
Embedding Alternating-time Temporal Logic in Strategic STIT Logic of Agency

JAN BROERSEN, *Department of Information and Computing Sciences, Universiteit Utrecht, Utrecht, The Netherlands.*
E-mail: broersen@cs.uu.nl

ANDREAS HERZIG, *Institut de Recherche en Informatique de Toulouse, CNRS, Université Paul Sabatier, Toulouse, France.*
E-mail: herzig@irit.fr

NICOLAS TROQUARD, *Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, Toulouse, France & Laboratorio di Ontologia Applicata, Università degli Studi di Trento, Trento, Italy.*
E-mail: troquard@irit.fr

Abstract

Seeing To It That (STIT) logic is a logic of agency, proposed in the 1990s in the domain of philosophy of action. It is the logic of constructions of the form ‘agent a sees to it that φ ’. We believe that STIT theory can contribute to the logical analysis of multiagent systems. To support this claim, we show that there is a close relationship with more recent logics for multiagent systems. This work extends Broersen *et al.* (2006, *Electron Notes Theor. Comput. Sci.*, Vol. 157, pp. 23–35) where we presented a translation from Pauly’s Coalition Logic to Chellas’ STIT logic. Here we focus on Alur, Henzinger and Kupferman’s Alternating-time Temporal Logic (ATL), and the logic of the ‘fused’ $\diamond_s[_scstit : _]$ operator for strategic ability, as described by Horty. After a brief presentation of ATL and the definition of a discrete-time strategic STIT framework slightly adapted from Horty, we give a translation from ATL to the STIT framework, and prove that it determines correct embedding.

Keywords: Modal logic, multi-agent systems, agency, strategic choice, philosophical foundations.

1 Introduction

Seeing To It That (STIT) logics have been proposed in the 1990s in the domain of the philosophy of action [9]. They are logics of constructions of the form ‘agent a sees to it that φ ’. Several versions of this modality have been studied in the philosophical literature. Here we use a *strategic* one, namely, the *fused* STIT operator $\diamond_s[_scstit : _]$ [17, p. 152] which combines a modality \diamond_s for strategic possibility with a strategic version of Chellas’ version of the STIT operator $[_cstit : _]$. This modality aims to model the *ability* of groups of agents to ensure something by means of a strategy.

The semantics of the STIT operator is based on branching-time temporal structures. In this sense it extends the logical ontology of the so-called ‘bringing it about’ operators [12, 22] that abstract from the temporal aspect of agency. This results in weak modal logics which can be given semantics in various ways, for instance by means of neighbourhood models. However,

we see the interaction with time as a crucial ingredient of agency that deserves a central place in the ontology of logics for agency and logics for multiagent systems.

In the philosophical literature the STIT operator has been used in the analysis of agency and in the analysis of deontic concepts [10, 17]. We believe that the philosophical intuitions underlying STIT theory are equally relevant for logical models developed to analyse and design multiagent systems. To support this claim, in this article we show that there is a close relationship with more recent temporal logics for specification and verification of multiagent systems. In particular, we will study here the relation between Alternating-time Temporal Logic (ATL) proposed by Alur *et al.* [1–3] and the logic of the fused $\diamond_s[_scstit : _]$ operator, as described by Horty [17]. ATL was designed as an extension of CTL. CTL is a branching-time temporal logic with modal operators quantifying (universal (**A**) and existential (**E**)) over sets of paths. In ATL, quantification is with respect to strategies, and quantification over paths is implicit as quantification over all paths that are in the *outcome* of a certain strategy. In particular, $\langle\langle A \rangle\rangle$, where A is a group of agents ($A \subseteq \mathcal{A}gt$, where $\mathcal{A}gt$ is the set of all agents), stands for existential quantification over strategies. In ATL, $\langle\langle A \rangle\rangle$ is always followed by one of the temporal operators **X** (next), **G** (henceforth) or **U** (until). Evaluation of these temporal operators is with respect to paths that are in the outcome of a strategy. For example, $\langle\langle A \rangle\rangle \mathbf{X}\varphi$ reads: ‘group A has a strategy to ensure that next φ ’. This setting allows for refinements of the CTL quantification over paths, CTL **E** corresponding to the ATL $\langle\langle \mathcal{A}gt \rangle\rangle$ and **A** corresponding to $\langle\langle \emptyset \rangle\rangle$. It was shown by Goranko [14] that ATL is also an extension of Pauly’s Coalition Logic CL [21]. The latter is the logic of expressions of the form $[A]\varphi$, reading ‘group A can ensure that φ ’. Such expressions correspond to ATL formulas $\langle\langle A \rangle\rangle \mathbf{X}\varphi$.

In [6], we proposed the following translation from CL to STIT. The box is the operator for historic necessity.

$$\begin{aligned} tr_{CL}(p) &= \Box p, \text{ for } p \in \mathcal{A}tm \\ tr_{CL}(\neg\varphi) &= \neg tr(\varphi) \\ tr_{CL}(\varphi \vee \psi) &= tr(\varphi) \vee tr(\psi) \\ tr_{CL}([A]\varphi) &= \Box[A \text{ cstit} : \mathbf{X}tr(\varphi)] \end{aligned}$$

In this article we propose a translation from ATL [14] to a discrete version of strategic STIT logic.

In [27], a close examination of the differences and similarities of the models of STIT theory and ATL is undertaken. It is shown that, under the addition of some specific conditions (e.g. discreteness), the models of the two systems can be seen to obey similar properties, like tree-likeness, uniformity and ‘restrictedness’ (Section 4). However, these properties are not necessarily expressible in the logics of STIT or ATL. So, although, from a philosophical point of view, it is interesting to look at properties of models as such, here we are essentially interested only in those properties that are expressible in the logics. Where [27] only compares the models for ATL and STIT, we also compare the logics of both systems.

In Section 2, we offer a brief presentation of ATL. Section 3 deals with an adapted discrete-time STIT framework. We prove a semantic equivalence result on ATL frames in Section 4. Section 5 presents the main result of this note: we describe a translation from ATL to STIT, and prove that it is correct.¹ We conclude with a discussion and some perspectives of investigation in Section 6.

¹ A correct embedding is a sound and complete translation to a fragment of a stronger logic.

2 Alternating-time Temporal Logic

The first paper on ATL was [1]. This preliminary work is restricted to turn-based games, i.e. games where each transition is governed by a single agent. Alur *et al.* [2] come with general structures called *alternating transition systems* (ATs), where choices are expressed as sets of possible outcomes. In [3], the authors change the models into *concurrent game structures* (CGSs),² where choices are identified with explicit labels. ATs and CGSs have been proven equivalent by Goranko and Jamroga [13]. Hence, defining the semantics of ATL in terms of either ATs or CGSs is a matter of convenience.

In what follows, $\mathcal{A}tm$ represents a set of atomic propositions, and $\mathcal{A}gt$ is the finite set of all agents.

SYNTAX. Given that p ranges over $\mathcal{A}tm$, and that A ranges over $2^{\mathcal{A}gt}$, the language of ATL is defined by:

$$\varphi, \psi, \dots ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid \langle\langle A \rangle\rangle X\varphi \mid \langle\langle A \rangle\rangle G\varphi \mid \langle\langle A \rangle\rangle \varphi U \psi$$

The intended reading of $\langle\langle A \rangle\rangle \eta$, with η a linear temporal formula (branch formula), is that ‘group A can ensure η whatever agents in $\mathcal{A}gt \setminus A$ do’.

MODELS. We present models for ATL as in [2], that is, in terms of alternating transition systems which are tuples $\mathcal{M} = \langle W, \delta, v \rangle$, where:

- W is a non-empty set of states (alias worlds, alias moments).
- $\delta : S \times \mathcal{A}gt \rightarrow 2^{2^W}$ is a transition function mapping each moment and agent to a non-empty family of sets of possible successor moments.
- $v : \mathcal{A}tm \rightarrow 2^W$ is a valuation function.

Each $Q \in \delta(w, a)$ may be seen as the choice by an agent of a particular action in its repertoire.

We use *lock-step synchronous* ATs, which means that in every state, all agents proceed simultaneously (as opposed to the particular case of *turn-based synchronous* ATs). The δ function is *non-blocking* (agent’s actions are always compatible) and the simultaneous choice of every agent in $\mathcal{A}gt$ determines a *unique next state*: assuming $\mathcal{A}gt = \{a_1, \dots, a_n\}$, for every state $w \in W$ and every set $\{Q_1, \dots, Q_n\}$ of choices $Q_i \in \delta(w, a_i)$, the intersection $Q_1 \cap \dots \cap Q_n$ is a singleton.

A *strategy for an agent a* is a mapping $f_a : W^+ \rightarrow 2^W$, such that it associates to each sequence of states $w_0 \dots w_k$ an element of $\delta(w_k, a)$.³ A collective strategy, for a set of agents $A \subseteq \mathcal{A}gt$ is a tuple $F_A = \langle f_{a_1}, f_{a_1}, \dots, f_{a_n} \rangle$ of strategies, one for each agent in $\mathcal{A}gt$. The outcome of F_A from w is defined as:

$$out(w, F_A) = \{\lambda \mid \lambda = w_0 w_1 w_2 \dots, w_0 = w, \forall i \geq 0 (w_{i+1} \in \bigcap_{a \in A} f_a(w_0 \dots w_i))\}$$

DEFINITION 1 (Strategy profile/choice profile)

A strategy profile is a collective strategy $F_{\mathcal{A}gt}$ for all agents of $\mathcal{A}gt$. Analogously, a tuple $\langle Q_1, \dots, Q_n \rangle$ (one Q_i for each $i \in \mathcal{A}gt$) is called a *choice profile*.

² An alternative name from the literature is ‘multi-player game model’, abbreviated ‘MGM’.

³ It actually suffices to use mappings $f_a : W \rightarrow 2^W$ [13]. Also in the STIT setting of Section 3, strategies will be defined as functions from states into choices (and not from sequences of states into choices). However, the current definition is the customary one.

SEMANTICS AND AXIOMATIZATION. $\lambda[i]$ is the i -th position in the path λ . A formula is evaluated with respect to an ATS $\mathcal{M} = \langle W, \delta, \nu \rangle$ and a moment $w \in W$.

$$\begin{aligned} \mathcal{M}, w \models \langle\langle A \rangle\rangle \mathbf{X}\varphi &\iff \exists F_A, \forall \lambda \in \text{out}(w, F_A), \mathcal{M}, \lambda[1] \models \varphi \\ \mathcal{M}, w \models \langle\langle A \rangle\rangle \mathbf{G}\varphi &\iff \exists F_A, \forall \lambda \in \text{out}(w, F_A), \mathcal{M}, \lambda[i] \models \varphi, \forall i \geq 0 \\ \mathcal{M}, w \models \langle\langle A \rangle\rangle \varphi \mathcal{U} \psi &\iff \exists F_A, \forall \lambda \in \text{out}(w, F_A), \\ &\quad \exists i \geq 0 (\mathcal{M}, \lambda[i] \models \psi, \forall j \in [0, i], \mathcal{M}, \lambda[j] \models \varphi) \end{aligned}$$

Validity is defined as usual. The following complete axiomatization of ATL (as an extension of any axiomatization for propositional logic) is given in [15]. $\mathcal{M}, w \models \langle\langle \emptyset \rangle\rangle \eta$ means that η holds irrespective of the choices made by A .

$$\begin{aligned} (\perp) \quad &\neg \langle\langle A \rangle\rangle \mathbf{X}\perp \\ (\top) \quad &\langle\langle A \rangle\rangle \mathbf{X}\top \\ (N) \quad &\neg \langle\langle \emptyset \rangle\rangle \mathbf{X}\neg\varphi \rightarrow \langle\langle \text{Agt} \rangle\rangle \mathbf{X}\varphi \\ (S) \quad &\langle\langle A_1 \rangle\rangle \mathbf{X}\varphi \wedge \langle\langle A_2 \rangle\rangle \mathbf{X}\psi \rightarrow \langle\langle A_1 \cup A_2 \rangle\rangle \mathbf{X}(\varphi \wedge \psi) \text{ if } A_1 \cap A_2 = \emptyset \\ (FP_G) \quad &\langle\langle A \rangle\rangle \mathbf{G}\varphi \equiv \varphi \wedge \langle\langle A \rangle\rangle \mathbf{X}\langle\langle A \rangle\rangle \mathbf{G}\varphi \\ (GFP_G) \quad &\langle\langle \emptyset \rangle\rangle \mathbf{G}(\theta \rightarrow (\varphi \wedge \langle\langle A \rangle\rangle \mathbf{X}\theta)) \rightarrow \langle\langle \emptyset \rangle\rangle \mathbf{G}(\theta \rightarrow \langle\langle A \rangle\rangle \mathbf{G}\varphi) \\ (FP_U) \quad &\langle\langle A \rangle\rangle \psi \mathcal{U} \varphi \equiv \varphi \vee (\psi \wedge \langle\langle A \rangle\rangle \mathbf{X}\langle\langle A \rangle\rangle \psi \mathcal{U} \varphi) \\ (LFP_U) \quad &\langle\langle \emptyset \rangle\rangle \mathbf{G}((\varphi \vee (\psi \wedge \langle\langle A \rangle\rangle \mathbf{X}\theta)) \rightarrow \theta) \rightarrow \langle\langle \emptyset \rangle\rangle \mathbf{G}(\langle\langle A \rangle\rangle \psi \mathcal{U} \varphi \rightarrow \theta) \\ (\langle\langle A \rangle\rangle \mathbf{X}\text{-Mon}) \quad &\text{from } \varphi \rightarrow \psi \text{ infer } \langle\langle A \rangle\rangle \mathbf{X}\varphi \rightarrow \langle\langle A \rangle\rangle \mathbf{X}\psi \\ (\langle\langle \emptyset \rangle\rangle \mathbf{G}\text{-Nec}) \quad &\text{from } \varphi \text{ infer } \langle\langle \emptyset \rangle\rangle \mathbf{G}\varphi \end{aligned}$$

Note that the (N) axiom follows from the determinism of choice profiles (actions constituted by simultaneous choices for every agent in the system): when every agent opts for a choice, the next state is fully determined, thus, if something is not settled, the coalition of all agents (Agt) can always work together to make its negation true. The axiom (S) says that two coalitions can combine their efforts to ensure a conjunction of properties if they are disjoint. Note that from (S) it follows that $\langle\langle A_1 \rangle\rangle \varphi \wedge \langle\langle A_2 \rangle\rangle \neg\varphi$ is not satisfiable for disjoint A_1 and A_2 . So, two disjoint coalitions cannot ensure inconsistent propositions. Axiom (FP_G) characterizes the global modality as a fixpoint of the next modality, and axiom (GFP_G) says that this is the greatest fixpoint. Axiom (FP_U) characterizes the until operator as a (special kind of) fixpoint of the next operator, and axiom LFP_U expresses that the semantics dictates that we take the least fixpoint.

3 A logic for strategic STIT ability

STIT theory originates from philosophy. Probably the first study to refer to the logic of ‘seeing to it that’ is [8]. It analyses the needs for a general theory of ‘an agent making a choice among alternatives that lead to an action’. The thesis is that the best way to meet this goal is to augment the language with a class of sentences. The proposed class is one of the sentences of the form ‘Ishmael sees to it that Ishmael sailed on board the Pequod’ paraphrasing the sentence ‘Ishmael sailed on board the Pequod’. Thus, from any sentence involving an action (e.g. *sailing*) we can reformulate it into an agentive one stating that an *agent a* sees to it that a

state of affairs φ holds, formally: $[astit : \varphi]$. Ref. [8] is a road map towards an outstandingly rich and justified theory of agency compiled in [10] and [17].

It is worth noting that, often puzzlingly, STIT is widely influenced by the observation that in a branching-time framework, future-tensed statements are ambiguous to evaluate if not impossible. Suppose a moment w_0 and two different moments w_1 and w_2 lying in the future of w_0 on two different courses of time. φ is true at w_1 and false at w_2 and everywhere before and after. What truth value should be assigned to the sentence ‘ φ is true in the future of w_0 ’? Indeed, φ really does lie in the future of w_1 , but what if the course of time happens to go through w_2 instead? In general, in branching time, a moment alone does not provide enough information to determine the truth value of a sentence about the future. Prior [23] and Thomason [24, 25] hence proposed to evaluate future-tensed sentences with respect to a moment *and* a particular course of time running through it. This is why, as we will see, states of the world in STIT models consist of ‘fragmentized’ moments; moments fragmentize into as much indexes as there are courses of time through it. In this section, we present the elements of the theory that are relevant in this work. Some extra assumptions are made with respect to the original STIT theory. They are discussed and motivated in Section 6.

Next we define the syntax of strategic STIT ability, as defined in chapter 6 of Horty’s book [17], augmented by LTL reasoning capabilities.

SYNTAX Given that p ranges over Atm , and that A ranges over 2^{Agt} , a language of strategic STIT is defined by:

$$\varphi, \psi, \dots ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi \mid \mathbf{X}\varphi \mid \mathbf{G}\varphi \mid \varphi\mathcal{U}\psi \mid \diamond_s[A\ scstit : \varphi]$$

First we have to explain why we call the logic defined relative to the above syntax a logic of ‘strategic STIT *ability*’ instead of a logic of ‘strategic STIT’. The intuitive reading of $\diamond_s[A\ scstit : \varphi]$ is ‘it is strategically possible that agents A see to it that φ ’. The operator $\diamond_s[A\ scstit : \varphi]$, suggested by Horty [17, p. 152], is thus a special (fused) operator that is ‘built’ from an operator for strategic possibility ($\diamond_s\varphi$) and a strategic version of Chellas’ STIT operator $[A\ cstit : \varphi]$. However, in Horty’s work these separate operators are not given a formal semantics individually; the operators are syntactically forced to occur only in combination (in a recent proposal [7] we propose a solution to this problem by evaluating with respect to strategy/state pairs). Yet, to understand the semantics of the fused operator, next we discuss the intended semantics of the individual operators.

The semantics of STIT is embedded in the branching-time framework. It is based on structures of the form $\langle W, < \rangle$, in which W is a non-empty set of moments, and $<$ is a tree-like ordering of these moments: for any w_1, w_2 and w_3 in W , if $w_1 < w_3$ and $w_2 < w_3$, then either $w_1 = w_2$ or $w_1 < w_2$ or $w_2 < w_1$.

A maximal set of linearly ordered moments from W is a *history*. Thus, $w \in h$ denotes that moment w is *on the history* h . We define *Hist* as the set of all histories of a STIT structure. $H_w = \{h \mid h \in \text{Hist}, w \in h\}$ denotes the set of histories passing through w . An *index* is a pair w/h , consisting of a moment w and a history h from H_w (i.e. a history and a moment in that history).

To enable a comparison with ATL we make the following assumption:

ASSUMPTION 1 (Countably infiniteness)

Every history is isomorphic to the set of natural numbers.

By assuming that histories are countably infinite sets of moments we will be able to reason about temporal properties as in LTL.

A STIT model is a tuple $\mathcal{M} = \langle W, Choice, <, v \rangle$, where:

- $\langle W, < \rangle$ is a branching-time structure.
- $Choice : Agt \times W \rightarrow 2^{2^{Hist}}$ is a function mapping each agent and each moment w into a partition of H_w . The equivalence classes belonging to $Choice_a^w$ can be thought of as possible choices or actions available to agent a at w . Given a history $h \in H_w$, $Choice_a^w(h)$ represents the particular choice from $Choice_a^w$ containing h , or in other words, the particular action performed by a at the index w/h . We must have $Choice_a^w \neq \emptyset$ and $Q \neq \emptyset$ for every $Q \in Choice_a^w$.
- v is valuation function $v : Atm \rightarrow 2^{W \times Hist}$.

REMARK

In STIT models, moments may have different valuations, depending on the history they are living in (cf. [16, footnote 2, p. 586]). Thus, at any specific moment, we might have different valuations corresponding to the results of the different (non-deterministic) actions possibly taken at that moment.

DEFINITION 2 (Current moment/current choice)

At index w/h we shall call w the current moment and $Choice_a^w(h)$ the current choice/action.

In order to deal with group agency, Horty defines in [17, section 2.4], the notion of collective choice. Horty first introduces action selection functions s_w from Agt into 2^{H_w} satisfying the condition that for each $w \in W$ and $a \in Agt$, $s_w(a) \in Choice_a^w$. So, a selection function s_w selects a particular action for each agent at w .

Then, for a given w , $Select_w$ is the set of all selection functions s_w . For every $s_w \in Select_w$, it is assumed that $\bigcap_{a \in Agt} s_w(a) \neq \emptyset$. This constraint corresponds to the assumption that the agents' choices are independent, in the sense that agents can never be deprived of choices due to the choices made by other agents.

Moreover, in order to match ATL, we make the following assumption stating that the intersection of choices of agents in Agt must exactly be the set of histories passing through some immediate next moment:

ASSUMPTION 2 (Determinism)

$$\forall w \in W, \exists w' \in W (w < w' \text{ and } \bigcap_{a \in Agt} s_w(a) = H_{w'})$$

Note that because STIT frames are trees, the moment w' is always a next moment.

Using choice selection functions s_w , the *Choice* function can be generalized to apply to groups of agents ($Choice : 2^{Agt} \times W \rightarrow 2^{2^{Hist}}$). A collective choice for a group of agents $A \subseteq Agt$ is defined as:

$$Choice_A^w = \left\{ \bigcap_{a \in A} s_w(a) \mid s_w \in Select_w \right\}$$

Again, $Choice_A^w(h) = \{h' \mid \text{there is } Q \in Choice_A^w \text{ such that } h, h' \in Q\}$.

SEMANTICS. We conclude $\models_{\text{STIT}} \varphi$ if $\mathcal{M}, w/h \models \varphi$ for every STIT model \mathcal{M} , h in \mathcal{M} and moment w in h . A formula is evaluated with respect to a model and an index.

$$\begin{aligned} \mathcal{M}, w/h \models p & \iff w/h \in v(p), p \in \text{Atm} \\ \mathcal{M}, w/h \models \neg\varphi & \iff \mathcal{M}, w/h \not\models \varphi \\ \mathcal{M}, w/h \models \varphi \vee \psi & \iff \mathcal{M}, w/h \models \varphi \text{ or } \mathcal{M}, w/h \models \psi \end{aligned}$$

Historical necessity (or inevitability) at a moment w in a history is defined as truth in all histories passing through w :

$$\mathcal{M}, w/h \models \Box\varphi \iff \mathcal{M}, w/h' \models \varphi, \forall h' \in H_w.$$

When $\Box\varphi$ holds at w then φ is said to be *settled true* at w . $\Diamond\varphi$ is defined in the usual way as $\neg\Box\neg\varphi$, and stands for historical possibility.

There are several STIT operators; the so-called Chellas' STIT is defined as follows:

$$\mathcal{M}, w/h \models [A \text{ cstit} : \varphi] \iff \mathcal{M}, w/h' \models \varphi, \forall h' \in \text{Choice}_A^w(h).$$

Intuitively it means that group A 's current choices ensure φ , whatever other agents outside A do. The more complex *deliberative* STIT is defined as $[A \text{ dstit} : \varphi] \equiv [A \text{ cstit} : \varphi] \wedge \neg\Box\varphi$.

As shown in [9], both Chellas' STIT and historical necessity are S5 modal operators, and $\models_{\text{STIT}} \Box\varphi \rightarrow [A \text{ cstit} : \varphi]$.

As time is discrete in our present setting, we can define the temporal operator **X** (*next*). We also introduce operators **G** (*always*) and **U** (*until*):

$$\begin{aligned} \mathcal{M}, w/h \models \mathbf{X}\varphi & \iff \exists w' \in h (w < w', \mathcal{M}, w'/h \models \varphi, \\ & \quad \nexists w'' \in h (w < w'' < w')). \\ \mathcal{M}, w/h \models \mathbf{G}\varphi & \iff \forall w' \in h (w \leq w', \mathcal{M}, w'/h \models \varphi) \\ \mathcal{M}, w/h \models \varphi \mathbf{U} \psi & \iff \exists w' \in h (w < w', \mathcal{M}, w'/h \models \psi, \\ & \quad \forall w'' (w \leq w'' < w', \mathcal{M}, w''/h \models \varphi) \end{aligned}$$

STRATEGIES. Belnap *et al.* [10] and Horty [17] introduce strategies into STIT theory: a *strategy* for an agent a is a partial function σ on W such that $\sigma(w) \in \text{Choice}_a^w$ for each moment w from $\text{Dom}(\sigma)$, the domain of σ . In STIT theory, it is assumed that σ may be a partial function. The reason is that there is no need to account for choices at states an agent never arrives at by following σ . In [10, p. 350] it says 'A strategy need not tell us what to do at moments that the strategy itself forbids'. This contrasts with ATL, where it is implicitly assumed that strategies are total. But, as the present comparison between both systems reveals, for the basic ATL modalities this is not at all necessary.⁴

As we can see in the definition of the $[_\text{cstit} : _]$ operator, an agent's choice restricts the set of possible futures, in particular it restricts the histories to those corresponding with the

⁴ However, if we extend ATL with strategic STIT operators, as we did in [7], totality of strategy functions with respect to the domain of states is indeed necessary.

choice being made. We expect a strategy to be a generalization of this, in particular, we want a strategy to restrict the possible histories to those corresponding to a series of choices being made at successive moments.

DEFINITION 3 (Admitted histories)

A strategy σ *admits* a history h if and only if (i) $Dom(\sigma) \cap h \neq \emptyset$ and (ii) for each $w \in Dom(\sigma) \cap h$ we have $h \in \sigma(w)$. The set of all histories admitted by a strategy σ is denoted $Adh(\sigma)$.

We will often use the notation σ_a , to name a particular strategy of an agent a .

DEFINITION 4 (Collective strategy)

A collective strategy for $A \subseteq \mathcal{Agt}$ is a tuple $\sigma_A = \langle \sigma_a \rangle_{a \in A}$, and $Adh(\sigma_A) = \bigcap_{a \in A} Adh(\sigma_a)$.

Horty [17] also proposes strategies with a limited scope. To this end, he introduces the notion of *field* which is a $<$ -backward-closed subset M of $Tree_w = \{w' \mid w < w' \text{ or } w = w'\}$. With $Adm(\sigma) = \{w \mid w \in h, h \in Adh(\sigma)\}$, a strategy is properly formed in the field M if it is *complete* in M ($Adm(\sigma) \cap M \subseteq Dom(\sigma)$) and *irredundant* ($Dom(\sigma) \subseteq Adm(\sigma)$). Thus, an *ability* operator should be evaluated with respect to a field.

In this work, we do not need such a refinement. Therefore, for any strategy at a moment w we will always consider the field to be the complete set $Tree_w$, that is, the backward-closed sub-tree having w as root. For evaluation of formulas in the strategic setting we will use the same models and indexes as for the non-strategic setting.⁵

As discussed in [17], global effectivity by means of a strategy differs from local effectivity induced by a unique (possibly collective) choice. Available choices at a moment form a partition of that moment: one history lies in one and only one choice. But, the sets of admitted histories of the strategies available at a given moment do *not* necessarily partition that moment. One history can lie in the sets of admitted histories of two different strategies. Therefore, since a history alone does not tell us which strategy we have to consider, we cannot evaluate global effectivity as we have done for local effectivity (the $[_{cstit} : _]$ operator). However, those semantic difficulties are outside the scope of this paper. We refer the reader to [17, Section 7.2.1] and to [7], where we propose a solution to this problem in the ATL-setting.

Horty points out that we can return to a natural evaluation by using an operator quantifying over strategies. In particular, we can define a *fused* operator for long-term strategic ability of groups of agents as follows:

$$\begin{aligned} \mathcal{M}, w/h \models \diamond_s[A \text{ scstit} : \varphi] &\iff \\ \exists \sigma \in \text{Strategy}_A^w \text{ s.t. } \forall h' \in Adh(\sigma), \mathcal{M}, w/h' &\models \varphi \end{aligned}$$

where $\text{Strategy}_A^w = \{\sigma \mid Dom(\sigma) = Tree_w\}$, is the set of strategies open to A at moment w .⁶

Intended readings for $\diamond_s[A \text{ scstit} : \varphi]$ are: ‘it is strategically possible that agents A see to it that φ ’, or ‘ A has the ability to guarantee the truth of φ by carrying out an available strategy’.

⁵ It is easy to see that actually histories are not needed to evaluate the strategic ability operator. Horty calls this moment-determinateness of the fused operator. We nevertheless keep the histories for uniformity purposes.

⁶ In the original definition, a set of strategies is denoted Strategy_A^M , where M is a field having w as root. Since we have assumed that M is always $Tree_w$, our notation Strategy_A^w suffices.

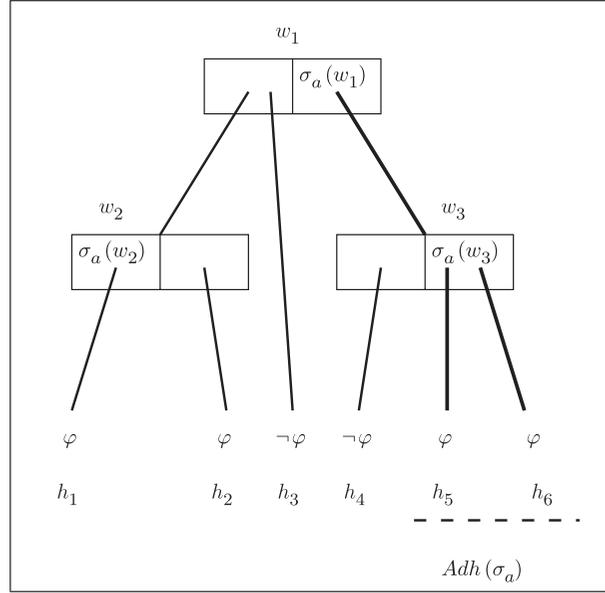


FIGURE 1. Example of strategic STIT with one agent. It is strategically possible that agent a sees to it that some time in the future φ

Horty uses a slightly different syntax and writes this fused operator as $\diamond[A\ scstit : \varphi]$. We use the s -subscript for the diamond to emphasize that it does not reflect *historical* possibility (written without the s -subscript as $\diamond\varphi$) but *strategic* possibility. For enlightenment, we mention the connections of this operator with Chellas' STIT operator and the historical necessity operator.

The strategic ability operator $\diamond_s[A\ scstit : \varphi]$ can be seen to be stronger than the local ability operator $\diamond[_{cstit} : _]$. In particular, it holds that:

$$\models_{\text{STIT}} \diamond[A\ cstit : \varphi] \rightarrow \diamond_s[A\ scstit : \varphi].$$

This property ensures that the translation we propose in Section 5 embeds the translation we did for CL (cf. the definition of tr_{CL} in Section 1).

However, $\diamond[_{cstit} : _]$ and $\diamond_s[_{scstit} : _]$ are not equivalent: in the example of Figure 1, we can imagine a strategy σ_a such that $\sigma_a(w_1) = \{h_4, h_5, h_6\}$, $\sigma_a(w_2) = \{h_1\}$ and $\sigma_a(w_3) = \{h_5, h_6\}$. h_1 , h_2 and h_3 are not admitted because they do not lie in $\sigma_a(w_1)$. $Dom(\sigma_a) \cap h_4 = \{w_1, w_3\}$, but $h_4 \notin \sigma_a(w_3)$, so $h_4 \notin Adh(\sigma_a)$. However, h_5 and h_6 are in $Adh(\sigma_a)$, and there are no other histories in $Adh(\sigma_a)$. So, there exists a strategy σ_a such that for every history in $Adh(\sigma_a)$, φ is true some time in the future. So, for all $h \in H_{w_1}$, $\mathcal{M}, w_1/h \models \diamond_s[a\ scstit : \top\mathcal{U}\varphi]$. However, for any $h \in H_{w_1}$ we also have $\mathcal{M}, w_1/h \not\models \diamond[a\ cstit : \top\mathcal{U}\varphi]$.

Note that the strategy σ'_a with $\sigma'_a(w_1) = \{h_1, h_2, h_3\}$, $\sigma'_a(w_2) = \{h_1\}$ and $\sigma'_a(w_3) = \{h_4\}$ cannot ensure that φ some time in the future, because $Adh(\sigma'_a) = \{h_1, h_3\}$, and $\mathcal{M}, w_1/h_3 \not\models \top\mathcal{U}\varphi$.

In combination with the standard STIT property $\Box\varphi \rightarrow \diamond[A\ cstit : \varphi]$, for nonempty coalitions $A \subseteq \mathcal{A}gt$ we arrive at the following property for strategic ability:

$$\models_{\text{STIT}} \Box\varphi \rightarrow \diamond_s[A\ scstit : \varphi]$$

For empty coalitions this implication strengthens to an equivalence.

PROPOSITION 1

$$\models_{\text{STIT}} \diamond_s[\emptyset.scstit : \varphi] \equiv \Box\varphi$$

PROOF. Since the empty coalition of agents is not assigned any choices, at each moment w' , the empty coalition has no alternative but $H_{w'}$. Hence, $Strategy_{\emptyset}^w = \{\sigma_{\emptyset}\}$ with $\sigma_{\emptyset}(w') = H_{w'}$ for all $w' \in Tree_w$. Therefore, for all σ in $Strategy_{\emptyset}^w$, we have $Adh(\sigma) = H_w$.

Thus $\mathcal{M}, w/h \models \diamond_s[\emptyset.scstit : \varphi] \iff \forall h' \in H_w, \mathcal{M}, w/h' \models \varphi$. Which corresponds to the semantics of the operator of historical necessity ■.

This proposition is instrumental in our proof of Theorem 2.

4 Semantic equivalences for ATL

As a first step towards the embedding, we discuss semantic equivalence results for interpreting ATL on ATSS. First we introduce some convenient notations:

DEFINITION 5 (Successor states/tree-order)

Given an ATS $\mathcal{M} = \langle W, \delta, v \rangle$ and an agent $a \in \mathcal{Agt}$:

- $Succ_a(w) \triangleq \{w' \mid w' \in Q_a, Q_a \in \delta(w, a)\}$,
- $Succ(w) \triangleq \bigcap_{a \in \mathcal{Agt}} Succ_a(w)$,
- $w <_{\delta} w' \triangleq w' \in Succ(w)$,
- $<_{\delta}$ is the transitive closure of $<_{\delta}$.

Intuitively, $Succ_a(w)$ gives the possible successor states from the point of view of agent a , and $Succ(w)$ gives possible successor states for the complete system of agents.

The first steps on the issue of semantical equivalence have already been made by Wölfl [27], who, among other things, shows how any ATS can be unraveled into $\langle W, \delta, v \rangle$ in such a way that $\langle W, <_{\delta} \rangle$ is a tree. From any ATS we can thus construct a tree-like ATS that is bisimilar. Therefore we may restrict our study to tree-like ATSS.

DEFINITION 6 (Tree-like ATSS)

An ATS $\mathcal{M} = \langle W, \delta, v \rangle$ where $\langle W, <_{\delta} \rangle$ is a tree, is called a tree-like ATS.

Now, for ATSS it is not necessarily the case that $Succ_a(w) = Succ(w)$. The only condition on ATSS is that each intersection of choices by all members of \mathcal{Agt} results in a unique state. This does not guarantee that choices for individual agents do not overlap, and it does not guarantee that there are worlds that seem reachable from the point of view of some agents but are actually not reachable in any simultaneous step by all agents in the system. To be more precise, if $\delta(w, a) = \{Q_1 \dots Q_n\}$, then both $Q_i \cap Q_j \neq \emptyset$ and $\bigcup_{1 < i < n} Q_i \not\subseteq Succ(w)$ for some i and j in $[1, n]$ are allowed. These properties would not hold if, like in STIT, choices for individual agents would partition the set of possible reachable worlds. In this section, we will work towards tree-like choice partitioned ATSS and show that they are bisimilar for ATL. For these models we thus have $Succ_a(w) = Succ(w)$, for all $a \in \mathcal{Agt}$.

Wölfl explicitly constrains ATSS with the condition that for each agent a and each state w , $\delta(w, a)$ is a partition of the set of successor states of w . Here we show that this explicit restriction is not necessary.

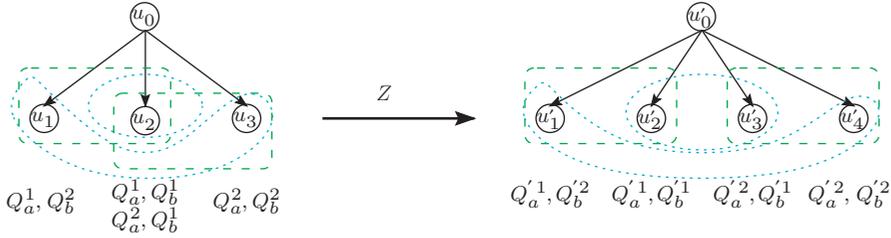


FIGURE 2. Construction of a semantically equivalent choice partitioned ATS. Dotted boxes correspond to $\delta(u_0, a)$ (respectively $\delta(u'_0, a)$) and closed curves correspond to $\delta(u_0, b)$ (respectively $\delta(u'_0, b)$)

DEFINITION 7 (Choice partitioned ATSS)

An ATS $\mathcal{M} = \langle W, \delta, v \rangle$ is called a choice partitioned ATS if for all agent $a \in \text{Agt}$ and for all state $w \in W$ the choices $\delta(w, a)$ partition the set $\text{Succ}(w)$.

LEMMA 1

For any ATS $\mathcal{M} = \langle W, \delta, v \rangle$ we can construct a bisimilar tree-like and choice partitioned ATS $\mathcal{M}' = \langle W', \delta', v' \rangle$.

PROOF. We roughly follow the proof of [4, Prop. 2.15]. Elements of W' are sequences

$$(u_0, \langle u_1, \langle Q_1^0 \dots Q_n^0 \rangle \rangle, \dots, \langle u_k, \langle Q_1^{k-1} \dots Q_n^{k-1} \rangle \rangle)$$

satisfying $k \geq 0$, $\text{Agt} = \{a_1, \dots, a_n\}$, $u_i \in W$, $u_{i+1} \in \bigcap_{a \in \text{Agt}} Q_a^i$, $Q_a^i \in \delta(a, u_i)$. u_0 is intended as the root of \mathcal{M} , and every u_i is a state reached from u_{i-1} by agents of Agt , applying the choice profile $\langle Q_1^i \dots Q_n^i \rangle$. Then, for every agent a and for all $w' = (u_0, \langle u_1, \langle Q_1^0 \dots Q_n^0 \rangle \rangle, \dots, \langle u_k, \langle Q_1^{k-1} \dots Q_n^{k-1} \rangle \rangle)$ of W' , we define $\delta'(a, w') = \{Q_1' \dots Q_d'\}$ with:

$$Q_j' = \{(u_0, \langle u_1, \langle Q_1^0 \dots Q_n^0 \rangle \rangle, \dots, \langle u_k, \langle Q_1^{k-1} \dots Q_n^{k-1} \rangle \rangle), \\ \langle u_{k+1}, \langle Q_1^k \dots Q_a \dots Q_n^k \rangle \rangle \mid \delta(a, u_k) = \{Q_1 \dots Q_a \dots Q_d\}, u_{k+1} \in Q_a\}$$

For all $w' = (u_0, \langle u_1, \langle Q_1^0 \dots Q_n^0 \rangle \rangle, \dots, \langle u_k, \langle Q_1^{k-1} \dots Q_n^{k-1} \rangle \rangle)$, the valuation function v' is defined by $v'(w') = v(u_k)$.

Let $w' = (u_0, \langle u_1, \langle Q_1^0 \dots Q_n^0 \rangle \rangle, \dots, \langle u_k, \langle Q_1^{k-1} \dots Q_n^{k-1} \rangle \rangle)$ and $Z : W \rightarrow 2^{W'}$ defined such that $w' \in Z(u_k)$. Clearly Z is a bisimulation between \mathcal{M} and \mathcal{M}' . ■

As an illustration, consider a *pre-ATS*⁷ \mathcal{M} over two agents a and b (left part of Figure 2). From u_0 , agent a can choose either $Q_a^1 = \{u_1, u_2\}$ or $Q_a^2 = \{u_2, u_3\}$. Agent b can choose either $Q_b^1 = \{u_2\}$ or $Q_b^2 = \{u_1, u_3\}$. Clearly \mathcal{M} is not choice partitioned since $\{Q_a^1, Q_a^2\}$ is not a partition of $\text{Succ}(u_0)$ ($Q_a^1 \cap Q_a^2 \neq \emptyset$).

We construct the equivalent choice partitioned ATS $\mathcal{M}' = \langle W', \delta', v' \rangle$ by duplicating u_2 which can be reached by applying two different choice profiles (right part of Figure 2).

⁷ Valuation and transition functions from u_1 , u_2 and u_3 are irrelevant.

Members of W' are thus $u'_0 = (u_0)$, $u'_1 = (u_0, \langle u_1, \langle Q_a^1, Q_b^2 \rangle \rangle)$, $u'_2 = (u_0, \langle u_2, \langle Q_a^1, Q_b^1 \rangle \rangle)$, $u'_3 = (u_0, \langle u_2, \langle Q_a^2, Q_b^1 \rangle \rangle)$ and $u'_4 = (u_0, \langle u_3, \langle Q_a^2, Q_b^2 \rangle \rangle)$. The transition function at u'_0 is represented by $\delta'(u'_0, a) = \{\{u'_1, u'_2\}, \{u'_3, u'_4\}\}$ and $\delta'(u'_0, b) = \{\{u'_1, u'_4\}, \{u'_2, u'_3\}\}$.

Lemma 1 permits us, without loss of generality, to consider only tree-like choice partitioned ATSS. Wölfl calls these ATSS ‘restricted’. However, as the semantic equivalence shows, this restriction is not a restriction from the viewpoint of modal logic. We come back to the equivalence property in Section 6.

5 From ATL to STIT logic

We define the translation tr from ATL formulae to STIT formulae as:

$$\begin{aligned} tr(p) &= \Box p, \text{ for } p \in \mathcal{A}tm \\ tr(\neg\varphi) &= \neg tr(\varphi) \\ tr(\varphi \vee \psi) &= tr(\varphi) \vee tr(\psi) \\ tr(\langle\langle A \rangle\rangle \mathbf{X}\varphi) &= \diamond_s[A \text{ scstit } \mathbf{X}tr(\varphi)] \\ tr(\langle\langle A \rangle\rangle \mathbf{G}\varphi) &= \diamond_s[A \text{ scstit } \mathbf{G}tr(\varphi)] \\ tr(\langle\langle A \rangle\rangle \varphi \mathcal{U} \psi) &= \diamond_s[A \text{ scstit } tr(\varphi) \mathcal{U} tr(\psi)] \end{aligned}$$

Translating an atom p into a modal formula $\Box p$ may seem odd, but is motivated by the remark in Section 3. All other clauses of the translation are straightforward, given the intended interpretation of the operators. The remaining section is devoted to the proof of the correctness of tr .

Given a tree-like choice partitioned ATS $\mathcal{M}_{ATL} = \langle W_{ATL}, \delta, v_{ATL} \rangle$ we associate to it a STIT model $\mathcal{M}_{STIT} = \langle W_{STIT}, \text{Choice}, <, v_{STIT} \rangle$, as follows:

- $W_{STIT} = W_{ATL}$,
- $w < u \iff \exists u_1, \dots, u_n (u_1 = w, u_n = u, \forall i < n (\exists a \in \mathcal{A}gt, Q_a \in \delta(u_i, a), u_{i+1} \in Q_a))$,
- $\text{Choice}_a^w = \{\{h \mid Q_a \cap h \neq \emptyset\} \mid Q_a \in \delta(w, a)\}$ for all a and w ,
- $\forall h \in H_w, v_{STIT}(w/h) = v_{ATL}(w)$.

It is clear that the tree property is instrumental for $\langle W_{STIT}, < \rangle$ being a tree. We inherit the branching-time structure of STIT directly from the tree structure of the ATS. Furthermore, the condition concerning partitions underlying choice partitioned ATSS prevents that two choices of the same agent have a non-empty intersection, and therefore every Choice_a^w is a partition of H_w . If intersections would possibly be non-empty, we could not have constructed the Choice function as we did: the same history could have been in two different sets of Choice_a^w .

PROPOSITION 2

\mathcal{M}_{STIT} is a discrete STIT model, and \mathcal{M}_{STIT} is unique.

PROOF. Straightforward. ■

In the following, \mathcal{M}_{STIT} histories are *maximal* sequences of ATL states respecting $<$. Given a history $h = \{w_0, w_1, \dots\}$ we can construct an infinite sequence of states $\lambda = q_0 q_1 \dots$ such that: $\forall q_i \in \lambda, \exists w_j \in h$ s.t. $q_i = w_j$, $q_i < q_{i+1}$ and $\nexists w \in h, q_i < w < q_{i+1}$ (since we have identified W_{ATL} with W_{STIT} , we can thus order members of W_{ATL} with the relation $<$). At such a

condition we will say that $h = \lambda$ (slightly abusing notation). Thus, we will indifferently use a STIT history and the corresponding ATL sequence of states.

LEMMA 2

Let $u \in W_{\text{ATL}}$ be a state in \mathcal{M}_{ATL} . For every collective ATL strategy F_A from \mathcal{M}_{ATL} , there is a collective STIT strategy $\sigma_A \in \text{Strategy}_A^u$ such that $\text{out}(u, F_A) = \text{Adh}(\sigma_A)$.

PROOF. We assume w.l.o.g. that the ATS of \mathcal{M}_{ATL} is a tree-like choice partitioned structure. Let $\text{path} : W_{\text{STIT}} \rightarrow W_{\text{ATL}}^+$ map each moment w into the (unique) *maximal* ordered sequence of states terminated by w . For all f_a of the tuple F_A we construct σ_a s.t.: for all $u \in W_{\text{STIT}}$ and $w' \in \text{Tree}_u$ we have

$$\sigma_a(w') = \{h | f_a(\text{path}(w')) \cap h \neq \emptyset\}$$

We let $\sigma_a(w')$ undefined for w' outside Tree_u . Let $\sigma_A = \langle \sigma_a \rangle_{a \in A}$, we want to show that $\text{out}(u, F_A) = \text{Adh}(\sigma_A)$.

(\Rightarrow) Suppose $\lambda \in \text{out}(u, F_A)$. It means $\lambda = q_0 q_1 \dots$ with $q_0 = u$ and $\forall i \geq 0, q_{i+1} \in \bigcap_{a \in A} f_a(q_0 \dots q_i)$. According to the construction of σ_a , we can say that $\forall i \geq 0, \forall a \in A, \{h | q_{i+1} \in h\} \subseteq \sigma_a(q_i)$, and then $\{h | q_{i+1} \in h\} \subseteq \sigma_A(q_i)$. Then the concatenation $\text{path}(u)\lambda \in \text{Adh}(\sigma_A)$ and thus $\text{out}(u, F_A) \subseteq \text{Adh}(\sigma_A)$.

(\Leftarrow) Suppose $h \in \text{Adh}(\sigma_A)$, and $\sigma_A \in \text{Strategy}_A^u$. This means that $h \in \text{Adh}(\sigma_a)$ for all $a \in A$: therefore we have (i) $\text{Dom}(\sigma_a) \cap h \neq \emptyset$ and (ii) $\forall w \in \text{Dom}(\sigma_a) \cap h, h \in \sigma_a(w)$. By definition, $u \in \text{Dom}(\sigma_a) \cap h, \forall a \in A$. According to the construction of σ_a we can say that for all $w \in h$ that appear in Tree_w , $f_a(\text{path}(w)) \cap h \neq \emptyset$, and therefore $(\bigcap_{a \in A} f_a(\text{path}(w))) \cap h \neq \emptyset$. Because h is a maximal set of linearly ordered moments from W containing u , we have that $h = \text{path}(u)q_1 q_2 \dots$ with $q_1 = \bigcap_{a \in A} f_a(\text{path}(u))$, and such that $q_{i+1} \in \bigcap_{a \in A} f_a(\text{path}(q_i))$. Then $h \in \text{out}(u, F_A)$ and $\text{Adh}(\sigma_A) \subseteq \text{out}(u, F_A)$.

We conclude that $\text{out}(u, F_A) = \text{Adh}(\sigma_A)$. ■

THEOREM 1

If φ is ATL-satisfiable then $\text{tr}(\varphi)$ is STIT-satisfiable.

PROOF. Suppose given an ATS $\mathcal{M}_{\text{ATL}} = \langle W_{\text{ATL}}, \delta, v_{\text{ATL}} \rangle$ and $w \in W_{\text{ATL}}$ s.t. $\mathcal{M}_{\text{ATL}}, w \models \varphi$. W.l.o.g. \mathcal{M}_{ATL} is tree-like. We translate it into $\mathcal{M}_{\text{STIT}} = \langle W_{\text{STIT}}, \text{Choice}, <, v_{\text{STIT}} \rangle$, as described above. Hence by Proposition 2, $\mathcal{M}_{\text{STIT}}$ is a STIT model. We prove by structural induction on φ that $\mathcal{M}_{\text{ATL}}, w \models \varphi$ iff $\mathcal{M}_{\text{STIT}}, w/h \models \text{tr}(\varphi), \forall h \in H_w$.

Cases of atomic formulae, negations and disjunctions are trivial, and we here only present the cases of the modal operators.

- Case $\psi = \langle \langle A \rangle \rangle \mathbf{X}\gamma$. This means that there is an F_A s.t. for all $\lambda \in \text{out}(w, F_A)$ we have $\mathcal{M}_{\text{ATL}}, \lambda[1] \models \gamma$. So by induction hypothesis, for all $\lambda \in \text{out}(w, F_A)$ we have $\mathcal{M}_{\text{STIT}}, \lambda[1]/h \models \text{tr}(\gamma)$ for all $h \in H_{\lambda[1]}$. By Lemma 2, we know that we can construct a collective strategy $\sigma_A \in \text{Strategy}_A^w$ s.t. $\text{out}(w, F_A) = \text{Adh}(\sigma_A)$. So, there is σ_A s.t. for all $h \in \text{Adh}(\sigma_A)$, we have $\mathcal{M}_{\text{STIT}}, \lambda[1]/h \models \text{tr}(\gamma)$. By construction of $<$, and according to the definition of the \mathbf{X} -operator, this means that $\mathcal{M}_{\text{STIT}}, w/h \models \mathbf{X}\text{tr}(\gamma)$, and we obtain that $\mathcal{M}, w/h \models \diamond_s[A \text{scstit} : \mathbf{X}\text{tr}(\gamma)]$.
- Case $\psi = \langle \langle A \rangle \rangle \mathbf{G}\gamma$. This means that there is an F_A s.t. for all $\lambda \in \text{out}(w, F_A)$ we have $\mathcal{M}_{\text{ATL}}, \lambda[i] \models \gamma, \forall i \geq 0$. By induction hypothesis, for all $\lambda \in \text{out}(w, F_A)$ we have

$\mathcal{M}_{\text{STIT}}, \lambda[i]/h \models \gamma, \forall i \geq 0, \forall h \in H_{\lambda[i]}$. By Lemma 2, there is $\sigma_A \in \text{Strategy}_A^w$ s.t. for all $h \in \text{Adh}(\sigma_A)$, we have $\mathcal{M}_{\text{STIT}}, \lambda[i]/h \models \text{tr}(\gamma), \forall i \geq 0$. By construction of $<$, and according to the definition of the **G**-operator, this means that $\mathcal{M}_{\text{STIT}}, w/h \models \mathbf{Gtr}(\gamma), \forall h \in \text{Adh}(\sigma_A)$, and we obtain that $\mathcal{M}_{\text{STIT}}, w/h \models \diamond_s[A \text{ scstit} : \mathbf{Gtr}(\gamma)]$.

- Case $\psi = \langle\langle A \rangle\rangle \gamma_1 \mathcal{U} \gamma_2$. This means that there is an F_A s.t. for all $\lambda \in \text{out}(w, F_A)$ there exists an $i \geq 0$ s.t. we have $\mathcal{M}_{\text{ATL}}, \lambda[i] \models \gamma_2$ and $\forall j, 0 \leq j < i, \mathcal{M}_{\text{ATL}}, \lambda[j] \models \gamma_1$. Using the same arguments as before, we get $\mathcal{M}_{\text{STIT}}, w/h \models \diamond_s[A \text{ scstit} : \text{tr}(\gamma_1) \mathcal{U} \text{tr}(\gamma_2)]$ for all h in H_w . ■

In addition, for the STIT-fragment corresponding to ATL, it holds that evaluation of formulas does not depend on the history (Horty calls this ‘moment determinedness’). This corresponds with the following property.

PROPOSITION 3

$$\models_{\text{STIT}} \text{tr}(\varphi) \equiv \square \text{tr}(\varphi)$$

PROOF. The proof is done by induction on the form of φ . It uses the fact that the logic of historical necessity \square is S5. ■

We need this proposition in our proof of Theorem 2.

THEOREM 2

If $\models_{\text{ATL}} \varphi$ then $\models_{\text{STIT}} \text{tr}(\varphi)$.

PROOF. We use the ATL axiomatization of [15], and prove that translation of the axioms are valid, and that the translated inference rules preserve validity.

(\perp), (\top), (N), (S) and ($\langle\langle A \rangle\rangle \mathbf{X}$ -Monotonicity) are axioms of Coalition Logic. Their translation to STIT preserves validity, as we have shown in [6, Theorem 4.2].⁸

If a formula is STIT-valid, it is true at each index of each STIT model. Then, it is obvious that the translation of ($\langle\langle \emptyset \rangle\rangle \mathbf{G}$ -Necessitation) preserves validity.

- The translation of (FP_G) is

$$\diamond_s[A \text{ scstit} : \mathbf{Gtr}(\varphi)] \equiv \text{tr}(\varphi) \wedge \diamond_s[A \text{ scstit} : \mathbf{X} \diamond_s[A \text{ scstit} \mathbf{Gtr}(\varphi)]]$$

(\Rightarrow) The left side of the equivalence implies that there is an index where $\text{tr}(\varphi)$ holds. By Proposition 3, $\models_{\text{STIT}} \text{tr}(\varphi) \rightarrow \square \text{tr}(\varphi)$, and thus $\text{tr}(\varphi)$ is true at any index of the current moment. If there exists a strategy such that $\text{tr}(\varphi)$ is globally true along admitted histories, then the same strategy also satisfies the right part of the equivalence. (\Leftarrow) The right side says there is a strategy σ_A at w , let us say with $\sigma_A(w) = Q, Q \in \text{Choice}_A^w$, s.t. at the next step, there is a strategy σ'_A s.t. $\text{tr}(\varphi)$ is globally true. Hence, the strategy σ''_A at w , defined as $\sigma''_A(w) = Q$ and $\forall u \in \text{Dom}(\sigma'_A) \setminus \{w\}, \sigma''_A(u) = \sigma'_A(u)$ satisfies that A can ensure at w that $\text{tr}(\varphi)$ is globally true along histories in $\text{Adh}(\sigma''_A)$.

- The translation of (GFP_G) is

$$\begin{aligned} \diamond_s[\emptyset \text{scstit} : \mathbf{G}(\text{tr}(\theta) \rightarrow (\text{tr}(\varphi) \wedge \diamond_s[A \text{ scstit} : \mathbf{X} \text{tr}(\theta)]))] \rightarrow \\ \diamond_s[\emptyset \text{scstit} : \mathbf{G}(\text{tr}(\theta) \rightarrow \diamond_s[A \text{ scstit} : \mathbf{Gtr}(\varphi)])] \end{aligned}$$

⁸ The proof of the validity of the translation of the axiom (N) involves Assumption 2 about determinism.

The left member means that it is settled that globally, if we have $tr(\theta)$ then we also have $tr(\varphi)$, and there is strategy s.t. $tr(\theta)$ is true at the next step. It implies that whenever $tr(\theta)$ is true, it exists a choice partition that ensures that $tr(\theta)$ holds at the next step. Thus the strategy σ_A which as soon as $tr(\theta)$ is true, chooses at each step such a choice partition, ensures that $tr(\varphi)$ is globally true along histories of $Adh(\sigma_A)$ (and this, whatever we choose before getting $tr(\theta)$).

- The translation of (FP_U) is

$$\begin{aligned} \diamond_s[A \text{ scstit} : tr(\psi)\mathcal{U}tr(\varphi)] &\equiv \\ tr(\varphi) \vee (tr(\psi) \wedge \diamond_s[A \text{ scstit } \mathbf{X} \diamond_s[A \text{ scstit} : tr(\psi)\mathcal{U}tr(\varphi)]]) & \end{aligned}$$

We prove its validity by using the same arguments as for (FP_G) .

- The translation of (LFP_U) is

$$\begin{aligned} \diamond_s[\emptyset \text{ scstit} : \mathbf{G}((tr(\varphi) \vee (tr(\psi) \wedge \diamond_s[A \text{ scstit } \mathbf{X}tr(\theta)])) \rightarrow tr(\theta))] &\rightarrow \\ \diamond_s[\emptyset \text{ scstit} : \mathbf{G}(\diamond_s[A \text{ scstit} : tr(\psi)\mathcal{U}tr(\varphi)] \rightarrow tr(\theta))] & \end{aligned}$$

We use the fact that $\diamond_s[\emptyset \text{ scstit} : \varphi] \equiv \Box\varphi$ (Proposition 1), that $\Box\mathbf{G}(\varphi \rightarrow \psi) \rightarrow (\Box\mathbf{G}\varphi \rightarrow \Box\mathbf{G}\psi)$ and that $(\alpha \rightarrow (\beta \rightarrow \gamma)) \equiv (\beta \rightarrow (\alpha \rightarrow \gamma))$. Thus, we have to prove that $\beta \rightarrow (\alpha \rightarrow \gamma)$ with $\beta \equiv \Box\mathbf{G}\diamond_s[A \text{ scstit} : tr(\psi)\mathcal{U}tr(\varphi)]$, $\alpha \equiv \Box\mathbf{G}((tr(\varphi) \vee (tr(\psi) \wedge \diamond_s[A \text{ scstit} : \mathbf{X}tr(\theta)])) \rightarrow tr(\theta))$ and $\gamma \equiv \Box\mathbf{G}tr(\theta)$.

Suppose that $\mathcal{M}, w/h \models \diamond_s[A \text{ scstit} : tr(\psi)\mathcal{U}tr(\varphi)]$. This means that there is a strategy σ_A s.t. $\forall h \in Adh(\sigma_A), \exists w_1 \in h, w < w_1$ s.t. $\mathcal{M}, w_1/h \models tr(\varphi)$ and $\forall w_2, w \leq w_2 < w_1, \mathcal{M}, w_2/h \models tr(\psi)$. By α , $tr(\theta)$ is true at w_1 . If $w_1 = w$ then it is sufficient to conclude. Else, we have $tr(\psi)$ true at the immediate predecessor of w_1 on h . So by α , we also have $tr(\theta)$, since $\diamond_s[A \text{ scstit} : tr(\theta)]$ is true. Still, recursively (this induction is allowed by countably infiniteness of Assumption 1) as $tr(\psi)$ is true at each $w_3 \in h$ s.t. $w \leq w_3 < w_1$, we also get $tr(\theta)$ at w_3 and in particular $\mathcal{M}, w/h \models tr(\theta)$. ■

COROLLARY 1

φ is satisfiable in ATL iff $tr(\varphi)$ is satisfiable in STIT.

PROOF. As an immediate corollary of Theorems 1 and 2. ■

6 Discussion

The main contribution of this work has been, we believe, to build a bridge between two formalisms with a rather different background; the STIT formalism originating in philosophy, and ATL originating in computer science (multiagent systems). In this section, we discuss details of our embedding. We address in what sense, and under what assumptions, ATL appears to be a well-identified fragment of a more general and philosophically grounded theory of agency. These assumptions are then insightful and suggestive of a shared core between computer science and the philosophy of agency/action.

It should be noted first that Horty's strategic ability only applies to *individual* agency. Hence, we had to define admitted histories for a *collective* strategy, as the intersection of individual ones. However, this is a straightforward extension of the definition of collective

choices; we believe we have neither violated a fundamental aspect of STIT nor forced the embedding by adding too much to the semantics.

We also added some constraints to the original STIT to guarantee that the proposed translation works well. We view these constraints as both relevant and harmless. The constraints are:

1. Histories are isomorphic to the set of natural numbers.
2. $\forall w \in W, \exists w' \in W (w < w' \text{ and } \bigcap_{a \in \mathcal{A}gt} s_w(a) = H_{w'})$
 Intersection of agents of $\mathcal{A}gt$'s choices is *not only* non-empty (which is the only restriction in the original STIT) but must exactly be the set of histories passing through a next moment.

The second condition is the simple counterpart of the ATL constraint stating that when every agent in $\mathcal{A}gt$ opts for an action then the next state of the world is completely determined. Here we just say that in STIT, the intersection of all agents of $\mathcal{A}gt$'s choices must be exactly the set of histories passing through this very completely determined moment.⁹ As discussed in [13], the condition of determinism is not a limitation of the modelling capabilities of the language, since we could introduce a neutral agent 'nature', in order to accommodate non-deterministic transitions. Hence, this constraint on ATSS should not be considered a fundamental distinction between the two formalisms.

The main difference then concerns the first constraint, that permits us to define the **X** operator, and then to grasp the concept of *next* moments and outcomes. More generally, it allows us to stick to standard LTL expressivity for temporal properties of paths. This same assumption applies to the temporal component of ATL. This imposes a particular view on time. However, deliberately, Belnap and colleagues do not take a position on the nature of time.

'For this reason the present theory of agency is immediately applicable regardless of whether we picture succession as discrete, dense, continuous, well-ordered, some mixture of these, or whatever; and regardless of whether histories are finite or infinite in one direction or the other.' ([10, p. 196].)

Although, from a philosophical point of view, it makes sense wanting to be as general as possible, in computer science it is very common and natural to model the temporal evolution of a system using a transition system. This brings with it a view on time as being discrete. Isomorphism with the natural numbers (and thus non-density) is often assumed in order to keep complexity within acceptable limits, and to avoid discussions about philosophical difficulties reminiscent of problems raised by presocratic philosophers typified by Zeno of Elea: how can time proceed (i.e. how can we interpret a 'next' operator) if there is always a moment between two moments? This justifies the assumption concerning isomorphism with the natural numbers.

However the differences in the temporal fragments of both frameworks do not only concern the models, but also the syntax. In particular, note that in STIT we can nest temporal operators without any restriction. In ATL this is syntactically disallowed. In ATL* we do not have this restriction. However, in some definitions for this stronger logic we cannot unravel ATSS into trees under preservation of satisfaction of formulas.

⁹ Actually, this condition does not explicitly refer to the *next* moment, but to a future moment. It is nevertheless sufficient, because for all $h \in H_w$, and for all $w' < w$, we have $h \in H_{w'}$. ($(W, <)$ is a tree.)

Obviously, STIT and ATL have some striking resemblances. The concepts of *agent* and *choice* are the same in both theories. In STIT agents are ‘individuals thought of as making choices, or acting, in time’ ([10, p. 33]). Belnap, as a founding father of STIT theory, has stressed that STIT agency is not restricted to persons or intentional agents and could equally be applied to processes making random choices. Actions are thus idealized in a way that ignores any mental state. STIT is only interested in the causal structure of choice, regardless of its content. To put it in yet other words, choices are just *objective possibilities* of an agent, selecting some possible courses of time and ruling out some others. All of this equally applies to ATL, where each agent selects a set of next states, and time will go through a state in the intersection of every agent’s selection.

Also the notion of *independence of choices* (or equally *independence of agents*) applies to both frameworks. Agent’s choices must be non-blocking, i.e. for each possible choice of some agent, the intersection with all possible choices of other agents is non-empty. Belnap *et al.* admit this to be a fierce constraint. For instance, it follows that two agents cannot possibly have identical sets of choices at the same moment (except the vacuous one). It also follows that in STIT, there are not less than $\prod_{a \in \text{Agt}} |\text{Choice}_a^w|$ histories passing through a moment w . Nevertheless the constraint is considered commonplace. In STIT theory it has been argued that if an agent can deprive other agents of some of their choices, then, regardless possible priorities in the causal order, ‘we shall need to treat in the theory of agency a phenomenon just as exotic as those discovered in the land of quantum mechanics by Einstein, Podolsky, and Rosen’ [10, p. 218].

ATL structures are not limited to trees. But, as described in [27], an ATS $\langle W, \delta, v \rangle$ can easily be unraveled to an ATS where the transition function δ in $\langle W, <_\delta \rangle$ is a tree. ATL, like all other modal formalisms,¹⁰ cannot distinguish the original model from its unraveling into a tree. STIT and ATL thus both embed in branching-time structures limited to trees. However, what we show in Section 4 is stronger. Lemma 1 tells us that we can unravel any ATS in a tree satisfying the property that choices of every agent, represented as sets of ‘possibly chosen next states’, are partitioning the ‘possible next states’. Hence, from any ATS, we can construct a bisimilar ATS that meets the constraint STIT imposes to the Choice_w^a functions. There is no need to enforce this on ATL frames as in [27]. The property of *empty intersection* of the different simultaneous classes of choice in STIT is not expressible in modal logic.

It is worth noting that the present translation is compatible with the one we have proposed in [6] for Coalition Logic, together with Goranko’s translation of $[A]\varphi$ to $\langle\langle A \rangle\rangle X\varphi$.

With the completeness result for ATL in [15], one immediate benefit of our translation is to identify a complete axiomatization of a fragment of STIT. In this sense it completes Ming Xu’s work, compiled in [10, Part VI] about decidability, soundness and completeness of fragments of the achievement STIT and deliberative STIT logics. As an interesting perspective, ATL model checking can be applied to a fragment of the STIT language.

A challenging research avenue is to import deontic concepts that have been investigated in the STIT framework [16, 17] into ATL. It appears to us that this can be done in a rather straightforward manner. We could then further the discussion initiated in [20] and [11] concerning how to model obligations in ATL. Likewise, the problem of how to accommodate

¹⁰ At least this is true according to van Benthem’s definition of ‘modal logic’ as the bisimulation-invariant subset of first-order logic [5].

epistemic notions in ATL may benefit from the link with the STIT framework. Alternating-time Temporal Epistemic Logic (ATEL) [26] adds to ATL operators representing knowledge. Its aim is to deal with strategies in a context of incomplete information. One of the challenges has been identified in [19]: ATEL is not expressive enough when it questions availability of strategies. As far as we know, there is no satisfactory solution using ATL. In [18] we have proposed a solution in the STIT framework and thus proved that STIT is particularly relevant for the analysis of multi-agent systems.

Finally, a legitimate question would be: ‘can STIT also be embedded in ATL?’ We do not think. As mentioned, the temporal fragment of STIT allows arbitrary nesting of temporal operators. This kind of expressivity would require ATL^* as a target logic for an embedding. Another problem is that STIT operators are not moment determinate; they evaluate to different values depending on the history. This means that STIT theory has operators for two separate dimensions: historical necessity and possibility operators for the dimension of histories, and STIT operators for the dimension of moments. In ATL, these two dimensions are not both explicitly equipped with operators; the central operator is one-dimensional. In [7] we show how we can add STIT expressivity to ATL. Indeed this involves turning the semantics of ATL in a two-dimensional one.

We conclude with the remark [10, p. 18] that STIT theory should be understood as a formal characterization of agency, permitting to postpone an ontology. One merit of this work is then to push a significant justification for ATL as an elegant and well-founded framework of agency.

Acknowledgements

We are grateful to anonymous reviewers, whose critical remarks have improved the article. Nicolas Troquard is funded by the ‘Fondo provinciale per i progetti di ricerca’ of the Provincia Autonoma of Trento.

References

- [1] R. Alur, T. A. Henzinger and O. Kupferman. Alternating-time temporal logic. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, pp. 100–109. IEEE Computer Society Press, Los Alamitos, CA, 1997.
- [2] R. Alur, T. A. Henzinger and O. Kupferman. Alternating-time temporal logic. In *Compositionality: The Significant Difference*, Lecture Notes in Computer Science, Vol. 1536, pp. 23–60. Springer, Berlin, 1999.
- [3] R. Alur, T. A. Henzinger and O. Kupferman. Alternating-time temporal logic. *J. ACM*, **49**, 672–713, 2002.
- [4] P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, 2001.
- [5] J. F. A. K. van Benthem. Correspondence theory. In *Handbook of Philosophical Logic*, vol. II. D. M. Gabbay and F. Guenther, eds, Reidel, Dordrecht, 1984.
- [6] J. Broersen, A. Herzig and N. Troquard. From Coalition Logic to STIT. In *Proceedings of the Third International Workshop on Logic and Communication in Multi-Agent Systems (LCMAS 2005)*, W. van der Hoek, A. Lomuscio, E. de Vink and

- M. Wooldridge, eds, pp. 23–35. Edinburgh, Scotland, UK, Electronic Notes in Theoretical Computer Science, Vol. 157, No. 4, Elsevier, Amsterdam, 2006.
- [7] J. Broersen, A. Herzig and N. Troquard. A STIT extension of ATL. In *Proceedings Tenth European Conference on Logics in Artificial Intelligence (JELIA'06)*, M. Fisher, ed., Volume 4160 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, Berlin, 2006.
- [8] N. Belnap and M. Perloff. Seeing to it that: a canonical form for agentives. *Theoria*, **54**, 175–199, 1988.
- [9] N. Belnap and M. Perloff. Seeing to it that: A canonical form for agentives. In *Knowledge Representation and Defeasible Reasoning*, H. E. Kyburg, R. P. Loui and G. N. Carlson, eds, pp. 167–190. Kluwer, Boston, 1990.
- [10] N. Belnap, M. Perloff and M. Xu. *Facing the Future: Agents and Choices in our Indeterminist World*. Oxford University Press, Oxford 2001.
- [11] J. Broersen. Strategic deontic temporal logic as a reduction to ATL, with an application to Chisholm’s scenario. In *Proceedings 8th International Workshop on Deontic Logic in Computer Science (DEON 06)*, L. Goble and J.-J.Ch. Meyer, eds, Vol. 4048 of *Lecture Notes in Computer Science*, pp. 53–68. Springer, Berlin 2006.
- [12] D. Elgesem. *Action Theory and Modal Logic*. PhD Thesis, Department of philosophy, University of Oslo, 1993.
- [13] V. Goranko and W. Jamroga. Comparing semantics of logics for multi-agent systems. *Synthese*, **139**, 241–280, 2004.
- [14] V. Goranko. Coalition games and alternating temporal logics. In *TARK 01: Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 259–272. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001.
- [15] V. Goranko and G. van Drimmelen. Decidability and complete axiomatization of the alternating-time temporal logic. *Theor. Comput. Sci.*, **353**, 93–117, 2006.
- [16] J. F. Horty and N. D. Belnap, Jr. The deliberative stit: A study of action, omission, and obligation. *J. Philos. Logic*, **24**, 583–644, 1995.
- [17] J. F. Horty. *Agency and Deontic Logic*. Oxford University Press, Oxford, 2001.
- [18] A. Herzig and N. Troquard. Knowing how to play: uniform choices in logics of agency. In *5th International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS-06)*, G. Weiss and P. Stone, eds, pp. 209–216. ACM Press, New York, NY, Hakodate, Japan, 8–12 May 2006.
- [19] W. Jamroga. Some remarks on alternating temporal epistemic logic. In *Proceedings of the International Workshop on Formal Approaches to Multi-Agent Systems (FAMAS 03)*, 2003.
- [20] W. Jamroga, W. van der Hoek and M. Wooldridge. On obligations and abilities. In *Deontic Logic in Computer Science, Lecture Notes in Artificial Intelligence*, A. Lomuscio and D. Nute, eds, Vol. 3065, pp. 165–181, Springer Verlag, Berlin, 2004.
- [21] M. Pauly. A modal logic for coalitional power in games. *J. Logic Comput.*, **12**, 149–166, 2002.
- [22] I. Pörn. *The Logic of Power*. Blackwell, Oxford, 1970.
- [23] A. N. Prior. *Past, Present, and Future*. Clarendon Press, Oxford, 1967.
- [24] R. Thomason. Indeterminist time and truth-value gaps. *Theoria*, **36**, 264–81, 1970.
- [25] R. H. Thomason. Combinations of tense and modality. In *Handbook of Philosophical Logic: Extensions of Classical Logic*, D. Gabbay and F. Guentner, eds, pp. 135–165. Reidel, Dordrecht, 1984.

- [26] W. van der Hoek and M. Wooldridge. Tractable multi-agent planning for epistemic goals. In *AAMAS '02: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1167–1174, ACM Press, New York, 2002.
- [27] S. Wölfl. Qualitative action theory: a comparison of the semantics of alternating time temporal logic and the Kutschera-Belnap approach to agency. In *Proceedings of the Ninth European Conference on Logics in Artificial Intelligence (JELIA 04)*, J. Alferes and J. Leite, eds, Vol. 3229 of *Lecture Notes in Artificial Intelligence*, pp. 70–81. Springer, Berlin, 2004.

Received 10 February 2006