi_A(i)(\mathtt{lst}(\rho_j)) = \xi_A(i)(\mathtt{lst}(\rho_k))$. Otherwise, $\pi(i) = \mathtt{IR}$, i can choose any action at any state of ρ_j. Thus, $\xi_{\overline{A}}$ is a collective strategy of $\overline{A}$ in $\mathcal{M}'$ and $\mathtt{play}(s, \xi_A, \xi_{\overline{A}}) = \mathtt{lst}(\rho_0)\mathtt{lst}(\rho_1)\cdots$. The result immediately follows from the fact that $\lambda^*(\rho) = \lambda(\mathtt{lst}(\rho))$ for every $\rho \in S^+$.

($\Leftarrow$) Suppose $\mathcal{M}_s^*, s \models \langle\langle A\rangle\rangle\varphi$, let ξ_A^* be the collective strategy such that for every path $\rho \in O_{\mathcal{M}_s^*}(s, \xi_A^*)$, $\mathcal{M}_s^*, \rho \models \varphi$. We assume that there is a total order $\preceq$ on set S^+, and denote by $\min(U)$ the minimal one of the set of states $U \subseteq S^+$ with respect to the order $\preceq$. Let ξ_A be the function such that for every $i \in A$ and $s' \in S$, $\xi_A(i)(s') = \xi_A^*(i)(\min(\{\rho \in S^+ \mid \mathtt{lst}(\rho) = s'\}))$.

First, we show that ξ_A is a collective strategy of A in $\mathcal{M}'$. Consider an agent $i \in A$ and two states $s_1, s_2 \in S$ such that $s_1 \sim_i s_2$, if $\pi(i) \neq \mathtt{IR}$, then for each pair of states $\rho_1, \rho_2 \in S^+$ such that $\mathtt{lst}(\rho_1) = s_1$ and $\mathtt{lst}(\rho_2) = s_2$, we have $\rho_1 \sim_i^* \rho_2$, which implies that $\xi_A^*(i)(\rho_1) = \xi_A^*(i)(\rho_2)$, hence $\xi_A(i)(s_1) = \xi_A(i)(s_2)$. Otherwise $s_1 = s_2$ and $\pi(i) = \mathtt{IR}$. We choose $\xi_A(i)(s_1) = \xi_A(i)(s_2) = \xi_A^*(i)(\min(\{\rho \in S^+ \mid \mathtt{lst}(\rho) = s_1\}))$. Therefore, ξ_A is a collective strategy of A in $\mathcal{M}'$.

Consider a collective strategy $\xi_{\overline{A}}$ of $\overline{A}$ in $\mathcal{M}'$, let $\rho = \mathtt{play}(s, \xi_A, \xi_{\overline{A}})$, then we have $\rho_{[0..0]}\rho_{[0..1]}\rho_{[0..2]}\cdots \in O_{\mathcal{M}_s^*}(s, \xi_A^*)$. The result immediately follows from the fact that $\lambda^*(\rho) = \lambda(\mathtt{lst}(\rho))$ for every $\rho \in S^+$. □

Recalling that [Alur *et al.*, 2002] observed that both semantics of ATL under $\mathtt{Ir}$-strategies and $\mathtt{IR}$-strategies coincide for *CGS*. This result was generalized and formally proved in infinite *CGS* (i.e., no finiteness with respect to the set of states and actions) (cf. Proposition 1 [Bulling and Jamroga, 2014]). Proposition 2 can be seen as a generalization of the result of [Alur *et al.*, 2002] and could be extended to the infinite ACGS similar to [Bulling and Jamroga, 2014].

C Effects of Strategy Types

By restricting all the strategy types to $\mathtt{IR}$, straightforwardly we have:

Proposition 3 *Let $\mathcal{M} = (\mathcal{G}, \pi)$ be an ACGS where for each $i \in Ag$, $\pi(i) = \mathtt{IR}$. For each state s of $\mathcal{M}$ and ATL* formula φ, $\mathcal{G}, s \models_{\mathtt{IR}} \varphi$ iff $\mathcal{M}, s \models \varphi$.*

Proof. By applying structural induction, it suffices to show that the result holds for formulae of the form $\langle\langle A\rangle\rangle\phi$. By applying induction hypothesis, for every path ρ, the following holds: $\mathcal{G},\rho \models_{\mathtt{IR}} \varphi$ iff $\mathcal{M},\rho \models \varphi$. For each pair $(\xi_A, \upsilon^\sigma_A)$ of collective strategies such that $\xi_A = \upsilon^\sigma_A$, $O_{\mathcal{M}}(s,\xi_A) = O^\sigma_{\mathcal{G}}(s,\upsilon^\sigma_A)$. Each $i \in A$ has same sets of possible IR-strategies in $\mathcal{G}$ and $\mathcal{M}$, hence $\mathcal{G},\rho \models_{\mathtt{IR}} \langle\langle A\rangle\rangle\phi$ iff $\mathcal{M},\rho \models \langle\langle A\rangle\rangle\phi$. □

In light of Proposition 1 and Proposition 3, we shall investigate the effects of strategy types by considering ACGS with various different setups of strategy types.

Given a set A of agents, for two functions $\pi_1,\pi_2 : Ag \to \mathrm{T}_{\mathtt{str}}$, π_1 is *coarser* than π_2 with respect to A, denoted by $\pi_1 \preceq_A \pi_2$, if for every $i \in A$, $\pi_1(i) = \pi_2(i)$ and for every $j \in \overline{A}$, one of the following conditions holds:

- $\pi_1(j) = \mathtt{IR}$, $\pi_2(j) = \mathtt{IR}$;
- $\pi_1(j) = \mathtt{Ir}$, $\pi_2(j) \in \{\mathtt{IR}, \mathtt{Ir}\}$;
- $\pi_1(j) = \mathtt{iR}$, $\pi_2(j) \in \{\mathtt{IR}, \mathtt{iR}\}$;
- $\pi_1(j) = \mathtt{ir}$, $\pi_2(j) \in \{\mathtt{IR}, \mathtt{Ir}, \mathtt{iR}, \mathtt{ir}\} = \mathrm{T}_{\mathtt{str}}$.

Lemma 4 *Let A be a set of agents and s be a state of a CGS $\mathcal{G}$. For two functions $\pi_1,\pi_2 : Ag \to \mathrm{T}_{\mathtt{str}}$ with $\pi_1 \preceq_A \pi_2$, and any collective strategy ξ_A of A, we have:*

$$O_{(\mathcal{G},\pi_1)}(s,\xi_A) \subseteq O_{(\mathcal{G},\pi_2)}(s,\xi_A).$$

Lemma 4 reveals the effect of strategy types of $\overline{A}$ on the outcomes. It is easy to observe that if $\pi_2(i) = \sigma$ for all $i \in A$, then for every collective σ-strategy υ^σ_A such that $\xi_A = \upsilon^\sigma_A$, we have that $O_{(\mathcal{G},\pi_2)}(s,\xi_A) \subseteq O^\sigma_{\mathcal{G}}(s,\upsilon^\sigma_A)$. Moreover, if $\pi_2(i) = \mathtt{IR}$ for all $i \in \overline{A}$, then $O_{(\mathcal{G},\pi_2)}(s,\xi_A) = O^\sigma_{\mathcal{G}}(s,\upsilon^\sigma_A)$.

An ATL* formula φ is *positive* if (1) for each occurrence of $\langle\langle A\rangle\rangle\phi$ in φ, ϕ is an LTL formula, (2) there is no occurrence of $[[A]]\phi$ in φ, and (3) negations $\neg$ only appear in front of atomic propositions. For example, $\langle\langle A\rangle\rangle\mathbf{X}\, q$ is positive, while $\neg\langle\langle A\rangle\rangle\mathbf{X}\, q$ is not positive. Given a formula φ, let Ag_φ denote the set of agents that appear in φ. By Lemma 4, we have:

Proposition 4 *Let $\mathcal{G}$ be a CGS, s be a state of $\mathcal{G}$ and φ be a positive ATL/ATL* formula. For $\pi_1,\pi_2 : Ag \to \mathrm{T}_{\mathtt{str}}$ such that $\pi_1 \preceq_{Ag_\varphi} \pi_2$, if $(\mathcal{G},\pi_2), s \models \varphi$, then $(\mathcal{G},\pi_1), s \models \varphi$.*

Proof. By applying structural induction, it suffices to show that the result holds for formulae of the form $\langle\langle A\rangle\rangle\phi$. We suppose $(\mathcal{G},\pi_2), s \models \langle\langle A\rangle\rangle\phi$, otherwise Item 1 immediately holds.

Since $(\mathcal{G},\pi_2), s \models \langle\langle A\rangle\rangle\phi$, then there exists a collective strategy ξ_A of agents A such that for each path $\rho \in O_{(\mathcal{G},\pi_2)}(s,\xi_A)$, $(\mathcal{G},\pi_2),\rho \models \phi$ holds. Since $A \subseteq Ag_\varphi$ and for every $i \in Ag_\varphi$, $\pi_1(i) = \pi_2(i)$ and $\pi_1 \preceq_{Ag_\varphi} \pi_2$, then $\pi_1 \preceq_A \pi_2$. By Lemma 4, we get that $O_{(\mathcal{G},\pi_1)}(s,\xi_A) \subseteq O_{(\mathcal{G},\pi_2)}(s,\xi_A)$.

By applying induction hypothesis, for every state formula ψ in ϕ and state s' of $\mathcal{G}$, if $(\mathcal{G},\pi_2), s' \models \psi$, then $(\mathcal{G},\pi_1), s' \models \psi$. Therefore, for each path $\rho \in O_{(\mathcal{G},\pi_1)}(s,\xi_A)$, we have $(\mathcal{G},\pi_1),\rho \models \phi$. The result follows.

□

For positive ATL/ATL* formulae φ, even if the agents of Ag_φ have the same strategy types in ACGS $(\mathcal{G},\pi)$ and CGS $\mathcal{G}$, verifying $\mathcal{G}$ against φ under σ will examine more behavior than verifying $(\mathcal{G},\pi)$ against φ, where $i \in Ag_\varphi$ and $\pi(i) = \sigma$. Therefore, if the behavior of a MAS is exactly modeled as an ACGS $\mathcal{M}$ rather than a CGS $\mathcal{G}$ with strategy type σ, verifying $\mathcal{G}$ against φ under σ may lead to incorrect result. However, more restrictions on strategy types and ATL/ATL* formulae can make them coincide, as the following proposition shows.

Proposition 5 *Let s be a state of $\mathcal{M} = (\mathcal{G},\pi)$ and $\sigma \in \mathrm{T}_{\mathtt{str}}$ be a strategy type. Assume ATL/ATL* formula φ satisfies (1) for every $i \in Ag_\varphi$, $\pi(i) = \sigma$, (2) for every $i \in Ag \setminus Ag_\varphi$, $\pi(i) = \mathtt{IR}$, and (3) for every occurrence of $\langle\langle A'\rangle\rangle\phi$ in φ, $Ag_\varphi = A'$. Then we have $\mathcal{G}, s \models_\sigma \varphi$ iff $\mathcal{M}, s \models \varphi$.*

D More Experimental Results

D.1 Experiment on Dining Cryptographers Protocol

Dining cryptographers protocol is one of anonymity protocols aimed at establishing the privacy of principals during an exchange [Lomuscio *et al.*, 2017]. The dining cryptographers protocol can be modeled as MAS. In this game, n cryptographers share a meal around a circular table. Either one of them or their employer paid for the meal. They are curious whether it was sponsored by their employer without revealing the identity of the payer (if one of them did pay). The protocol works as follows: each cryptographer 1) tosses a coin and shows the outcome to his right-hand neighbour, 2) announces whether the two coins agree or not if he/she is not payer, otherwise announces the opposite of what he/she sees. Their employer is the payer if an even number of cryptographers claiming that the two coins are different, otherwise not. For experimental purpose, we allow the cryptographer who paid for the meal announces either the two coins agree or not no matter what he/she saw.

In this experiment, n ranges from 3 to 10, a cryptographer whose is payer and one of other cryptographers use ir-strategies, the others all use IR-strategies. We verify three formulae ψ_1, ψ_2 and ψ_3, where ψ_i expresses that if the number of "saydifferent"

Table 2: Results of dining cryptographers protocol

#Crypts	#States	Lemma 1			Lemma 3		
		ψ_1	ψ_2	ψ_3	ψ_1	ψ_2	ψ_3
3	160	0.022	0.016	0.013	6.439	5.838	5.852
4	384	0.059	0.049	0.028	6.928	6.744	7.242
5	896	0.133	0.114	0.049	8.839	8.874	8.88
6	2048	0.315	0.328	0.163	12.567	12.724	12.865
7	4608	0.929	1.388	0.382	22.938	23.411	23.654
8	10240	3.463	4.022	0.834	60.642	60.583	63.064
9	22528	9.19	8.913	1.721	266.844	240.003	254.293
10	49152	21.988	21.927	5.094	1712.62	1588.06	1762.88

Table 3: Results of BSS

π	φ_1			φ_2		
	Lemma 1	Lemma 3	SAT	Lemma 1	Lemma 3	SAT
$(\mathtt{IR}, \mathtt{IR})$	4.237	11.264	Y	0.08	5.566	Y
$(\mathtt{IR}, \mathtt{Ir})$	4.102	12.185	Y	0.081	5.124	Y
$(\mathtt{IR}, \mathtt{ir})$	4.094	11.459	Y	0.081	5.26	Y
$(\mathtt{Ir}, \mathtt{IR})$	4.095	17.398	Y	0.081	6.096	Y
$(\mathtt{Ir}, \mathtt{Ir})$	4.086	30.649	Y	0.082	7.183	Y
$(\mathtt{Ir}, \mathtt{ir})$	4.112	32.985	Y	0.082	8.009	Y
$(\mathtt{ir}, \mathtt{IR})$	4.162	17.842	N	0.082	5.96	Y
$(\mathtt{ir}, \mathtt{Ir})$	4.144	31.155	N	0.082	7.592	Y
$(\mathtt{ir}, \mathtt{ir})$	4.157	30.73	N	0.082	7.473	Y

is odd and the i-th cryptographer is not the payer, then he/she knows that the bill is paid by one of the others, but cannot tell exactly who is the payer. For instance, in the three cryptographers case,

$$\psi_1 \equiv \langle\langle\emptyset\rangle\rangle\mathbf{G}((odd \wedge \neg c1paid) \rightarrow ((\mathbf{K}_{c1}(c2paid \vee c3paid)) \wedge \neg\mathbf{K}_{c1}c2paid \wedge \neg\mathbf{K}_{c1}c3paid)).$$

The results are shown in Table 2, where column 1 gives the number of cryptographers, column 2 gives the number of states, columns 3–5 (resp. columns 6–7) show the total time of respectively verifying ψ_1, ψ_2 and ψ_3 using Lemma 1 (resp. Lemma 3). Both ψ_1 and ψ_2 are satisfied by all the models, while ψ_3 not. We observe that Lemma 1 is more efficient than Lemma 3, as the coalitions in all the formulae are $\emptyset$. From this experiment, one may conclude the reasonable scalability of our tool.

D.2 Experiment on Book Store Scenario (BSS)

The BSS model depicts a deal between two agents: a supplier (S) and a purchaser (P) [Lomuscio *et al.*, 2017]. Initially, S is waiting for an order from P, and P is ready for initiating a trade. Upon receiving an order of some e-good from P, S can make a decision whether accepts the order or not do, and later notify P. If S accepts, then P can pay fee. Once paid, S can either reject the payment or accept and deliver the good. If P received the good, then trade is completed. During the trade, P can revoke the order, both S and P can terminate the trade, after which the information of the trade should be symmetric at any time. In this model, the S has 15 local states and 13 actions, and P has 12 local states and 7 actions. In this experiment, we verify the model against the following formulae.

$$\varphi_1 \equiv \langle\langle\emptyset\rangle\rangle\mathbf{G}((S\&P\_no\_T) \rightarrow (\mathbf{K}_S\langle\langle\{S,P\}\rangle\rangle\mathbf{F}\ trade\_end))$$

expresses that "if neither S nor P terminates the trade (i.e., $S\&P\_no\_T$ is ture), then S knows that they can cooperatively complete the trade eventually" always holds.

$$\varphi_2 \equiv \langle\langle\{S,P\}\rangle\rangle(S\&P\_no\_T\ \mathbf{U}(trade\_end \wedge \neg trade\_success))$$

expresses that the trade can ends by P asking for refund.

The results are shown in Table 3. Each row presents the result of one of strategy type combinations of S and P, for instance, $(\mathtt{IR}, \mathtt{ir})$ denotes that S has an $\mathtt{IR}$-strategy while P has $\mathtt{ir}$-strategy. Columns 2-4 (resp. columns 5-7) show total time and result of verifying φ_1 (resp. φ_2) using Lemma 1 (resp. Lemma 3). The results of φ_1 confirm that strategy types affect the truth of formulae. Lemma 1 performs better than Lemma 3 both on φ_1 and φ_2 in this experiment.