Knowledge and Strategic Ability for Model Checking: A Refined Approach

Wojciech Jamroga

Department of Informatics, Clausthal University of Technology, Germany
wjamroga@in.tu-clausthal.de

Abstract. We present a translation that reduces epistemic operators to strategic operators in the context of model checking. The translation is a refinement of the one from [4], and it improves on the previous scheme in two ways. First, it does not suffer any blowup in the length of formulae (the one from [4] did). Second, the new translation is defined in a more general setting: additional constraints can be imposed on strategy profiles that agents can execute. We show the applicability of such a general translation on the case of strategic abilities under imperfect information.

1 Introduction

Modal logics of multi-agent systems usually combine several dimensions. Knowledge, time, actions, strategic abilities, norms/obligations, intentions, desires etc. can all be involved in a description of an agent system. This way, modal logic can support sufficiently realistic descriptions of agents. But there is a price to pay: such multi-modal logics are usually harder to handle semantically as well as algorithmically. Thus, a designer is usually faced with the task of finding a good tradeoff between a "clean" logic with few modalities (and clear overall semantics) and a "realistic" language with many modalities (where it is not immediately visible how parts of the semantics interfere). A reduction method that allows to express one modality with the others offers two kinds of advantage. In terms of theory, it allows to make the logic "cleaner", and study its theoretical properties (semantics, computational complexity) in a simpler environment. On the practical side, we can reuse the advances in, say, model checking of one sort of modality to improve the techniques used for dealing with the other dimensions.

In [4], we proposed how epistemic modalities can be equivalently expressed by strategic operators of alternating-time temporal logic ATL [1] in the context of model checking. The reduction was polynomial in *almost* every respect. Unfortunately, the length of formulae could suffer exponential blowup (although the number of *different* subformulae in the formula increased only linearly). We argued that, for most model checking algorithms, it would not increase the verification time. Still, it was a flaw that made using the reduction awkward, at least for theoretical purposes. The aim of this paper is to propose a refinement of the reduction that does not suffer from the blowup any more. Moreover, we point out that the reduction can be used even if we impose some "behavioral constraints" on the strategies that can be played by agents. Thus, the method can be used also for variants of ATL where one assumes that the agents can only play in a uniform [7], socially acceptable [11], or rational way [6].

Our presentation here is based on some material from [4]. It should be also mentioned that the original reduction was inspired by [9], and shared some similarities with [13] (although the reduction proposed in the latter paper had a more limited scope). Similar translations of modal logics include [8, 3]. Our presentation of strategic constraints is based on the approach of [6].

2 Preliminaries

2.1 ATL: Abilities in Perfect Information Games

ATL [1] generalizes the branching time logic CTL [2] by replacing path quantifiers with so called *cooperation modalities*. The formula $\langle\langle A\rangle\rangle\varphi$ expresses that group of agents A have a collective strategy to enforce φ. ATL formulae include temporal operators: "$\bigcirc$" ("in the next state"), $\Box$ ("always from now on") and $\mathcal{U}$ ("until"). Operator $\Diamond$ ("now or sometime in the future") can be defined as $\Diamond\varphi \equiv \mathrm{T}\,\mathcal{U}\,\varphi$. Formally, the recursive definition of ATL formulae is:

$$\varphi ::= p \mid \neg\varphi \mid \varphi\wedge\varphi \mid \langle\langle A\rangle\rangle\bigcirc\varphi \mid \langle\langle A\rangle\rangle\Box\varphi \mid \langle\langle A\rangle\rangle\varphi\,\mathcal{U}\,\varphi.$$

A *concurrent game structure* (CGS) is a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o\rangle$ which includes a nonempty finite set of all agents $\mathbb{A}\mathrm{gt} = \{1,\dots,k\}$, a nonempty set of states St, a set of atomic propositions Π, a valuation of propositions $\pi : St \to 2^{\Pi}$, and a set of (atomic) actions Act. Function $d : \mathbb{A}\mathrm{gt} \times St \to (2^{Act}\setminus\emptyset)$ defines nonempty sets of actions available to agents at each state, and o is a transition function that assigns the outcome state $q' = o(q,\alpha_1,\dots,\alpha_k)$ to state q and a tuple of actions $\langle\alpha_1,\dots,\alpha_k\rangle$, $\alpha_i \in d(i,q)$, that can be executed by $\mathbb{A}\mathrm{gt}$ in q.

A (memoryless) *strategy* s_a of agent a is a conditional plan that specifies what a is going to do for every possible situation: $s_a : St \to Act$ such that $s_a(q) \in d(a,q)$. We denote the set of such functions by Σ_a. A *collective strategy* s_A for a group of agents A is a tuple of strategies, one per agent from A; the set of A's collective strategies is given by $\Sigma_A = \prod_{a\in A}\Sigma_a$. The set of all *strategy profiles* is given by $\Sigma = \Sigma_{\mathbb{A}\mathrm{gt}}$.

A *path* λ in model M is an infinite sequence of states that can be effected by subsequent transitions, and refers to a possible course of action (or a possible computation) that may occur in the system; by $\lambda[i]$, we denote the ith position on path λ. The set of all paths starting from state q is given by $\Lambda(q)$. Function $out(q, s_A)$ returns the set of all paths that may result from agents A executing strategy s_A from state q onward.

Formally, the semantics of cooperation modalities can be given via the following clauses:

$M, q \models \langle\langle A\rangle\rangle\bigcirc\varphi$ iff there is a collective strategy s_A such that, for every $\lambda \in out(q, s_A)$, we have $M, \lambda[1] \models \varphi$;

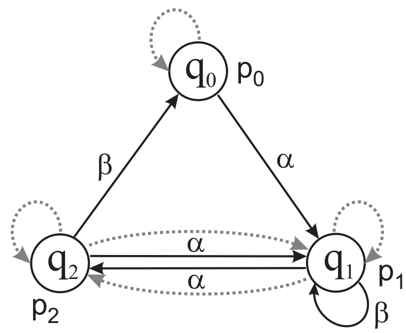

Fig. 1. Simple concurrent epistemic game structure M_1. Nodes represent states of the system, solid arrows depict transitions (labeled by the agent's actions), and dotted arrows show indistinguishability of states.

$M, q \models \langle\langle A \rangle\rangle \Box \varphi$ iff there exists s_A such that, for every $\lambda \in out(q, s_A)$, we have $M, \lambda[i] \models \varphi$ for every $i \geq 0$;
$M, q \models \langle\langle A \rangle\rangle \varphi \mathcal{U} \psi$ iff there exists s_A such that for every $\lambda \in out(q, s_A)$ there is an $i \geq 0$, for which $M, \lambda[i] \models \psi$, and $M, \lambda[j] \models \varphi$ for every $0 \leq j < i$.

2.2 Epistemic Logic: Knowledge and Imperfect Information

Epistemic logic uses operators $K_a\varphi$ ("agent a knows that φ"). Additional operators $E_A\varphi$, $C_A\varphi$, and $D_A\varphi$, where A is a set of agents, refer to *mutual knowledge* ("everybody knows"), *common knowledge*, and *distributed knowledge* among the agents from A. On the semantic side, uncertainty of agents is modeled by indistinguishability relations $\sim_1, \ldots, \sim_k \subseteq St \times St$ (one per agent). The semantics of K_a is defined as: $M, q \models K_a\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_a q'$.

Relations $\sim_A^E$, $\sim_A^C$ and $\sim_A^D$, used to model group epistemics, are derived from the individual relations of agents from A. First, $\sim_A^E$ is the union of relations $\sim_a$, $a \in A$. Next, $\sim_A^C$ is defined as the transitive closure of $\sim_A^E$. Finally, $\sim_A^D$ is the intersection of all the $\sim_a$, $a \in A$. Then, for $\mathcal{K} = C, E, D$, we define: $M, q \models \mathcal{K}_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^{\mathcal{K}} q'$.

A straightforward combination of ATL and epistemic logic, called ATEL was introduced in [12]. The language of ATEL allows to express knowledge about agents' (perfect information) abilities. Models of ATEL are called *concurrent epistemic game structures* (CEGS). A simple CEGS (with only one agent a) is depicted in Figure 1. For that model, we have for instance that $M_1, q_1 \models K_a \langle\langle a \rangle\rangle \Diamond \mathsf{p}_0$.

3 Restricting Strategies of Agents

In many cases, it seems appropriate to put some constraints on the "good" (allowed, legal etc.) behaviors. We define a class of such *strategic constraints* in this section. Our constraints are based on the idea of plausibility sets [6], and generalize the behavioral constraints from the framework of social laws [11].

3.1 Strategic Constraints

A behavioral constraint in [11] is a function $\beta : \mathbb{A}\mathrm{gt} \times St \to 2^{Act}$ that specifies which actions can be "legally" played by agents. More specifically, $\beta(a,q)$ is the set of actions that a is allowed to play at state q. Naturally, $\beta(a,q) \subseteq d(a,q)$, and the inclusion can be strict. $\beta(a,q)$ is assumed to implement a social norm: agent a (when in state q) may be forbidden to play some actions in his repertoire; if he decides to play them, he will violate the norm.

Note that using constraints of this type implies that norms only apply to actions of individual agents (independently). It is therefore not possible to specify e.g. that one is allowed to shoot in self-defense, i.e., *right at the moment* when another person is trying to harm him. Likewise, norms of that type specify legal actions independently for each state. Thus, if we do not accept lying, then making a false statement will be always forbidden, even if it is just a joke, and the speaker is going to disclose the truth in the very next moment.

Here, we are looking for a model that enables to cope with such interrelationships between the allowed actions of different agents at different states, too. Another rationale for this comes from game theory. Unlike in normative systems, we are interested in "rational" rather than "moral" behavior there, but the general pattern is the same. That is, some strategy profiles of agents (e.g., those in Nash equilibrium) are deemed "rational", while the others are rejected as "irrational". Note that, especially for Nash equilibrium, the rationality of an action does depend on what the agent is going to do at other states; moreover, it depends on what *the other agents* are going to do at this and other states. Thus, our requirements with respect to agents' behavior will be modeled as *sets of strategy profiles*.

When defining agents' behavior via strategy sets, one assumes implicitly that agents actually *play strategies*. In our case, it would for instance imply that each agent does the same action every time the system comes back again to one of the previous states (as memoryless strategies are used in our semantics of ATL). This is a very strong assumption, and we do not always want to make it with respect to *all* agents. Thus, our strategic constraints will also include the set of agents to whom the constraint should apply.

Definition 1. *A* strategic constraint *is a pair $\eta = \langle \Upsilon, A \rangle$, where $\Upsilon \subseteq \Sigma$ is a non-empty set of strategy profiles and $A \subseteq \mathbb{A}\mathrm{gt}$ is a set of agents.*

Definition 2 (Substrategy). *Let $A, B \subseteq \mathbb{A}\mathrm{gt}$, and let s_A be a collective strategy for A. We use $s_A[B]$ to denote the* substrategy *of s_A for agents from B only, i.e., strategy $t_{A\cap B}$ such that $t^a_{A\cap B} = s^a_A$ for every $a \in A \cap B$. We extend the notation to sets in a natural way: for a set of collective strategies $\Upsilon_A \subseteq \Sigma_A$, we define $\Upsilon_A[B] = \{t \in \Sigma_{A\cap B} \mid \exists s_A \in \Upsilon_A . t = s_A[B]\}$.*

Definition 3 (Consistency with a constraint). *Let s_A be a collective strategy of $A \subseteq \mathbb{A}\mathrm{gt}$, and $\eta = \langle \Upsilon, B \rangle$ be a strategic constraint. Strategy s_A is* consistent with constraint *η iff the part of s_A to which the constraint should apply occurs in Υ, i.e., $s_A[B] \in \Upsilon[A \cap B]$.*

Definition 4 (Outcome under constraint). *Let M be a* CGS*, and q a state in M. Furthermore, let s_A be a collective strategy, and $\eta = \langle \Upsilon, B\rangle$ be a strategic constraint. The outcome of s_A from q under constraint η contains all paths which may result from agents A executing s_A from q on, when the opponents are only allowed to play strategies which complement s_A in a way that complies with η. Formally, the set is defined as:*

$out(q, s_A, \eta) = \{\lambda \in \Lambda(q) \mid$ *there is* $t \in \Sigma_{A\cup B}$*, consistent with* η*, such that* $t[A] = s_A$ *and for every* $i = 1, 2, \ldots$ *there exists a tuple of agents' decisions* $\langle \alpha_1, \ldots, \alpha_k\rangle$ *for which:* $\alpha_a = t^a(\lambda[i-1])$ *for* $a \in A \cup B$*,* $\alpha_a \in d(a, \lambda[i-1])$ *for* $a \notin A \cup B$*, and* $o(\lambda[i-1], \alpha_1, \ldots, \alpha_k) = \lambda[i]\}$.

3.2 Abilities under Strategic Constraints: Semantics

The intuition behind strategic constraints is rather simple: for a constraint $\eta = \langle \Upsilon, B\rangle$ we assume that the actual collective strategy of agents B must occur somewhere in Υ. Note that the agents from B do not have to be all in the same coalition – B can collect both "proponents" and "opponents". The formal semantics of ATL formulae in the presence of strategic constraints is given by the clauses below.

$M, q, \eta \models p$ iff $p \in \pi(q)$ (for $p \in \Pi$);
$M, q, \eta \models \neg\varphi$ iff $M, q, \eta \not\models \varphi$;
$M, q, \eta \models \varphi \wedge \psi$ iff $M, q, \eta \models \varphi$ and $M, q, \eta \models \psi$;
$M, q, \eta \models \langle\langle A\rangle\rangle \bigcirc \varphi$ iff there is a collective strategy s_A, consistent with η, such that for every $\lambda \in out(q, s_A, \eta)$ we have $M, \lambda[1], \eta \models \varphi$;
$M, q, \eta \models \langle\langle A\rangle\rangle \Box \varphi$ iff there exists s_A consistent with η, such that for every $\lambda \in out(q, s_A, \eta)$ we have $M, \lambda[i], \eta \models \varphi$ for every $i \geq 0$;
$M, q, \eta \models \langle\langle A\rangle\rangle \varphi \mathcal{U} \psi$ iff there exists s_A consistent with η, such that for every $\lambda \in out(q, s_A, \eta)$ there is an $i \geq 0$, for which $M, \lambda[i], \eta \models \psi$, and $M, \lambda[j], \eta \models \varphi$ for every $0 \leq j < i$.

The semantics of knowledge under strategic constraints is defined in a straightforward way: agents know that φ under η iff φ holds under η in every indistinguishable state.

$M, q, \eta \models K_a \varphi$ iff $M, q', \eta \models \varphi$ for every q' such that $q \sim_a q'$.
$M, q, \eta \models \mathcal{K}_A \varphi$ iff $M, q', \eta \models \varphi$ for every q' such that $q \sim_A^{\mathcal{K}} q'$ (where $\mathcal{K} = C, E, D$).

A useful example of strategic constraints are so called *uniform strategies*, i.e., strategies that can be feasibly executed by an agent under imperfect information. We say that s_a is uniform iff, for every q, q', $q \sim_a q'$ implies that $s_a(q) = s_a(q')$; that is, agent a must specify same choices in states that look the same to him. A collective strategy s_A is uniform iff it consists only of uniform individual strategies. Let Σ_a^u denote the set of uniform strategies of agent a. Then $\Sigma_A^u = \prod_{a\in A} \Sigma_a^u$ is the set of collective uniform strategies of A, and $\Sigma^u = \Sigma_{\mathbb{A}\mathrm{gt}}^u$ is the set of uniform strategy profiles. Now, the requirement that agents from A should only use uniform strategies can be captured by the strategic constraint $\eta = \langle \Sigma^u, A \rangle$.

Consider CEGS M_1 from Figure 1. For that model, the requirement that the only agent sticks to executable (i.e., uniform) strategies can be captured by the constraint $\eta = \langle \{[q_0 \mapsto \alpha, q_1 \mapsto \alpha, q_2 \mapsto \alpha], [q_0 \mapsto \alpha, q_1 \mapsto \beta, q_2 \mapsto \beta]\}, \{a\} \rangle$. Then, we have for instance that $M_1, q_1, \eta \models K_a \neg \langle\langle a \rangle\rangle \Diamond \mathsf{p_0}$: no uniform strategy can guarantee that a gets from q_1 to q_0, and the agent knows about it.

4 Translating Knowledge to Strategic Ability

In this section, we show a satisfaction-preserving interpretation of ATEL formulae and models into ATL. The interpretation is an update of that proposed in [4]. Two things are changed. First, we slightly change the transformation of models so that, after visiting an "epistemic" state, the system *always* returns immediately to its corresponding "action" state. In consequence, it is possible to define the translation of formulae without exponential blowup in their length. Second, we show that the translation is also correct when we add constraints on the behavior of agents.

4.1 Idea of the Translation

ATEL consists of two orthogonal layers. The first one, inherited from ATL, refers to what agents can achieve in temporal perspective, and is underpinned by the structure defined via transition function o. The other layer is the epistemic component, reflected by epistemic indistinguishability relations. Our idea of the translation is to leave the original temporal structure intact, while extending it with additional transitions to "simulate" epistemic links. The simulation is achieved through adding new "epistemic" agents who can enforce transitions to special "epistemic" copies of "action" states (i.e., the states inherited from the original model). The "action" and "epistemic" states form separate strata in the resulting model, and are labeled accordingly to distinguish transitions that implement different modalities.

The interpretation consists of two independent parts: a transformation of models and a translation of formulae. First, we propose a construction that transforms every concurrent epistemic game structure M for a set of agents $\{1, ..., k\}$, into a (pure) concurrent game structure M' over a set of agents $\{1, ..., k, e_1, ..., e_k\}$. Agents $1, ..., k$ are the original agents from M (we will call them "real agents"). Agents $e_1, ..., e_k$ are "epistemic doubles" of the real agents: the role of e_i is to "point out" the states that were epistemically indistinguishable from the current state for agent i in M. In order to distinguish transitions referring to different modalities, we introduce additional states in model M'. States $q_1^{\mathsf{e_i}}, ..., q_n^{\mathsf{e_i}}$ satisfy new proposition $\mathsf{e_i}$ added to enable identifying moves of epistemic agent e_i. Moreover, epistemic state $q^{\mathsf{e_i}}$ has the same "epistemic" transitions as q (leading to epistemic copies of states indistinguishable from q), plus one outgoing transition leading to the corresponding action state q. The original states $q_1, ..., q_n$ are still in M' to represent targets of "action" moves of the real agents $1, ..., k$. We will use a new proposition act to label these states. Now, the type of a transition can be recognized by the label of its target state. The structure of the transformation can be seen in Figure 2.

Defining the transition function o so that both epistemic and "action" transitions can occur is the trickiest part of the construction. We achieve this by giving priority to the epistemic agents' decisions. Every epistemic agent can choose to be "passive" and let the others decide upon the next move, or may try to effect an epistemic move. The resulting transition leads to the state selected by the *first* non-passive epistemic agent. If all the epistemic agents have decided to be passive, the action transition chosen by the real agents follows. Epistemic states are given special treatment, as we assume that the real agents are always passive there. Thus, if all the epistemic agents decide to be passive at an epistemic state, the system proceeds to the corresponding action state.

With the above construction in mind, ATEL formulae can be translated to ATL according to the following scheme:

- $K_i\varphi$ can be rephrased as $\neg\langle\langle e_1, ..., e_i\rangle\rangle \bigcirc (\mathsf{e_i} \wedge \langle\langle e_1, ..., e_k\rangle\rangle \bigcirc (\mathsf{act} \wedge \neg\varphi))$: the epistemic moves to agent e_i's epistemic states do not lead to a state where φ fails (more precisely: where φ fails in the corresponding "action" state). Note that player e_i can select a state of his if, and only if, players $e_1, ..., e_{i-1}$ are passive (hence their presence in the cooperation modality).
- $\langle\langle A\rangle\rangle \bigcirc \varphi$ becomes $\langle\langle A \cup \{e_1, ..., e_k\}\rangle\rangle \bigcirc (\mathsf{act} \wedge \varphi)$ in a similar way.
- Translation of the other temporal operators is now more straightforward than in [4]: $\langle\langle A\rangle\rangle \Box\varphi$ can be rephrased as $\langle\langle A \cup \{e_1, ..., e_k\}\rangle\rangle \Box(\mathsf{act} \wedge \varphi)$, and $\langle\langle A\rangle\rangle \varphi \mathcal{U} \psi$ becomes $\langle\langle A \cup \{e_1, ..., e_k\}\rangle\rangle (\mathsf{act} \wedge \varphi) \mathcal{U} (\mathsf{act} \wedge \psi)$. This is possible because the construction of epistemic states (and the translation of K_a) ensures that strategic (sub)formulae will be always evaluated in "action" states. We observe that the new translation of $\Box$ and $\mathcal{U}$ does not involve exponential increase in the length of formulae (contrary to the construction from [4]).
- Translation of mutual knowledge (E_A) is analogous to the individual knowledge case. Translation of common knowledge refers to the definition of relation $\sim_A^C$ as the transitive closure of relations $\sim_i$ for $i \in A$: $C_A\varphi$ means that all the (finite) sequences of appropriate epistemic transitions must end up in a state where φ is true.

The only operator that does not seem to lend itself to a translation according to the above scheme is the distributed knowledge operator D_A, for which we seem to need more "auxiliary" agents. Thus, we will begin with presenting details of our interpretation for ATEL$_{CE}$ – a reduced version of ATEL that includes only common knowledge and "everybody knows" operators for group epistemics. Section 4.3 shows how to modify the translation to include distributed knowledge as well.

4.2 Interpreting Models and Formulae of ATEL$_{CE}$ into ATL

Transforming Models Given a concurrent epistemic game structure $M = \langle \mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, o, \sim_1, ..., \sim_k \rangle$, we construct a new concurrent game structure $M' = \langle \mathbb{A}\mathrm{gt}', St', \Pi', \pi', Act', d', o' \rangle$ as follows:

- $\mathbb{A}\mathrm{gt}' = \mathbb{A}\mathrm{gt} \cup \mathbb{A}\mathrm{gt}^e$, where $\mathbb{A}\mathrm{gt}^e = \{e_1, ..., e_k\}$ is the set of epistemic agents;
- $St' = St \cup St^{\mathsf{e}_1} \cup ... \cup St^{\mathsf{e}_k}$, where $St^{\mathsf{e}_i} = \{q^{\mathsf{e}_i} \mid q \in St\}$.
- $\Pi' = \Pi \cup \{\mathsf{act}, \mathsf{e}_1, ..., \mathsf{e}_k\}$;
- $\pi'(p) = \pi(p)$ for every proposition $p \in \Pi$. Moreover, $\pi'(\mathsf{act}) = St$ and $\pi'(\mathsf{e}_i) = St^{\mathsf{e}_i}$;
- $Act' = Act \cup St \cup \{pass\}$: the new model M' contains the original actions from M, plus epistemic actions (pointing indistinguishable states), and the "do nothing" action $pass$;
- $d'_a(q) = d_a(q)$ for $a \in \mathbb{A}\mathrm{gt}$, $q \in St$; $d'_a(q) = \{pass\}$ for $a \in \mathbb{A}\mathrm{gt}$, $q \in St' \setminus St$; $d'_{e_i}(q) = \mathrm{img}(q, \sim_i) \cup \{pass\}$ for $q \in St'$;
- the new transition function is defined as follows:

$$o'(q, \alpha_1, ..., \alpha_k, \alpha_{e_1}, ..., \alpha_{e_k}) = \begin{cases} o(q, \alpha_1, ..., \alpha_k) & \text{if } q \in St \text{ and} \\ & \alpha_{e_1} = ... = \alpha_{e_k} = pass \\ \\ q_0 & \text{if } q = q_0^{\mathsf{e}_i} \in St^{\mathsf{e}_i} \text{and} \\ & \alpha_{e_1} = ... = \alpha_{e_k} = pass \\ \\ (\alpha_{e_i})^{e_i} & \text{if } e_i \text{ is the first active} \\ & \text{epistemic agent.} \end{cases}$$

We assume that all the epistemic agents from $\mathbb{A}\mathrm{gt}^e$, states from $St^{\mathsf{e}_1} \cup ... \cup St^{\mathsf{e}_k}$, and propositions from $\{\mathsf{act}, \mathsf{e}_1, ..., \mathsf{e}_k\}$, are new and were absent in the original model M.

The transformation of the simple CEGS from Figure 1 is shown in Figure 2.

Translation of Formulae Now, we define a translation of formulae from ATEL$_{CE}$ to ATL corresponding to the above transformation of models:

$$\begin{aligned}
tr(p) &= p, \qquad \text{for } p \in \Pi \\
tr(\neg\varphi) &= \neg tr(\varphi) \\
tr(\varphi \wedge \psi) &= tr(\varphi) \wedge tr(\psi) \\
tr(\langle\!\langle A \rangle\!\rangle \bigcirc \varphi) &= \langle\!\langle A \cup \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc (\mathsf{act} \wedge tr(\varphi)) \\
tr(\langle\!\langle A \rangle\!\rangle \Box \varphi) &= \langle\!\langle A \cup \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \Box (\mathsf{act} \wedge tr(\varphi)) \\
tr(\langle\!\langle A \rangle\!\rangle \varphi \mathcal{U} \psi) &= \langle\!\langle A \cup \mathbb{A}\mathrm{gt}^e \rangle\!\rangle (\mathsf{act} \wedge tr(\varphi)) \mathcal{U} (\mathsf{act} \wedge tr(\psi)) \\
tr(K_i\varphi) &= \neg\langle\!\langle e_1, ..., e_i \rangle\!\rangle \bigcirc \big(\mathsf{e}_i \wedge \langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc (\mathsf{act} \wedge \neg tr(\varphi))\big) \\
tr(E_A\varphi) &= \neg\langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc \big((\bigvee_{a_i \in A} \mathsf{e}_i) \wedge \langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc (\mathsf{act} \wedge \neg tr(\varphi))\big) \\
tr(C_A\varphi) &= \neg\langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc \langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \\
&\qquad (\bigvee_{a_i \in A} \mathsf{e}_i) \,\mathcal{U} \big((\bigvee_{a_i \in A} \mathsf{e}_i) \wedge \langle\!\langle \mathbb{A}\mathrm{gt}^e \rangle\!\rangle \bigcirc (\mathsf{act} \wedge \neg tr(\varphi))\big).
\end{aligned}$$

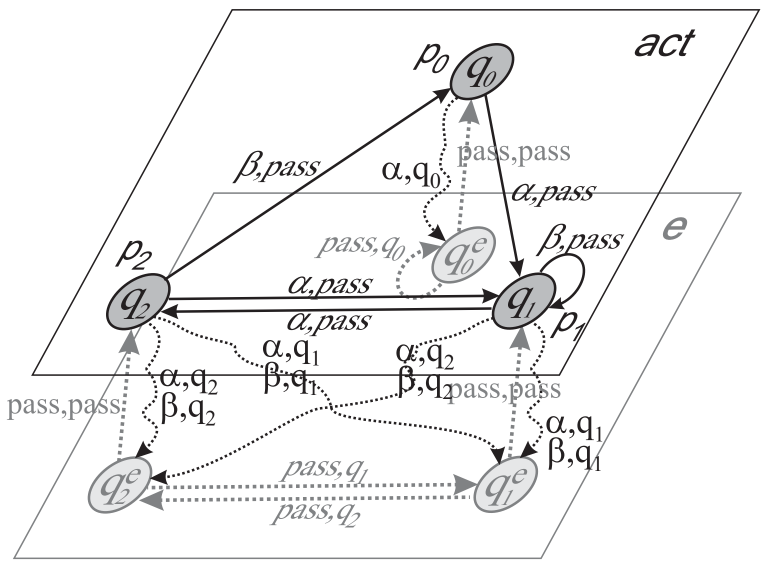

Fig. 2. Reconstruction for the concurrent epistemic game structure from Figure 1.

Extending Strategic Constraints Given a strategic constraint $\eta = \langle \Upsilon, B \rangle$ in M, we must extend it to match the type of constraints in M' (because M' includes more agents than M, and in consequence the elements of Υ, which are full strategy profiles in M, are only partial profiles in M').[1] This can be done in many ways; here, we explicitly assume that the additional (epistemic) agents can use any strategies they like. The new constraint must apply to the agents from B, plus (possibly) to some of the new agents from $\mathbb{A}\mathrm{gt}^e$. That is, agents from B are constrained in the same way as before, agents from $\mathbb{A}\mathrm{gt} \setminus B$ are unconstrained in the same way as before, and the new agents can be put under constraints or not – but even if they are, they can play any available strategy.[2]

Definition 5. *Let $\eta = \langle \Upsilon, B \rangle$ be a strategic constraint in concurrent epistemic game structure M, and let M' be the concurrent game structure obtained from M by the construction presented in Section 4.2. We say that constraint $\eta' = \langle \Upsilon', B' \rangle$* extends η in M' *iff: (1) $\Upsilon' = \Upsilon \times \Sigma_{\mathbb{A}\mathrm{gt}^e}$, and (2) $B \subseteq B' \subseteq B \cup \mathbb{A}\mathrm{gt}^e$.*

Soundness and Complexity of the Translation

Theorem 1. *Let φ be a formula of* ATEL_{CE}*, M be a* CEGS*, $q \in St$ a state in M, and M' the* CGS *resulting from the transformation. Furthermore, let η be a*

---

[1] Note that the old agents from $\mathbb{A}\mathrm{gt}$ have no real choice in the new states ($St' \setminus St$), so extending the set of states is not a problem (for every $s_a : St \to Act$ there is a unique $s'_a : St' \to Act$ that extends s_a).

[2] We recall that the assumption that a player plays a memoryless strategy is itself a restriction on the agent's behavior.

behavioral constraint in M, and let η' extend η in M'. Then, $M, q, \eta \models \varphi$ iff $M', q, \eta' \models tr(\varphi)$.

A proof of the theorem can be found in the technical report [5].

Note that the construction used above has several nice complexity properties. In the following list, k denotes the number of agents, p the number of propositions, n the number of states, m the number of transitions, and $\overline{m}$ the number of epistemic links in the original CEGS M. Likewise, k', p', n', m' denote the number of agents, propositions, states and transitions in the resulting CGS M'.

- The vocabulary (set of propositions Π) and the set of agents only increase linearly: $p' = p + k + 1 = O(p + k)$ and $k' = 2k = O(k)$.
- The set of states in an ATEL-model grows linearly, too: $n' = (k+1)n = O(kn)$.
- We have $m' = m + k(\overline{m} + 1) = O(m + k\overline{m})$ transitions in M' (m "action" transitions and $\overline{m}$ epistemic transitions from "action" states, plus $\overline{m} + 1$ transitions from each "epistemic" state).
- The length of formulae also increases linearly: $l \leq l' \leq l(8 + 5k) = O(kl)$.

The transformation of models and formulae is straightforward, and in consequence its complexity is no worse than the complexity of the resulting structures.

4.3 Handling Distributed Knowledge

In order to interpret the full ATEL we modify the construction from Section 4.2 by introducing additional epistemic agents (and states) indexed with coalitions which occur with a distributed knowledge operator:

- $\mathbb{A}\mathrm{gt}^e = \{e_i \mid i \in \mathbb{A}\mathrm{gt}\} \cup \{e_A \mid D_A \in \varphi\}$;
- $St' = St \cup \bigcup_{i \in \mathbb{A}\mathrm{gt}} St^{\mathsf{e_i}} \cup \bigcup_{D_A \in \varphi} St^{\mathsf{e_A}}$.

Accordingly, we extend the language with new propositions $\{\mathsf{e_i} \mid i \in \mathbb{A}\mathrm{gt}\}$ and $\{\mathsf{e_A} \mid D_A \in \varphi\}$. The choices of collective epistemic agents e_A refer to the (epistemic copies of) states accessible via distributed knowledge relations:

- $d'_{e_A}(q) = \{pass\} \cup img(q, \sim^D_A)^{\mathsf{e_A}}$.

The new transition function extends the one from Section 4.2 with choices of agents e_A (putting them in any predefined order, e.g. alphabetical order):

$$o'(q, \alpha_1, ..., \alpha_k, \alpha_{e_1}, ..., \alpha_{e_k}, ..., \alpha_{e_A}, ...) = \begin{cases} o(q, \alpha_1, ..., \alpha_k) & \text{if } q \in St \text{ and} \\ & \alpha_a = pass \text{ for all } a \in \mathbb{A}\mathrm{gt}^e \\ q_0 & \text{if } q = q_0^{\mathsf{e_i}} \in St^{\mathsf{e_i}} \text{and} \\ & \alpha_a = pass \text{ for all } a \in \mathbb{A}\mathrm{gt}^e \\ (\alpha_{e_a})^{e_a} & \text{if } e_a \text{ is the first active} \\ & \text{epistemic agent.} \end{cases}$$

The translation of formulae for all operators of ATEL$_{CE}$ remains the same as well, and the translation of D_A is:

$$tr(D_A\varphi) = \neg\langle\langle \mathbb{A}\mathrm{gt}^e \rangle\rangle \bigcirc \big(\mathsf{e_A} \wedge \langle\langle \mathbb{A}\mathrm{gt}^e \rangle\rangle \bigcirc (\mathsf{act} \wedge \neg tr(\varphi))\big).$$

Theorem 2. *Let φ be a formula of* ATEL*, M be a* CEGS*, and $q \in St$ an "action" state in M. Furthermore, let η be a behavioral constraint in M, and let η' extend η in M'. Then, $M, q, \eta \models \varphi$ iff $M', q, \eta' \models tr(\varphi)$.*

This construction, too, does not involve any substantial increase of complexity. Still, it has one disadvantage when compared to the construction from Section 4.2: there, models and formulae could be translated independently; here, the transformation of a model depends on the formula which will be model-checked. Thus, it is not possible any more to "pre-compile" a given CEGS in advance, and then model-check on the fly any formulae that will become relevant.

4.4 Reducing Knowledge to Strategic Ability: Example

Since the transformation of models and formulae involves only linear increase of their size, it can be used for an efficient reduction of model checking when we want to get rid of epistemic operators from formulae. Strategic constraints, on the other hand, enable realistic approach to the semantics of abilities. The idea behind indistinguishability relations is that they capture agents' uncertainty about the current state of the game, so our analysis of abilities should be in most cases restricted to uniform strategies.

Let $\langle\!\langle A \rangle\!\rangle_{u}$ be a "uniform" version of cooperation modality, similar to the operator $\langle\!\langle A \rangle\!\rangle_{ir}$ from [10]. The semantics of $\langle\!\langle A \rangle\!\rangle_{u} \gamma$ is the same as for $\langle\!\langle A \rangle\!\rangle \gamma$ except that only uniform strategies can be used by A. It is easy to see that $\langle\!\langle A \rangle\!\rangle_{u}$ can be rephrased as an ordinary cooperation modality with the strategic constraint that requires A's choices to be uniform: $M, q \models \langle\!\langle A \rangle\!\rangle_{u} \gamma$ iff $M, q, \langle \Sigma^{u}, A \rangle \models \langle\!\langle A \rangle\!\rangle \gamma$.

For example, we have that $M_1, q_1 \models K_a \neg \langle\!\langle a \rangle\!\rangle_{u} \Diamond \mathsf{p_0}$ for the CEGS from Figure 1. This can be rephrased as $M_1, q_1, \langle \Sigma^{u}, \{a\} \rangle \models K_a \neg \langle\!\langle a \rangle\!\rangle \Diamond \mathsf{p_0}$, which is by Theorem 1 equivalent to $M_1', q_1, \langle \Sigma^{u}, \{a\} \rangle \models \neg \langle\!\langle e_a \rangle\!\rangle \bigcirc \big(\mathsf{e_a} \wedge \langle\!\langle e_a \rangle\!\rangle \bigcirc (\mathsf{act} \wedge \langle\!\langle a \rangle\!\rangle \Diamond \mathsf{p_0})\big)$, where M_1' is the concurrent game structure from Figure 2. Thus, we have reduced the original property (and model) to ones that include no epistemic dimension.

Note that we can incorporate the uniformity constraints back into the cooperation modalities if we keep epistemic links in the reconstructed model. Let M_1'' be M_1' with epistemic links retained from the original model M_1 (plus reflective epistemic links added for the epistemic agent e_a to indicate that e_a has perfect information in every state). Then, $M_1, q_1 \models K_a \neg \langle\!\langle a \rangle\!\rangle_{u} \Diamond \mathsf{p_0}$ iff $M_1'', q_1 \models \neg \langle\!\langle e_a \rangle\!\rangle_{u} \bigcirc \big(\mathsf{e_a} \wedge \langle\!\langle e_a \rangle\!\rangle_{u} \bigcirc (\mathsf{act} \wedge \langle\!\langle a \rangle\!\rangle_{u} \Diamond \mathsf{p_0})\big)$. On a more general level, Theorem 1 implies that adding explicit operators K_a for describing agents' knowledge does not increase the complexity of model checking agents' abilities also in the case of imperfect information strategies.

5 Conclusions

In this paper, we propose an update of the reduction scheme that was presented in [4]. The original reduction allowed to get rid of epistemic operators by translating them to cooperation modalities of ATL which made use of additional "epistemic" agents. The new version has two new features. First, we avoid the exponential blowup of formulae, which was to some extent present in the original reduction. Second, we show that the reduction is valid also if we specify *strategic constraints* which restrict collective strategies that some (or all) agents are allowed to use. Thus, the applicability of the new reduction scheme goes well beyond ATEL (i.e., perfect information ATL + knowledge operators). We can use the scheme to translate knowledge to strategic ability for agents playing under imperfect information (like in ATL_{ir} from [10]), acting in the presence of social norms [11], or choosing only rational play [6]. It seems that many other extensions of alternating-time logic should submit to the reduction, too.

References

1. R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002.
2. E.M. Clarke and E.A. Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Proceedings of Logics of Programs Workshop*, volume 131 of *Lecture Notes in Computer Science*, pages 52–71, 1981.
3. J. Gerbrandy. *Bisimulations on Planet Kripke*. PhD thesis, University of Amsterdam, 1999.
4. V. Goranko and W. Jamroga. Comparing semantics of logics for multi-agent systems. *Synthese*, 139(2):241–280, 2004.
5. W. Jamroga. Reducing knowledge operators in the context of model checking. Technical Report IfI-07-09, Clausthal University of Technology, 2007.
6. W. Jamroga and N. Bulling. A framework for reasoning about rational agents. In *Proceedings of AAMAS'07*, pages 592–594, 2007.
7. W. Jamroga and W. van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 63(2–3):185–219, 2004.
8. J.-J.Ch. Meyer. A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136, 1988.
9. K. Schild. On the relationship between BDI logics and standard logics of concurrency. *Autonomous Agents and Multi Agent Systems*, pages 259–283, 2000.
10. P. Y. Schobbens. Alternating-time logic with imperfect recall. *Electronic Notes in Theoretical Computer Science*, 85(2), 2004.
11. W. van der Hoek, M. Roberts, and M. Wooldridge. Social laws in alternating time: Effectiveness, feasibility and synthesis. *Synthese*, 2005.
12. W. van der Hoek and M. Wooldridge. Cooperation, knowledge and time: Alternating-time Temporal Epistemic Logic and its applications. *Studia Logica*, 75(1):125–157, 2003.
13. S. van Otterloo, W. van der Hoek, and M. Wooldridge. Knowledge as strategic ability. *Electronic Lecture Notes in Theoretical Computer Science*, 85(2), 2003.