Knowing How to Play: Uniform Choices in Logics of Agency

Andreas Herzig
CNRS, Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier
Toulouse, France
herzig@irit.fr

Nicolas Troquard
Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier, Toulouse, France
and
Laboratorio di Ontologia Applicata
Università degli Studi di Trento, Trento, Italy
troquard@irit.fr

ABSTRACT

Reasoning about capabilities, strategies and knowledge is important in the analysis of multiagent systems. Alternating-time Temporal Epistemic Logic (ATEL) was designed with this aim. Nevertheless, the original interpretation of the language suffered from some counterintuitive properties. These are due to the fact that the strategies the agent applies in worlds that he cannot distinguish may not be uniform, in the sense that the same action is applied in all indistinguishable worlds. Several refinements of the original ATEL semantics were proposed since then. In this paper we argue that the STIT framework can easily account for uniform strategies. STIT is a logic of agency that has been proposed in the 90ies in the domain of philosophy of action. It is the logic of constructions of the form "agent a sees to it that φ". To support our claim, we first present a straightforward solution in STIT logic augmented by a modal operator of knowledge. Then we offer a simplification, by introducing a modal logic of knowledge-based uniform agency, for *one-step* strategies, alias choices.

Categories and Subject Descriptors

I.2.11 [**Distributed artificial intelligence**]: multiagent systems

General Terms

Theory

Keywords

formal models of agency, logics for agent systems, modal logic

1. INTRODUCTION

In the last years there has been increasing interest in logics enabling reasoning about strategies of agents and coalitions of agents, and the agents' knowledge about such strategies. Such logics combine two kinds of modal logics:

- logics of knowledge such as S5, and multiagent versions thereof; such logics have modal operators K_a, where $K_a\varphi$ is read "agent a knows that φ";
- logics of agency, including in particular Coalition Logic (CL) and Alternating-time Temporal Logic (ATL); such logics have constructions such as CL's $[A]\varphi$ or ATL's $\langle\langle A\rangle\rangle\mathbf{X}\varphi$, both (roughly) reading "group of agents A has an action to ensure that φ holds (whatever the other agents choose to do)".

While each of these logics is by now well-established, the interaction between knowledge and agency is less consensual. A straightforward combination of for example ATL and epistemic logic (called ATEL) was proposed in [14]. In ATEL one can express things such as "agent a has an action to ensure that φ, but ignores that". It turned out that ATEL has some counterintuitive properties. The problem can be highlighted by the following example.

Example 1. *There is a switch, a lamp, and a blind agent a_1, which ignores whether the light is on or off. a_1 can toggle the switch (and he knows that), and a_1 can remain passive.*

Clearly, $\langle\langle\{a_1\}\rangle\rangle\mathbf{X}\mathtt{light}$ holds here, i.e., a_1 can ensure that the light is on (viz. by toggling the switch if the light is off, and by doing nothing if the light is already on). We should also be able to conclude that a_1 does not *know* which action to perform in order to do this.

ATEL makes us conclude here that $K_{a_1}\langle\langle\{a_1\}\rangle\rangle\mathbf{X}\mathtt{light}$, i.e. the blind agent a_1 knows that he has an action to ensure the light is on. The problem is that this action corresponds to w a strategy that is what has been called *non-uniform*: it makes a_1 choose different actions in possible worlds that are indistinguishable for him.

Multiagent variants of our example can be devised.

Example 2. *There are two blind agents and two toggle switches such that the light is on exactly when both switches are in the same position. a_1 can toggle the first switch and a_2 the second one.*

Contrary to intuitions, ATEL makes us conclude that there is both distributed and common knowledge that there is a joint strategy to switch the light on. Another version that we will study involves one blind agent and one lame agent.

EXAMPLE 3. *There is a switch, a lamp, a blind agent a_1 and a lame agent a_2 (who knows whether the light is on, but cannot toggle). Initially the light is off. This is known by a_2, and ignored by a_1.*

This example is less problematic since there is distributed knowledge about a joint strategy, however ATEL also claims that there is common knowledge.

Several authors have proposed modified versions of ATEL, trying to accommodate in one way or another the notion of uniform strategy [8, 9, 15]. It seems to be fair to say that all these tentatives resulted in rather complex formalisms with heavy notations, and that there is no consensus up to now what the appropriate logic of knowledge and strategies is.

In this paper we take as our starting point a slightly different logic of agency that has been developed in philosophical logic. Just as ATL, the logic of "Seeing To It That" (STIT) is a modal logic enabling us to speak about time and agents' choices of actions. In STIT, CL's and ATL's $\exists$-$\forall$-quantification ("*there is* a strategy of group A such that *for all* actions of the other agents") is split up into two different modal operators:

- an operator of historical possibility $\Diamond$;
- an operator of "seeing to it that" *Stit*.

In previous work [3] we have shown that STIT is at least as powerful as Coalition Logic (CL). We have proved this by giving a translation t from CL into STIT. The main clause of the translation maps CL's $[A]\varphi$ ("group A has an action to ensure that φ") into STIT's $\Diamond Stit_A \mathbf{X} t(\varphi)$ ("it is possible that group A sees to it that next $t(\varphi)$"). This translation can be extended to ATL [4].

If by coalition we mean a set of agents that are able to deliberate about their knowledge and the collective strategy to adopt then reasoning about *common knowledge* does not seem to be relevant when dealing with coalitions. Suppose a_1 cannot distinguish two worlds w_1 and w_2, and a_2 cannot distinguish w_2 and w_3, by common knowledge they together cannot distinguish w_1, w_2 and w_3. However, when standing at w_2, agents a_1 and a_2 by *communicating*, should be able to discriminate incompatible worlds: here w_1 and w_3. By definition, the intersection of a_1's and a_2's knowledge state corresponds to distributed knowledge of the coalition $\{a_1, a_2\}$.

We assume that a coalition can ensure φ if by *sharing their knowledge* and acting together they can ensure that φ. For that purpose, we introduce an operator $Kstit_A\varphi$ of *knowledge-based agency.* "Agent a knows an action to ensure that φ" can thus be expressed in our epistemic extension of STIT by: "there is an action such that the agent knows that φ holds next". In formulas: $\Diamond Kstit_a \mathbf{X}\varphi$. Thus in Example 1 we expect $\neg\Diamond Kstit_{a_1}\mathbf{X}\mathtt{light} \wedge K_{a_1}\Diamond Stit_{a_1}\mathbf{X}\mathtt{light}$ to hold. This corresponds to the classical distinction between de re sentences such as $\exists x K_a p(x)$ and de dicto sentences such as $K_a \exists x p(x)$.

We present a discrete version of STIT logic in Section 2 and augment it with an epistemic operator in Section 3. In Section 4, we show how the problem of enforcing uniform strategies can be treated straightforwardly in the resulting framework. We simplify it in Section 5 by presenting a logic of knowledge-based agency for groups of agents. Section 6 offers a formulation of "knowing how to play" and examples to illustrate the logic. We show in Section 7 how to handle static knowledge (or knowledge in the sense of Hintikka), and future work is discussed in Section 8.

2. DISCRETE STIT FRAMEWORK

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 nonempty set of moments, and $<$ is a tree-like ordering of these moments, such that 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.* $w \in h$ denotes that the moment w is *on* the history h. We define $Hist$ as the set of all histories of a STIT structure. $H_w = \{h | h \in Hist, w \in h\}$ 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). Moreover, we make the following hypothesis (that is not made in the original STIT for reasons of generality):

HYPOTHESIS 1 (DISCRETENESS). $<$ *is* discrete.

That is to say, given a moment w_1 and a history h with $w_1 \in h$, there exists a successor moment $w_2 \in h$ such that $w_1 < w_2$, and there is *no* moment w_3 such that $w_1 < w_3 < w_2$.

2.1 Models of individual agency

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

- $\langle W, < \rangle$ is a branching time structure;
- $Choice : \mathcal{A}gt \times W \to 2^{2^{Hist}}$ is a function mapping each agent and each moment w into a partition of H_w;
- v is a valuation function $v : \mathcal{A}tm \to 2^{W \times Hist}$.

The equivalence classes belonging to $Choice^w_a$ can be thought of as possible choices or actions available to agent a at w.

Given a history $h \in H_w$, $Choice^w_a(h)$ represents the particular choice from $Choice^w_a$ containing h, or in other words, the particular action performed by a at the index w/h. We must have $Choice^w_a \neq \emptyset$ and $Q \neq \emptyset$ for every $Q \in Choice^w_a$.

DEFINITION 1 (CURRENT MOMENT / CURRENT CHOICE). *At index w/h we shall call w the* current moment *and $Choice^w_a(h)$ the* current choice/action.

In STIT-models, moments may have different valuations, depending on the history they are living in (cf. [7, footnote 2 p. 586]). Thus, at any specific moment, we have different valuations. This was first suggested by Prior and Thomason, and permits a natural evaluation of statements about the future.

2.2 Models of group agency

In order to deal with group agency, Horty defines in [6, section 2.4], the notion of collective choice. Horty first introduces *action selection functions* s_w from $\mathcal{A}gt$ into 2^{H_w} associating to each $w \in W$ and $a \in \mathcal{A}gt$, some $s_w(a) \in Choice_a^w$. So, a given 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 $\bigcap_{a \in \mathcal{A}gt} s_w(a) \neq \emptyset$. This constraint corresponds to the hypothesis that the agents' choices are independent, in the sense that agents can never be deprived of some choice due to the choices made by other agents.[1] Moreover, we suppose that intersection of *all* agents' choices must exactly be the set of histories passing through an immediate next moment:

HYPOTHESIS 2 (DETERMINISM). $\forall w \in W, \exists w' \in W\, (w < w'$ *and* $\nexists w'' \in W, w < w'' < w', \bigcap_{a \in \mathcal{A}gt} s_w(a) = H_{w'})$

As explained for ATL in [5], this determinism is not a limitation of the modeling capabilities of the language, since we could introduce a neutral agent *nature*, in order to accommodate non-deterministic transitions.

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

$$Choice_A^w = \{ \bigcap_{a \in A} s_w(a) | s_w \in Select_w \}.$$

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

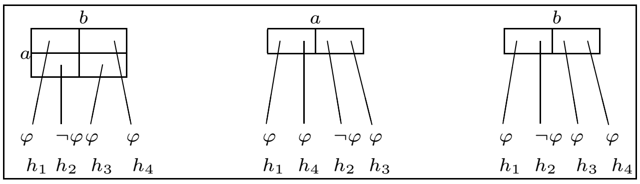

Figure 1: Group agency. $Choice_a^w = \{\{h_1, h_4\}, \{h_2, h_3\}\}$, $Choice_b^w = \{\{h_1, h_2\}, \{h_3, h_4\}\}$, $Choice_{\{a,b\}}^w = \{\{h_1\}, \{h_4\}, \{h_2\}, \{h_3\}\}$.

In Figure 1, a can choose between the upper and the lower row, and b can choose between the left and the right column. In the left model we can see that when a and b opt for a choice, the outcome is deterministic. However, actions of individual agents are not deterministic, as we can see on the two others schemes. And while both a and b can *see to it* that φ independently, they need to form a coalition in order to ensure $\neg\varphi$.

2.3 Truth conditions

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

$$\begin{array}{lll} \mathcal{M}, w/h \models p & \iff & w/h \in v(p), p \in \mathcal{A}tm. \\ \mathcal{M}, w/h \models \neg\varphi & \iff & \mathcal{M}, w/h \not\models \varphi \\ \mathcal{M}, w/h \models \varphi \vee \psi & \iff & \mathcal{M}, w/h \models \varphi \text{ or } \mathcal{M}, w/h \models \psi \end{array}$$

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; we here introduce the so-called Chellas' STIT [7] which is defined as follows:

$$\mathcal{M}, w/h \models Stit_A\varphi \iff \mathcal{M}, w/h' \models \varphi, \forall h' \in Choice_A^w(h).$$

Intuitively it means that group A's current choices ensure φ, whatever other agents outside A do.[2] $Stit_A\varphi$ corresponds to $[A\, cstit\colon \varphi]$ in original notation.

As it is shown in [7], both Chellas' STIT and historical necessity are S5 modal operators, and $\models_{\mathsf{STIT}} \Box\varphi \to Stit_A\varphi$.

Since we here have discrete time, we can define the temporal operator $\mathbf{X}$ (*next*). We also introduce operator $\mathcal{U}$ (*until*):

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

$\mathbf{F}\varphi$ ("at some point in the future φ") is an abbreviation of $\top\mathcal{U}\varphi$, and $\mathbf{G}\varphi$ ("always in the future φ") is $\neg F\neg\varphi$.

3. ADDING EPISTEMIC FEATURES

In this section, we add imperfect knowledge to the STIT framework, assuming a particular structure of epistemic relations which will be constructed above agent choices' relations.

As a related work, we refer the reader to Pacuit et al. [11], which is as far as we know, the only to deal with epistemic notions in the STIT framework: an agent a is said to know that φ at a history h, if φ is true in every history in a's *local view*. We can enlighten differences with our notion of knowledge by the motivations, which are to investigate the relation between knowledge and obligation in Pacuit et al.'s work, while we are interested in epistemic based uniform agency.

3.1 Individual knowledge

If two moments are not distinguishable by an agent according to his knowledge, these moments are said to be *epistemically equivalent*.

Since the purpose is here to reason about actions and choices under imperfect knowledge, we should be able to identify identical choices (or choices involving actions of the same type)[3] at epistemically equivalent worlds.

[1] Note that from this constraint it follows that two agents cannot possibly have an identical set of choices at the same moment. It also follows that there are not less than $\prod_{a \in \mathcal{A}gt} |Choice_a^w|$ histories passing through a moment w. Moreover, at moments where the minimal number of histories satisfies this constraint, choices at future moments will be vacuous.

[2] The more complex *deliberative* STIT can be defined as $Dstit_A\varphi =_{def} Stit_A\varphi \wedge \neg\Box\varphi$ [7].

[3] Intuitively, two actions are of the same *type* if *the way to*

Thus, in order to identify choices of the same type executable in moments epistemically equivalent, let $R_{UC_a} : W \times Hist \to 2^{W \times Hist}$ be a relation between indexes lying in equivalent choice types of epistemically equivalent moments for agent a. UC stands for "*U*niform *C*hoice". If two moment/history pairs are linked by a R_{UC_a} relation, the same type of choice is done by agent a at both indexes. We require

$$Choice_a^w(h) \subseteq \{h' \mid w'/h' \in R_{UC_a}(w/h)\}.$$

In Example 1, $R_{UC_{a_1}}$ respectively links indexes of the two moments that are possible for a_1 where the action `toggle` is executed, and indexes where a_1 does nothing (action `skip`).

For the sake of brevity we introduce $R_{UC} : \mathcal{A}gt \to (W \times Hist \to 2^{W \times Hist})$, which is a function associating a R_{UC_a} relation with each agent a. Hence, we obtain an *Epistemic STIT-model* (*ESTIT-model*) by adding R_{UC} to a *STIT-model*:

$$\mathcal{M}_{\mathsf{ESTIT}} = \langle W, <, Choice, R_{UC}, v \rangle$$

We can then offer a new kind of STIT operator $Kstit_a\varphi$, which reads "agent a knows that he ensures that φ". $Kstit_a$ can be seen as a Chellas' STIT adapted to imperfect knowledge:

$$\mathcal{M}, w/h \models Kstit_a\varphi \iff \forall w'/h' \in R_{UC_a}(w/h),\ \mathcal{M}, w'/h' \models \varphi.$$

It says that an agent ensures φ uniformly with its knowledge if among indistinguishable moments, and whatever the *current* one is, $Stit_a\varphi$ is true by doing the *current* choice type. Just as (static) knowledge, $Kstit_A$ is an $S5$ modality.

HYPOTHESIS 3. *Two moments w_1 and w_2 are indistinguishable for an agent a only if for every choice open to a at w_1, there is a choice of the same type open to him at w_2.*

A corollary is that each agent knows which choices are open to him, or still that two moments with different sets of choices cannot be epistemically indistinguishable. Some authors have already used such an assumption ([13, 11]).

3.2 Group knowledge

We assume that agents of a group share their knowledge. Distributed knowledge of a group A is usually defined as the intersection of knowledge of individual agents of A. Hence,

$$R_{UC_A}(w/h) = \bigcap_{a \in A} R_{UC_a}(w/h).$$

In a straightforward manner, we will thus say that a group of agents A knows it ensures φ if among their indistinguishable moments, and whatever the current one is, $Stit_A\varphi$ is true by doing the *current* action type.

$$\mathcal{M}, w/h \models Kstit_A\varphi \iff \forall w'/h' \in R_{UC_A}(w/h),\ \mathcal{M}, w'/h' \models \varphi.$$

In Example 3 we intuitively have that $\Diamond Kstit_{\{a_1,a_2\}}\mathbf{Xlight}$, while in Example 2 we have $\neg\Diamond Kstit_{\{a_1,a_2\}}\mathbf{Xlight}$.

produce them or *the bodily movement part of the action* is the same. For example *turning off the light*, *turning on the light*, *toggle the switch* or *push the switch* are of the same type. Our aim is not to debate about types, so we stay at this intuitive definition.

4. UNIFORM STRATEGIES

In this section, we expose the problem of *uniform strategy* with a short summary of the state of the art of existing solutions. In order to support our claim that the STIT framework is appropriate for that purpose, we then show that we can adapt the semantics of the strategic STIT to obtain a rather simple solution to that problem.

4.1 The uniform strategies problem

Van der Hoek and Wooldridge's Alternating-time Temporal Epistemic Logic (ATEL) [14] was designed with the aim of reasoning about strategies and knowledge. Nevertheless, the original interpretation of the language suffered from counterintuitive properties. Situations such as Example 1 cannot be captured in such a way.

Some refinements were proposed [8, 9, 15]. They all aim at capturing the notion of a *uniform strategy*. Informally, a strategy σ is uniform if, when it is defined on a moment w epistemically equivalent to a moment w', then σ is also defined on w' and the choice type planned by the strategy is the same in both moments. In Example 1, a_1 has *no* uniform strategy to ensure `light`.

As far as we know, the most complete answer to the problem is offered in [9]. The authors propose two extensions to ATEL: ATOL, a logic of observation for agents with bounded recall of the past, and ATEL-R* for agents with both perfect and imperfect recall, and possibly reasoning about the past. Both formalisms are rather complex. In particular, the agency operator $\langle\langle A \rangle\rangle$ is indexed by epistemic formulas. This is corrected in a very recent work by Jamroga and Ågotnes [10]. They propose an interesting solution, but at the price of a non-standard negation operator (called *constructive* negation) which seems necessary to handle reasoning about non-strategic properties in their semantics.

4.2 Strategies and uniform strategies in STIT

Horty [6] and Belnap et al. [2] introduce strategies into STIT theory.

DEFINITION 2. *A* strategy *for an agent a is a partial function σ such that $\sigma(w) \in Choice_a^w$ for each moment w from $Dom(\sigma)$, the domain of σ.*[4]

As we can see in the definition of the *Stit* operator, an agent's choice restricts the set of possible futures, in particular it restricts the histories to those corresponding with the choice being made. We expect a strategy to be a generalization of this, restricting histories to a set compatible to every choices at every moment of the domain.

DEFINITION 3. *We say that a strategy σ* admits *a history h if and only if*

- *$Dom(\sigma) \cap h \neq \emptyset$ and*
- *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 noted $Adh(\sigma)$.

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

[4] σ is a partial function because there is no need to represent in σ an agent's choices in moments he will never reach if he follows σ.

Definition 4 (collective strategy). *A collective strategy for $A \subseteq \mathcal{A}gt$ is a tuple $\sigma_A = \langle \sigma_a \rangle_{a \in A}$ (one strategy σ_a for every agent), and $Adh(\sigma_A) = \bigcap_{a \in A} Adh(\sigma_a)$.*

[6] proposes a characterization of a sort of properly-formed strategies with limited scope. It is necessary to introduce the notion of *field* at a moment w, which is a subset containing w of the subtree $Tree_w = \{w' \mid w < w' \text{ or } w = w'\}$ with a downward closure. With $Adm(\sigma) = \{m \mid m \in h, h \in Adh(\sigma)\}$, a strategy is properly-formed in M if it is *complete in the field* ($Adm(\sigma) \cap M \subseteq Dom(\sigma)$) and *irredundant* ($Dom(\sigma) \subseteq Adm(\sigma)$). Thus, an *ability* operator should be evaluated with respect to a field. But for simplicity, we do not use it here.[5]

However, just as it is done in [6] for "properly-formed" strategies, we here constrain strategies open to a group A at moment w to a set $Strategy^w_A$, such that every strategy in it is uniform. For notational convenience we say that two moments w_1 and w_2 are *indistinguishable*, if and only if every choice in w_1 has an indistinguishable choice in w_2, formally stated by

$$w_1 \in R_{K_A}(w_2) \text{ iff } \forall h \in H_{w_1}, \exists h' \in H_{w_2}, w_1/h \in R_{UC_A}(w_2/h').$$

R_{K_A} is a relation of equivalence among moments.

$$\begin{array}{rl} Strategy^w_A = \{\sigma \mid & Dom(\sigma) \subseteq \bigcup_{w' \in R_{K_A}(w)} Tree_{w'}, \\ & \text{and } \forall w' \in Dom(\sigma), \forall w'' \in R_{K_A}(w'), \\ & (w'' \in Dom(\sigma) \text{ and } \exists h' \in \sigma(w'), \\ & \exists h'' \in \sigma(w''), w''/h'' \in R_{UC_A}(w'/h'))\} \end{array}$$

The set $Strategy^w_A$ contains strategies defined on moments lying in the future of the indistinguishable moments of w (included). Those strategies, if defined at a moment w' must also be defined at each moment w'' indistinguishable from w', where the strategy must apply the same choice type. The semantics of the operator for *S*trategic *A*bility with imperfect knowledge is then:

$$\begin{array}{rcl} \mathcal{M}, w/h \models SAstit_A\varphi & \iff & \exists \sigma_A \in Strategy^w_A, \\ & & \forall h' \in Adh(\sigma_A), \\ & & \forall w' \in R_{K_A}(w) \cap h', \\ & & \mathcal{M}, w'/h' \models \varphi \end{array}$$

A possible reading is : "A has the ability to guarantee the truth of φ by carrying out an available strategy, and knows it". $SAstit_A\varphi$ corresponds to $\Diamond[A\, scstit : \varphi]$ in the original notation. We believe that this notation can be misleading, since the operator is atomic: Horty refers to it as a *fused* operator.

Figure 2 presents a scenario where a *blind* agent a enters a room. He does not know whether the light is on. At moments w_2 and w_3 he has the choice between switching the light (action `toggle`), remaining passive (action `skip`) or touching the light bulb (action `touch`). This last action is *informative*. Let strategy $\sigma = \{\langle w_1, \{h_1 \dots h_{12}\}\rangle, \langle w_2, \{h_3, h_4\}\rangle, \langle w_3, \{h_9, h_{10}\}\rangle, \langle w_5, \{h_4\}\rangle, \langle w_8, \{h_9\}\rangle\}$.[6] We can ckeck that σ is indeed in $Strategy^{w_1}_{\{a\}}$, i.e., that it is a uniform strategy for a at the moment w_1, and $Adh(\sigma) = \{h_4, h_9\}$. l holds

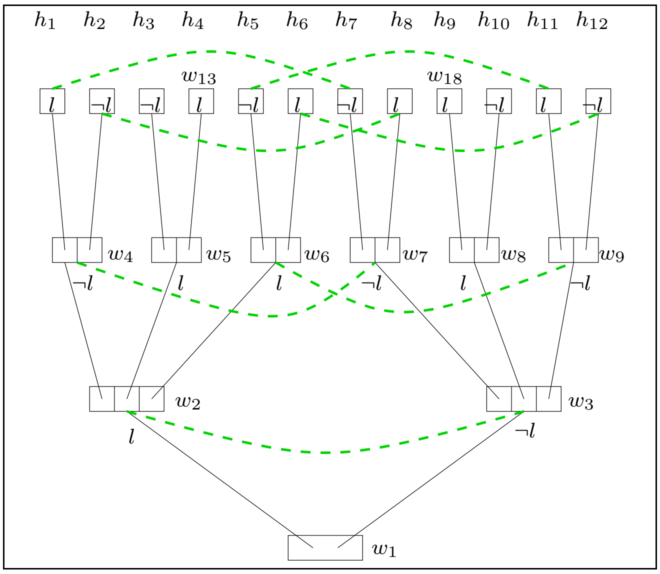

Figure 2: Suppose a blind agent a. For clarity, dotted lines are intended for R_{K_a} relations over moments, and R_{UC_a} relations can be deduced since choices of indistinguishable moments are the same: at w_1 $\{h_1 \dots h_{12}\}$ is for `enter`, at w_2 $\{h_1, h_2\}$ is for `toggle`, $\{h_3, h_4\}$ for `touch` and $\{h_5, h_6\}$ for `skip` (and identically for choices at w_3), at w_4 $\{h_1\}$ is for `toggle` and $\{h_2\}$ for `skip` (the same for moments $w_2 \dots w_9$). l stands for "the light is on", and time goes upward.

at both w_{13}/h_4 and w_{18}/h_9. Thus $\mathbf{F}l$ is true at w_1/h_4 and w_1/h_9 and it follows that considering σ, agent a knows how to play to ensure that the light is on in the future: $SAstit_a\mathbf{F}l$.

We can thus easily grasp the notion of uniform strategies in the STIT framework. This is due to the fact that the underlying semantics is more versatile than that of ATEL, because it allows to relate a moment/history pair w_1/h_1 to another w_2/h_2 via R_{UC_a} without relating all w_1/h'_1 to w_2/h'_2. We here have changed the domain of quantification over strategies such that we just consider strategies satisfying the definition of uniformity. In the following, and for this preliminary work, we decide to simplify the framework into a STIT-like one, for *one-step* strategies.

5. TOWARDS A SIMPLIFICATION

The STIT branching time is based on the so called T × W logic for *historical necessity* (see [1]). Such a logic is a combination of tense and modal logic for worlds with the same time order. T×W semantics can be seen as a grid with tense on one dimension and set of histories on the other. In Figure 3 we have a grid with temporal aspects represented longitudinally, time going upward.

The aim of this section is to simplify the STIT-framework with imperfect knowledge into a semantics that does not involve histories. We focus on *one-step* strategies where the problem of uniform strategies already appears.

5.1 A logic of knowledge-based uniform agency

We here deal with a framework permitting us to talk about agency, and more precisely, about agency with respect to

[5] It is easy to see that even histories are superfluous. We nevertheless keep them for uniformity.

[6] The notation is abusive: σ should be defined on the complete set $\{w_1 \dots w_{12}\}$, but it does not influence the set of admitted histories.

knowledge about actions. Thus, we do not deal with static knowledge in Hintikka's tradition: we are not interested here in sentences like "Mary knows that the light is on", but in sentences such as "Mary knows her current action makes sure that the light will be on". Nevertheless, we will briefly see in Section 7 how we can model the first case in that framework, too.

We propose a logic for *knowledge-based uniform agency* in which the knowledge scope stops at effects of actions. We keep on using the terminology of STIT. Thus, historically possible *indexes* lie in a same *moment*.

5.2 Syntax and Semantics

Given a finite set of agents $\mathcal{A}gt$ and a set of atomic propositions $\mathcal{A}tm$, a formula φ can have the following syntactic form:

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathbf{X}\varphi \mid \Box\varphi \mid Kstit_A\varphi \mid Stit_A\varphi$$

with $p \in \mathcal{A}tm$ and $A \subseteq \mathcal{A}gt$.

A model is a tuple $\mathcal{M} = \langle W, R_X, R_\Box, R_{UC}, v\rangle$, where:

- W is a nonempty set of indexes;
- $R_X : W \to W$ is total function associating an index to its historical successor;
- $R_\Box$ is a relation between indexes;
- $R_{UC} : \mathcal{A}gt \to (W \to 2^W)$ is a function associating each agent $a \in \mathcal{A}gt$ to a relation. For every agent $a \in \mathcal{A}gt$ the relation $R_{UC}(a)$ will be noted R_{UC_a};
- $v : \mathcal{A}tm \to 2^W$ is an evaluation function.

$R_\Box(w)$ denotes the set of indexes lying in the same moment as w. R_{UC_a} is the relation between indexes of identical choices lying in moments indistinguishable by agent a.

For a group of agents A, we suppose that agents of A make choices collectively with distributed knowledge: they share their knowledge. We thus define the collective uniform choice relation as follows:

$$R_{UC_A}(w) = \bigcap_{a \in A} R_{UC_a}(w).$$

We make the same hypothesis as in STIT concerning the independence/consistency of choices of agents, and assume that for each index w, $R_{UC_{\mathcal{A}gt}}(w) \neq \emptyset$.

We assume models satisfy additional properties:

1. $(R_\Box \circ R_X)(w) \supseteq (R_X \circ R_\Box)(w)$
   If it is historically necessary that some state of affairs will hold next, then next, this state of affairs will be historically necessary.
2. $(R_\Box \circ R_{UC_A})(w) = (R_{UC_A} \circ R_\Box)(w)$
   The same types of choices are available at every indistinguishable moment.
3. $(R_{UC_A} \circ R_X)(w) \supseteq (R_X \circ R_\Box \circ R_{UC_A})(w)$
   If a coalition A makes a uniform choice to ensure a state of affairs at the next step, it will be historically inevitable at the next step that A knows that it ensures φ.

The first assumption forces a tree structure of moments. The second one is the translation of Hypothesis 3. The last assumption is motivated by the fact that it is not sufficient to say that a group of agents A knows how to play to ensure φ at moment w if A does not know when it has achieved φ. Nevertheless, agents know possible effects of their actions. We then adopt the principle that if agents know they ensure φ at the next step by doing the current choice, then at the very next step they will know that φ holds. It forces a *no forget* principle which seems realistic since we deal with *one-step* strategies.

A formula is evaluated with respect to a model and a world, which corresponds to STIT indexes and thus contains information about history.

$$\begin{array}{lcl}
\mathcal{M}, w \models p & \iff & w \in v(p),\ p \in \mathcal{A}tm \\
\mathcal{M}, w \models \neg\varphi & \iff & \mathcal{M}, w \not\models \varphi \\
\mathcal{M}, w \models \mathbf{X}\varphi & \iff & \mathcal{M}, R_X(w) \models \varphi \\
\mathcal{M}, w \models \Box\varphi & \iff & \forall w' \in R_\Box(w), \mathcal{M}, w' \models \varphi \\
\mathcal{M}, w \models Kstit_A\varphi & \iff & \forall w' \in R_{UC_A}(w), \mathcal{M}, w' \models \varphi \\
\mathcal{M}, w \models Stit_A\varphi & \iff & \forall w' \in (R_{UC_A} \cap R_\Box)(w), \\
 & & \mathcal{M}, w' \models \varphi
\end{array}$$

$w \models \mathbf{X}\varphi$ means that φ is true at the immediate successor of w on the history. $w \models \Box\varphi$ means that φ is true at each world lying in the same moment as w. $Kstit_a\varphi$ is an agency operator with an implicit epistemic feature. It means that a knows that if he performs his chosen action, he ensures that φ. $Stit_A\varphi$ still stands for "A's current choice makes sure that φ".

5.3 Some validities

$\mathbf{X}$ is a KD-modality, and $\Box$ and every $Stit_A$ and $Kstit_A$ are $S5$-modalities.

$(MonKstit)$ $Kstit_{A_1}\varphi \to Kstit_{A_2}\varphi$ if $A_1 \subseteq A_2$

$(MonStit)$ $Stit_{A_1}\varphi \to Stit_{A_2}\varphi$ if $A_1 \subseteq A_2$

$(KnowAct)$ $\Box Kstit_A\varphi \leftrightarrow Kstit_A\Box\varphi$

$(Linearity)$ $\neg\mathbf{X}\neg\varphi \to \mathbf{X}\varphi$

$(XBox)$ $\Box\mathbf{X}\varphi \to \mathbf{X}\Box\varphi$

$(NoForget)$ $Kstit_A\mathbf{X}\varphi \to \mathbf{X}\Box Kstit_A\varphi$

$(StitBox)$ $\Box\varphi \to Stit_A\varphi$

$(StitKstit)$ $Kstit_A\varphi \to Stit_A\varphi$

$(MonKstit)$ and $(MonStit)$ are valid by the construction of the collective uniform choice relation R_{UC_A}. $(KnowAct)$ is the immediate translation of the second constraint. $(Linearity)$ follows from the definition of the definition of the R_X function. $(XBox)$ is the translation of the first constraint. $(NoForget)$ is the direct translation of the third constraint.[7] $(StitBox)$ and $(StitKstit)$ are valid by the semantics of the $Stit$ operator.

[7]The link with the *no forget* principle is maybe not clear, but we say more about this in section 7.

6. KNOWING HOW TO PLAY

The fact that a group of agents A knows how to play to ensure φ at the next step can be captured by the formula

$$\Diamond Kstit_A \mathbf{X}\varphi.$$

A group of agents knows how to play to ensure φ at the next step if "A can make a choice such that it knows it ensures φ at the next step".

In order to illustrate the logic of Section 5, and particularly the capabilities of groups with incomplete knowledge, we present two scenarii constructed above examples 1 and 3. First we deal with one agent with lack of knowledge (the blind agent a_1). Then, we add another agent that cannot act (the lame agent a_2), and show how they can put their strengths together.

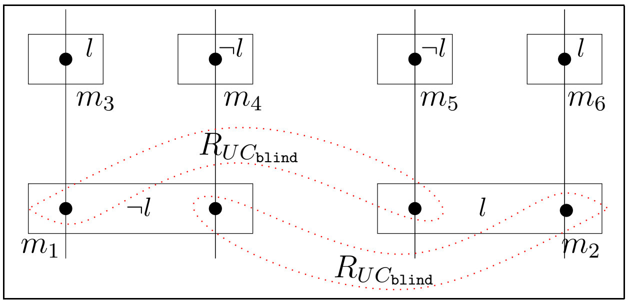

Figure 3: (Schematic of Example 1) Vertical lines represent histories, rectangles are historical necessity relations and delimit moments.

In Figure 3 the moments m_1 and m_2 are epistemically equivalent w.r.t. static knowledge. Suppose a blind person enters a room at moment m_0. The problem is to know whether he can ensure that the light is on. For that he has two possible actions, `toggle` and `skip`. If at m_1 or m_2 he can act such that the light is on, he cannot be aware of that: he has no uniform choice such that the result is always l. Let us call $w_{1,r}$ the right index of moment m_1, and $w_{1,l}$ the left one. The following holds:

- $w_{1,r} \models Stit_{\text{blind}}\mathbf{X}l \wedge \neg Kstit_{\text{blind}}\mathbf{X}l$
- $w_{1,l} \models \neg\Box\neg Stit_{\text{blind}}\mathbf{X}l \wedge \Box\neg Kstit_{\text{blind}}\mathbf{X}l.$

Figure 4 presents an example where the blind person from the previous example is helped by a lame person. While the blind does not know the current state of the light, the lame does. While the lame cannot act such that the light is on at the next step, the blind can. If we call $w_{1,l}$ the left index of the moment m_1 and $w_{1,r}$ the right one, the following holds:

- $w_{1,l} \models Stit_{\text{blind}}\mathbf{X}l \wedge \neg Kstit_{\text{blind}}\mathbf{X}l$
- $w_{1,r} \models \neg\Box\neg Stit_{\text{blind}}\mathbf{X}l \wedge \Box\neg Kstit_{\text{blind}}\mathbf{X}l$
- $w_{1,l} \models \neg Stit_{\text{lame}}\mathbf{X}l$
- $w_{1,r} \models \Box\neg Stit_{\text{lame}}\mathbf{X}l$

but,

- $w_{1,l} \models Kstit_{\{\text{blind,lame}\}}\mathbf{X}l$

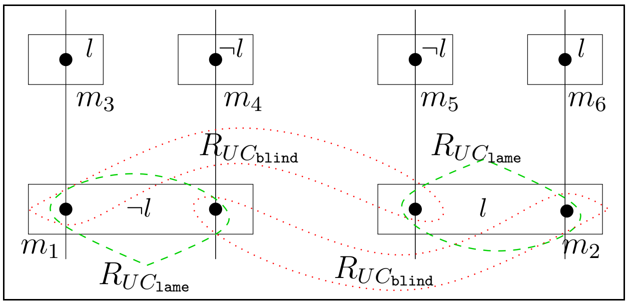

Figure 4: (Schematic of Example 3) The actual moment is m_1, where the light is off ($\neg l$). The lame agent cannot act but has complete knowledge. The blind agent does not know if the light is on or off, but can choose to toggle or not (`skip`). Individually, neither the lame agent nor the blind agent can ensure that the light is on, but together, they can form an effective coalition for that purpose.

- $w_{1,r} \models \neg\Box\neg Kstit_{\{\text{blind,lame}\}}\mathbf{X}l$

Now, we see that this last formula corresponds to the concept of *knowing how to play* that we wanted to grasp.

7. STATIC KNOWLEDGE

We have not included static knowledge in the language of our logic of knowledge-based uniform agency. However, it can be defined: we say that an agent a knows that φ if and only if whatever a chooses, it knows that it's choice makes sure that φ. We say that a group of agents A knows that φ if and only if whatever agents of A choose, they distributively know that their choices make sure that φ. We thus can define in our logic, a modal operator reading "group A knows that φ", as:

$$K_A\varphi =_{def} \Box Kstit_A\varphi.$$

"A knows that φ" is equivalent to "it is historically inevitable that A knows that it ensures φ". We also can reformulate the $(NoForget)$ validity by $Kstit_A\mathbf{X}\varphi \rightarrow \mathbf{X}K_A\varphi$, which makes sense now, and can be read as "if A knows it ensures that φ is true at the next step, then at the next step the group will know that φ holds". The following formulas are valid:

- $K_A\varphi \rightarrow \Box\varphi$
- $K_A Kstit_A\varphi \leftrightarrow K_A\varphi$
- $Kstit_A\neg\Box\neg\varphi \rightarrow K_A\neg\Box\neg\varphi$

Finally, the formulation of "a group of agents A know how to play to ensure that φ at the next step" is equivalent to

$$K_A\Diamond Kstit_A\mathbf{X}\varphi.$$

K_A focuses on what agents know at a moment, before doing a choice, or whatever the choice they do. $Kstit_A$ is about knowledge after having made a choice. The threshold from static knowledge to dynamic knowledge is monotonic: what is known before doing a choice is also known after ($K_A\varphi \rightarrow Kstit_A\varphi$ is valid). Thus $Kstit_A$ is a refinement of Hintikka's static knowledge.

This distinction is of interest even with perfect knowledge about the current world. Indeed, if Mary knows that the light is on, she is the only one to have the control on it, and that the switch is functioning well, she nevertheless ignores the state of the light at the next moment, until she makes her choice between toggling and remaining passive. But when she is committed to a choice, she knows the next state of the light, and this already before performing the action.

8. FUTURE WORK

A further extension is a strategic operator for global rather than local effectivity. Two paths open before us: (1) We can define an iterated operator $(\Diamond Kstit_A \mathbf{X})^* \varphi$, which we believe can be done in a rather simple manner, following [12, Chapter 4] and capitalizing on the *no forget* principle. However, this remains an idealization about knowledge of agents. (2) We can adapt strategies in our logic as we have done in Section 3.

Also, our aim was here to deal with cooperative coalitions of agents. We hence used distributed knowledge. It could be interesting to investigate refinements of group knowledge.

9. ACKNOWLEDGMENTS

This reasearch is part of the MOSTRO project, funded by the "Fondo provinciale per i progetti di ricerca" of the Provincia Autonoma of Trento.

We would like to thank Wojtek Jamroga and two anonymous reviewers for fruitful comments.

10. REFERENCES

[1] L. Åqvist. The logic of historical necessity as founded on two-dimensional modal tense logic. *Journal of Philosophical Logic*, 28(4):329–369, 1999.

[2] N. Belnap, M. Perloff, and M. Xu. *Facing the future: agents and choices in our indeterminist world*. Oxford, 2001.

[3] J. Broersen, A. Herzig, and N. Troquard. From Coalition Logic to STIT. In W. van der Hoek, A. Lomuscio, E. de Vink, and M. Wooldridge, editors, *Proceedings of the 3rd Int. Workshop on Logic and Communication in Multiagent Systems (LCMAS 2005)*, 2005. To appear in ENTCS.

[4] J. Broersen, A. Herzig, and N. Troquard. Embedding Alternating-time Temporal Logic in Strategic STIT Logic of Agency, 2006. Submitted.

[5] V. Goranko and W. Jamroga. Comparing semantics of logics for multi-agent systems. *Synthese*, 139(2):241–280, 2004.

[6] J. F. Horty. *Agency and Deontic Logic*. Oxford University Press, Oxford, 2001.

[7] J. F. Horty and N. D. Belnap, Jr. The deliberative STIT: A study of action, omission, and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.

[8] W. Jamroga. Some remarks on alternating temporal epistemic logic. In *Proc. of the Int. Workshop on Formal Approaches to Multi-Agent Systems (FAMAS'03)*, 2003.

[9] W. Jamroga and W. van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 2004.

[10] W. Jamroga and T. Ågotnes. What agents can achieve under incomplete information. In *Proceedings of AAMAS'06*, 2006.

[11] E. Pacuit, R. Parikh, and E. Cogan. The logic of knowledge based obligation, 2005. To appear in Knowledge, Rationality and Action. Early versions were presented at DALT 2004 and SEP 2004.

[12] M. Pauly. *Logic for Social Software*. PhD thesis, University of Amsterdam, 2001. ILLC Dissertation Series 2001-10.

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

[14] W. van der Hoek and M. Wooldridge. Tractable multiagent planning for epistemic goals. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 1167–1174, New York, NY, USA, 2002. ACM Press.

[15] S. van Otterloo and G. Jonker. On epistemic temporal strategic logic. *Electronic Notes in Theoretical Computer Science*, 85(2), 2004. In Proceedings of the 2nd Int. Workshop on Logic and Communication in Multiagent Systems (LCMAS 2004).