Performing conditional strategies in strategic STIT theory

Hein Duijf*

Department of Philosophy and Religious Studies, Universiteit Utrecht

Abstract. We introduce a formalization of conditional strategies in strategic STIT theory. This will turn out to have unexpected consequences, in particular it turns out that performing a strategy conditional on c is equivalent to performing that strategy conditional on a logically weaker condition. Hence it will turn out that performance of a strategy conditional on c can already commit you to performing that strategy if c is not the case. We will argue in favour of our formalization, this result and some further consequences. Our investigation points to a misunderstanding that the conditions in the conditional strategies are *moment-determinate*.

1 Introduction

In this paper we present a formal investigation on conditional strategies. These are strategies that include some conditional actions. We often describe the strategies we are performing in this way, for instance 'if it is hot outside, I eat lunch in the park' or 'I will buy you a present if you behave nicely'. These sentences describe my current strategy by mentioning how certain actions are triggered, this is at least a partial description of my current strategy. Take the example of 'if it is hot outside, I eat lunch in the park'. I may even utter this sentence on a snow day,[1] which implies that this action is not triggered, i.e. I will not eat lunch in the park today.

Furthermore, we want to examine these conditional strategies in a multi-agent setting. The main advantage of this is that it enables us to make a start in the study of interaction and collective action.[2] For instance in a coordination game I have to coordinate my actions with yours, that is to say that I want to perform a strategy conditional on your current strategy. This makes the current exposition applicable to these central concepts in interaction and collective agency.

* I gratefully acknowledge financial support from the ERC-2013-CoG project REINS, nr. 616512.

[1] As I was told, a snow day in Canada means a day off from school because of arctic weather conditions.

[2] Contrast collective action with mere cumulative action, which is represented in STIT by a coalition A in the agency operator $[A\ \texttt{sstit}]$. The formal study of collective action is still in its infancy.

There has been a lot of work on formalizations of strategies and strategic reasoning. The seminal work on ATL[1] adds strategies to temporal logic. However, the syntax is confined since the main operator only considers the existence of strategies to ensure certain properties. This work is extended in [14] to *Counterfactual* ATL (CATL). The essential difference is a syntactical extension of the language with an operator $C_i(\sigma, \varphi)$ with the intended meaning of 'suppose agent i chooses the strategy denoted by σ; then φ holds'. In [5] and [8] a first-order extension of temporal logic (both called *Strategy Logic*) is proposed in which one can quantify over the strategies. In [5] they introduce seperate quantifiers for each agent, while in [8] the quantification is over general strategies but an operator $(a, x)\varphi$ is added to say that 'bind agent a to the strategy associated by x'. All of these frameworks focus on the *existence* of strategies and lack talk about the *structure* of the strategies involved. In this paper we aim to provide a formalization of performing conditional strategies, so we have to be able to point out structural properties of strategies.

In [15] the semantics are similar to that of ATL, except strategies are defined as "rough plans" in contrast to the complete, deterministic strategies that are considered in ATL. So on the semantic side they propose a more liberal view on strategies. One of the main themes they discuss is the representation of strategies by formulas of the language. In the current paper we aim to represent the performance of conditional strategies by formulas of some language. Since their language is very elementary it is not rich enough for our purposes.

It is not surprising that Dynamic Logics have been proposed for modelling strategies and the structure of the strategies (see [12] and [13]). However, these logics seem only appropriate for finite extensive form games and it is not clear how the methods generalize to infinite games.

The logic proposed in [11] evaluates formulas at game-strategy pairs thereby combining aspects of *Game Logic* (see the original work [9] and the overview [10]) and strategic reasoning. A multi-sorted language is proposed to emphasize the *structure* of the strategy. The logic includes two 'types' of conditional strategies $[\psi \mapsto a]^i$ and $\pi \mapsto \sigma$, which are interpreted as 'player i chooses move a whenever p holds' and 'player i sticks to the specification given by σ if on the history of play, all moves made by $\bar{i}$ conform to π' respectively. The ψ in the first formula is of a special nature, namely it is a boolean combination of propositional letters, and the second formula represents that the other players *have acted* conform π. We, however, do not want to commit ourselves to these restrictions on conditions. For instance, it is not obvious whether a sentence like 'I will buy you a present if you behave nicely' can be represented in that logic. Indeed, it could be that at the very moment I buy you a present you misbehave. The proposed logic does not allow my current actions to be conditional on your current actions.

In order to adequately formalize conditional strategies we will use the framework G.STRAT (first introduced in [3]), which is an extension of basic STIT frameworks to a strategic and multi-agent setting. As is usual in STIT frameworks, we are able to talk about the *structure* of a strategy by explicating which (temporal) properties it ensures. In addition, this formalism allows us to talk about the dynamics of the world, in particular it allows us to say that an agent *is performing* an action. In consequence, we can use this to say that when a certain condition holds the agent *performs* a certain action, which means that the action is triggered by the condition.

The paper is organized as follows. In Section 2 we introduce the logical framework G.STRAT. Subsequently, in Section 3 we will present our formalization of conditional strategies, treat several examples to challenge our intuitions, derive some formal results and argue in favour of our formalization. The proofs of the formal results can be found in the Appendix A. We will conclude in Section 4 with some discussion, a summary of the results and key contributions of this paper. If the reader does not want to delve deep into formalisms, it is advisable to start from Section 3 and only look up formal aspects in Section 2 when needed.

2 Strategic stit theory: G.STRAT

Below we present the formal syntax and semantics of G.STRAT, a logic that was first presented in [3] (though our presentation is slightly different) and extends the classical STIT framework (cf. [2] and [7]) to strategic actions. The G in the acronym G.STRAT stands for 'Group' and STRAT stands for 'strategic'.

Definition 1 (Syntax)**.** Well-formed formulas of the language $\mathcal{L}_{\mathsf{G.STRAT}}$ are defined by:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [A\ \texttt{sstit}]\varphi \mid \Box\varphi \mid \texttt{X}\varphi \mid \texttt{G}\varphi$$

The p are elements from a countable infinite set of propositional symbols $\mathcal{P}$, and A is a subset of a finite set of agents Ags. We will often use i to refer to an arbitrary agent in this set. We use $\varphi, \psi, \ldots$ to represent arbitrary well-formed formulas. We use the standard propositional abbreviations, the standard notation for the duals of modal boxes (i.e., diamonds).

We now go on to define the semantic structures for G.STRAT. Here we take the viewpoint that strategies *are* sets of histories obeying certain structural properties. This is nothing more than a convenient shift of viewpoint and does not result in a fundamentally new type of strategies. If we define strategies as mappings from states / moments to choices, we have to define 'compliance' of a history to a strategy as a secondary concept. By defining strategies as sets of histories with a certain structure there is no need anymore to define a notion of 'compliance'.

We introduce a semantics for this language using standard techniques in modal logic. The units of evaluation are 'profiles'. A profile records the dynamic aspects of a system of agents. The profiles of our semantics take a moment, a history and a strategy profile (a list of strategies, one for each agent in the system)[3] as components. So, the formulas of G.STRAT are evaluated against

[3] Beware that we have 'profiles' and 'strategy profiles' in our terminology. A strategy profile is a part of a profile.

tuples $\langle m, h, \alpha_1, \alpha_2, \dots \alpha_n\rangle$, where m is a moment, h a history, and $\alpha_1, \alpha_2, \dots \alpha_n$ a strategy profile. Then, the truth of formulas is evaluated against the background of a *current* moment, a *current* history, and a *current* strategy-profile.

If, under this semantics, we want to consider truths that do *not* depend on dynamic aspects as represented by the histories and the strategies, we can use the historical necessity operator $\Box$. In particular, if $\Box\varphi$ holds, φ can be said to hold 'statically'. In *stit*-theory, one says that φ is 'moment determinate'. We also say that φ is 'settled'[4], which refers to the fact that it is completely independent of any action currently taken by any agent in the system.

Definition 2 (Strategic frames)**.** A frame is a tuple $\mathcal{F} = \langle M, H, \{Strat(i) \mid i \in Ags\}, \{R_A \mid A \subseteq Ags\}\rangle$ such that:

1. M is a non-empty set of moments, which are denoted by m, m', ...
2. H is a non-empty set of 'backwards bundled' histories. A history $h \in H$ is a sequence $\dots m, m', m'' \dots$ of mutually different elements from M. For m appearing strictly before m' on the history h we write $m <_h m'$. To denote that m' succeeds m on h we use a successor function *Succ* and write $m' = Succ(m, h)$. A history $h \in H$ and its successor function *Succ* are isomorphic to $(\mathbb{Z}, +1)$. Furthermore, let $H_n = \{h \mid h \in H \text{ and } n \text{ on } h\}$. The following constraint on the set H ensures a deterministic past:
   a. for all $h \in H$, if $m = succ(n, h)$ then $H_m \subseteq H_n$
3. $Strat(i)$ yields for each $i \in Ags$ a non-empty set of strategies. Strategies are non-empty sets of histories. A strategy profile[5] is a list of strategies $\alpha_1, \alpha_2, \dots \alpha_n$, where $\{1, 2, \dots, n\} = Ags$ and $\alpha_i \in Strat(i)$ for any i. For strategy profiles we will use the vector notation $\vec{\alpha}$ when we need to be more concise. We introduce the notion of *strategies for agent i at m*, denoted by $Strat_m(i)$, as the set of strategies that are compliant with m, formally $\{\alpha_i \in Strat(i) \mid \alpha \cap H_m \neq \emptyset\}$.[6] Strategies obey the constraints a., b., c., d. and e. below.
   a. for all $m \in M$ and every agent i, there is a strategy at m
   b. for all $m \in M$ and for any $i \in Ags$, if α_i is a strategy for agent i at m, and if for a history $h' \in H$ through m it holds that for all x on h' there is a $h'' \in \alpha_i$ such that x is on h'', then $h' \in \alpha_i$

[4] Settledness does *not* necessarily mean that a property is *always* true in the future (as often thought). Settledness may also apply to the condition that φ occurs 'some' time in the future, now, or indeed any condition expressible as linear time temporal formula. So, settledness is a universal quantification over the *branching* dimension of time, and *not* over the dimension of duration.

[5] In the game forms of game theory strategy profiles are referred to by means of names associated with the choices of agents in states. Here we abstract from names of choices.

[6] It makes sense to talk about strategies at a moment. Since we represent strategies by the set of histories that are compliant with that strategy, a strategy will most likely not be compliant with every moment. Note that this notion is redundant when we view strategies as functions from moments to choices, like is usually done in for instance ATL (see [1]), since in that case we would have $Strat_m(i) = Strat(i)$.

c. for all $m \in M$ and for any $i \in Ags$, if α_i is a strategy for agent i at m, and there is another strategy β_i for agent i at $n = Succ(m, h)$, then there is a strategy γ_i for agent i at m such that $\gamma_i = (\alpha_i \setminus H_n) \cup (\beta_i \cap H_n)$

d. for all $m \in M$ for all strategy profiles at m, i.e. lists of strategies $\alpha_1, \ldots, \alpha_n$ at m, we have that $\bigcap_{i \in Ags} \alpha_i$ contains at least one history through m

e. for all $m \in M$ for all strategy profiles at m, i.e. lists of strategies $\alpha_1, \ldots, \alpha_n$ at m, we have that $\bigcap_{i \in Ags} \alpha_i$ contains at most one history through m

4. Profiles are exactly the tuples $\langle m, h, \alpha_1, \alpha_2, \ldots \alpha_n \rangle$ such that m belongs to h and for all $i \in Ags,\ h \in \alpha_i$. These will be the units of evaluation (possible worlds) of our modal semantics.
5. The relations R_A are 'effectivity' equivalence classes over profiles such that $\langle m, h, \alpha_1, \alpha_2, \ldots \alpha_n \rangle R_A \langle m', h', \beta_1, \beta_2, \ldots \beta_n \rangle$ if and only if $m = m'$, and for all $i \in A, \alpha_i = \beta_i$

Item **1** gives the basic elements of the frames: the set of moments M.

Item **2** defines histories to be linearly ordered sets of moments. It also defines a bundling structure for the set H of all histories: histories that come together when going in the past direction will stay together in that direction. This implies that in the future direction, once bundles have separated they will never come together again.

Item **3** defines strategies. Strategies are sets of histories with certain properties. A strategy profile is a choice of strategy for each agent in the system. Condition **3(a)** ensures that for each moment, for each agent, its strategy at that moment is defined by some bundle of histories. Condition **3(b)** ensures that our bundled histories semantics behaves as a tree semantics on aspects relevant for the comparison with ATL (see [1]). It says that if an agent indefinitely complies to a strategy then he is performing that strategy. In particular, this property ensures that G.STRAT obeys ATL's induction axiom. Condition **3(c)** ensures that if at a certain moment an agent has two different strategies in his repertoire, recombinations of these strategies where first one strategy is followed and later on (next) the other, are always also in the repertoire of that agent at that moment. Condition **3(d)** ensures that we can always recombine strategies of individual agents contributing to a group strategy into a new joint strategy. This implements the *stit*-requirement of *independence of agency* (no agent can choose in such a way that some other agent is deprived of one of its choices). Condition **3(e)** ensures that there is *exactly one* history complying to all strategies of a strategy profile at a moment. This reflects the idea that a choice of strategy for each agent in the system completely determines the entire future.

Item **4** introduces profiles, which are the units of evaluation for our modal semantics. Profiles are tuples consisting of a moment, a history and a strategy for each agent.

Item **5** defines R_A to be a relation reaching all profiles that only deviate from the current profile in the sense that agents not among A perform a choice different from the current one. This reflects the basic idea of agency saying that acting or choosing is ensuring a condition irrespective of what other agents do or choose.

Now we are ready to define the formal semantics of the language $\mathcal{L}_{\mathsf{G.STRAT}}$. The semantics is multi-dimensional, and the truth conditions are quite standard. First we define models based on the frames of the previous definition.

Definition 3 (Strategic models)**.** A frame $\mathcal{F} = \langle M, H, \{Strat(i) \mid i \in Ags\}, \{R_A \mid A \subseteq Ags\}\rangle$ is extended to a model by adding a valuation π of atomic propositions:

– π is a valuation function assigning to each atomic proposition the set of profiles (see Definition 2) in which they are true.

Definition 4 (Truth, validity, logic)**.** Truth $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \models \varphi$, of a G.STRAT-formula φ in a profile $\langle m, h, \vec{\alpha}\rangle$ of a model $\mathcal{M} = \langle M, H, \{Strat(i) \mid i \in Ags\}, \{R_A \mid A \subseteq Ags\}, \pi\rangle$ is defined as (suppressing the model denotation '$\mathcal{M}$'):

$$
\begin{array}{ll}
\langle m, h, \vec{\alpha}\rangle \models p & \Leftrightarrow \langle m, h, \vec{\alpha}\rangle \in \pi(p) \\
\langle m, h, \vec{\alpha}\rangle \models \neg\varphi & \Leftrightarrow \text{not } \langle m, h, \vec{\alpha}\rangle \models \varphi \\
\langle m, h, \vec{\alpha}\rangle \models \varphi \wedge \psi & \Leftrightarrow \langle m, h, \vec{\alpha}\rangle \models \varphi \text{ and } \langle m, h, \vec{\alpha}\rangle \models \psi \\
\langle m, h, \vec{\alpha}\rangle \models \Box\varphi & \Leftrightarrow \text{for all profiles } \langle m, h', \vec{\beta}\rangle \text{ at } m \\
 & \quad\ \text{we have } \langle m, h', \vec{\beta}\rangle \models \varphi \\
\langle m, h, \vec{\alpha}\rangle \models \mathtt{X}\varphi & \Leftrightarrow \text{for } m' = Succ(m, h) \\
 & \quad\ \text{it holds that } \langle m', h, \vec{\alpha}\rangle \models \varphi \\
\langle m, h, \vec{\alpha}\rangle \models \mathtt{G}\varphi & \Leftrightarrow \text{for all } m' \text{ such that } m \leq_h m' \\
 & \quad\ \text{it holds that } \langle m', h, \vec{\alpha}\rangle \models \varphi \\
\langle m, h, \vec{\alpha}\rangle \models [A\ \mathtt{sstit}]\varphi & \Leftrightarrow \text{for all } h', \vec{\beta} \text{ such that} \\
 & \quad\ \langle m, h, \vec{\alpha}\rangle R_A \langle m, h', \vec{\beta}\rangle \\
 & \quad\ \text{it holds that } \langle m, h', \vec{\beta}\rangle \models \varphi
\end{array}
$$

Validity on a G.STRAT-model $\mathcal{M}$ is defined as truth in all profiles of the G.STRAT-model. General validity of a formula φ is defined as validity on all possible G.STRAT-models. The logic G.STRAT is the subset of all general validities of $\mathcal{L}_{\mathsf{G.STRAT}}$ over the class of G.STRAT-models.[7]

3 Conditional strategies

In order to adequately formalize conditional strategies we will use the G.STRAT framework as presented in Section 2. We briefly present key elements and aspects

[7] Little is known about the validities of this logic, for instance there is no axiomatization nor completeness result for this logic. To guide some of the formal intuitions of the reader we mention some validities without proof or conceptual motivation: $\Box$ and $[i\ \mathtt{sstit}]$ are S5-modalities, the $[i\ \mathtt{sstit}]$-operator is monotone in its agency argument, i.e. for $A \subseteq B$ we have $\models [A\ \mathtt{sstit}]p \rightarrow [B\ \mathtt{sstit}]p$, the temporal part is the standard discrete linear temporal logic containing X and G, some interaction principles are $\models \Box\mathtt{X}p \rightarrow \mathtt{X}\Box p$, $\models [i\ \mathtt{sstit}]\mathtt{X}p \rightarrow \mathtt{X}[i\ \mathtt{sstit}]p$, $\models [i\ \mathtt{sstit}]\Box p \leftrightarrow \Box p$, $\models [i\ \mathtt{sstit}]\mathtt{X}[i\ \mathtt{sstit}]p \leftrightarrow [i\ \mathtt{sstit}]\mathtt{X}p$, $\models [i\ \mathtt{sstit}]\mathtt{G}[i\ \mathtt{sstit}]p \leftrightarrow [i\ \mathtt{sstit}]\mathtt{G}p$ (see Lemma 10 in Appendix A), and $\Diamond[A\ \mathtt{sstit}]p \wedge \Diamond[B\ \mathtt{sstit}]q \rightarrow \Diamond[A \cup B\ \mathtt{sstit}](p \wedge q)$ for disjoint coalitions A and B (independence of agency).

of G.STRAT that are needed in order to understand our formalism. G.STRAT is a strategic STIT logic that evaluates formulas against tuples $\langle m, h, \vec{\alpha}\rangle$ that are referred to as *profiles*. The elements of these profiles are a moment m, a history h, and a strategy profile $\vec{\alpha}$ specifying a strategy α_i for each agent i in the language. The central agency operator is the modality $[i\ \texttt{sstit}]\varphi$ which stands for 'agent i strategically sees to it that φ'. Relative to a profile $\langle m, h, \vec{\alpha}\rangle$ the modality $[i\ \texttt{sstit}]\varphi$ is interpreted as 'agent i is in the process of executing α_i thereby ensuring the (temporal) condition φ'. In addition, the logic includes temporal modalities $\texttt{X}\varphi$ and $\texttt{G}\varphi$ which are interpreted, relative to a profile $\langle m, h, \vec{\alpha}\rangle$, as '$\varphi$ holds in the next moment after m on the current history h' and 'φ holds in every moment after m on the current history h' respectively. Finally, the logic includes a historical necessity operator $\Box\varphi$ which is interpreted, relative to a profile $\langle m, h, \vec{\alpha}\rangle$, as '$\varphi$ holds on any profile at m' or more informally 'φ is settled at m' which says that the truth does *not* depend on dynamic aspects as represented by the histories and the strategies. We now return to the main story concerning conditional strategies and our formalization thereof.

People often make conditional statements about their future choices: 'if it is hot outside, I eat lunch in the park', 'I will buy you a present if you behave nicely'. It is crucial to view these phrases as denoting aspects of the strategy I am currently performing.[8]

Definition 5. (Conditional strategy) We say that relative to a profile $\langle m, h, \vec{\alpha}\rangle$ in a model $\mathcal{M}$ an agent i performs a strategy ensuring φ conditional on c if and only if the following truth condition is satisfied:

$$\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \texttt{sstit}]\texttt{G}(c \to [i\ \texttt{sstit}]\varphi).$$[9]

The above formula means that agent i, along history h, performs strategy α_i that ensures that for all histories in the strategy the conditional is ensured. Informally, this says that agent i is currently performing a strategy that ensures that in case the condition holds he performs a strategy ensuring φ.

The nested $[i\ \texttt{sstit}]$ operator might look confusing, but this makes perfect sense. In order to argue in favour we break the formula down: the conditional strategy is formalized as $c \to [i\ \texttt{sstit}]\varphi$, but one should not forget that here we focus on an agent performing such a strategy. This is formalized by the second $[i\ \texttt{sstit}]$ operator and the $\texttt{G}$ operator.[10]

[8] One could argue that these phrases announce a commitment that I currently have. I believe the current framework can be useful for this interpretation so it might be useful in modeling commitments.

[9] In [4] Broersen proposes to formalize such a conditional strategy by using the formula $[i\ \texttt{sstit}]\texttt{G}(c \to [i\ \texttt{sstit}]\texttt{X}\varphi)$. We think that this is just a minor difference. Since our current research and findings also apply to that proposal, we will not discuss the differences.

[10] Note that it might make sense not to require an agent to ensure that *henceforth* some temporal condition holds, but that the agent ensures this temporal condition *until* a certain deadline. In light of Definition 5 this might be similarly formalized

The possibility to nest strategic action modalities is a crucial feature of G.STRAT. This allows us to express many interesting properties of strategies and consequently of abilities. For instance, the formula $\Diamond[i\ \texttt{sstit}][j\ \texttt{sstit}]\varphi$ says that agent i has the ability to ensure that agent j performs a strategy ensuring φ, or the formula $[i\ \texttt{sstit}]([i\ \texttt{sstit}]\varphi \wedge \texttt{X}[i\ \texttt{sstit}]\psi)$ says that agent i currently performs a strategy that ensures that he is performing a strategy ensuring φ now and immediately after that a strategy ensuring ψ.

Note that in case an agent is performing a strategy ensuring φ conditional on c, he might never perform a strategy ensuring φ. This is obvious since the condition may never be satisfied (for instance if $c = \bot$). It also shows that this conditional strategy of the agent might be interleaved with other strategies. The only guarantee that we have when the agent performs this conditional strategy is that he ensures that the conditional is met. The current formal setting enables us to investigate these conditional strategies in further detail.

Proposition 6. *Let $\mathcal{M}$ be a* G.STRAT *model and $\langle m, h, \vec{\alpha}\rangle$ be a profile in $\mathcal{M}$. Then relative to $\langle m, h, \vec{\alpha}\rangle$ agent i performs a strategy ensuring φ conditional on c if and only if*

$$\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \texttt{sstit}]\texttt{G}(\langle i\ \texttt{sstit}\rangle c \rightarrow [i\ \texttt{sstit}]\varphi),$$

that is, agent i performs a strategy ensuring φ conditional on $\langle i\ \texttt{sstit}\rangle c$.

This means that performing a strategy ensuring φ conditional on c is equivalent to performing a strategy ensuring φ conditional on $\langle i\ \texttt{sstit}\rangle c$. The latter condition is satisfied exactly when my strategy does not ensure $\neg c$, i.e. if there is a history h' in my strategy α_i that validates c. Note that we have $\vDash c \rightarrow \langle i\ \texttt{sstit}\rangle c$ and $\nvDash \langle i\ \texttt{sstit}\rangle c \rightarrow c$, so the above proposition says that performing a strategy ensuring φ conditional on c is equivalent to performing a strategy ensuring φ conditional on the logically weaker condition $\langle i\ \texttt{sstit}\rangle c$.

The following example is devised to show that this property might be very counterintuitive. However, we will subsequently locate where the intuitive mismatch comes from and argue in favour of our formalism by way of another example. But first we will treat an example revealing some unexpected results of our formalization:

Consider the strategy to kill Jones gently conditional on Smith killing Jones.[11] We consider an agent i and let Smith be agent j, and we write K for killing Jones and K_G for killing Jones gently. According to Definition 5 we formalize the performance by agent i of this conditional strategy as $[i\ \texttt{sstit}]\texttt{G}([j\ \texttt{sstit}]K \rightarrow [i\ \texttt{sstit}]K_G)$, which is equivalent to $[i\ \texttt{sstit}]\texttt{G}(\langle i\ \texttt{sstit}\rangle[j\ \texttt{sstit}]K \rightarrow [i\ \texttt{sstit}]K_G)$ by Proposition 6. Note that allowing Smith to kill Jones is logically equivalent to Smith having the ability to kill Jones, i.e. $\vDash \langle i\ \texttt{sstit}\rangle[j\ \texttt{sstit}]K \leftrightarrow$

using the formula $[i\ \texttt{sstit}](c \rightarrow [i\ \texttt{sstit}]\varphi)\texttt{U}\,\psi_{ddl}$ (where ψ_{ddl} denotes the deadline). In this paper we do not study the conceptual and formal adequacy of this proposal, however it is a direction worth pursuing in future work.

11 As the attentive reader may already notice, this example is inspired by the Forrester paradox (see [6]).

$\Diamond[j\ \texttt{sstit}]K$[12]. Hence we derive that performing this conditional strategy is equivalent to $[i\ \texttt{sstit}]\texttt{G}(\Diamond[j\ \texttt{sstit}]K \rightarrow [i\ \texttt{sstit}]K_G)$. In words, performing the strategy to kill Jones gently conditional on Smith killing Jones is logically equivalent to performing the strategy to kill Jones gently conditional on Smith having the ability to kill Jones. Note that Smith's ability to kill Jones does not imply that he actually kills Jones. So it feels like this does not meet our intuitions.

The discussion of the previous paragraph amounts to:

Proposition 7. *Let $\mathcal{M}$ be a* G.STRAT *model and $\langle m, h, \vec{\alpha}\rangle$ be a profile in $\mathcal{M}$. We see that relative to $\langle m, h, \vec{\alpha}\rangle$ agent i performs a strategy ensuring φ conditional on j performing a strategy ensuring ψ if and only if*

$$\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \models [i\ \texttt{sstit}]\texttt{G}(\Diamond[j\ \texttt{sstit}]\psi \rightarrow [i\ \texttt{sstit}]\varphi),$$

that is, agent i performs a strategy ensuring φ conditional on the (settled) condition that j is able to perform a strategy ensuring ψ.

What does this result and the motivating example teach us? I think we should draw the conclusion that conditional strategies are quite hard to perform. At the very least the performance of such a conditional strategy might incline you to perform some actions while the condition is not met. This happens especially when the conditions are not settled.

One may see this result as providing evidence against our formalization of performing a conditional strategy, therefore I now want to give an example in favor of our formal result:

Consider the example of the gentle killing of Jones conditional on Smith killing Jones. Suppose Smith has the ability to kill Jones by lethal injection. However, if he does it will take an hour before Jones actually dies. In this scenario I do not know whether Smith kills Jones by lethal injection and the only way to find out is running a test which takes two hours to process. Hence there is no way for me to find out whether Smith kills Jones before Jones dies. If I do not kill Jones gently before he dies, I could have failed to ensure the condition in the conditional strategy (since Smith could have injected Jones with a lethal dose). We conclude that in order to ensure that the conditional is met I have to kill Jones gently without being certain that Smith kills Jones. In fact, the mere ability of Smith implies that I have to kill Jones gently in order to ensure the conditional.

Having motivated our formalization of conditional strategies and the formal result of Proposition 7 we now try to locate the (or: a probable) source of our previous intuition that the formal result of Proposition 7 was flawed. In order to do this, we shift our attention to *settled* conditions.

A formula being settled means that it is determined only by the current moment, i.e. the property is *moment-determinate.* This means that it does not depend on the current history, nor on the strategies that the agents are currently performing. In contrast to other formalisms the current STIT framework

[12] This follows from the axiom of independence of agency.

allows us to differentiate between static and dynamic properties. These moment-determinate formulas correspond to static properties.

What happens if the conditions are settled?

Proposition 8. *Let $\mathcal{M}$ be a* G.STRAT *model and $\langle m, h, \vec{\alpha} \rangle$ be a profile in $\mathcal{M}$. Suppose $\mathcal{M} \vDash c \to \Box c$. Then relative to $\langle m, h, \vec{\alpha} \rangle$ agent i performs a strategy ensuring φ conditional on c if and only if*

$$\mathcal{M}, \langle m, h, \vec{\alpha} \rangle \vDash [i \ \texttt{sstit}]\mathtt{G}(\Box c \to [i \ \texttt{sstit}]\varphi),$$

that is, agent i performs a strategy ensuring φ conditional on c being settled.

Note that we have $\vDash \Box c \to c$ and $\nvDash c \to \Box c$. The above proposition says that, if in our model c implies that it is settled, performing a strategy ensuring φ conditional on c is equivalent to performing a strategy ensuring φ conditional on the logically stronger condition of c being settled.

I believe it is a key contribution of the current enterprise that we observe that our intuitions of conditional strategies could be flawed because we tend to think of conditions as being moment-determinate.

4 Discussion

In this paper we have presented a formalization of conditional strategies. The formal setting provides us with useful tools to examine consequences and equivalences of conditional strategies. We showed that performing a strategy ensuring φ conditional on c is equivalent to performing a strategy ensuring φ conditional on the logically weaker condition $\langle i \ \texttt{sstit} \rangle c$. This observation has some counterintuitive instances. I argued in favour of the presented formalization and the derived results through examples. I believe this mismatch with our intuitions stems from a misunderstanding that we regard the conditions to be *moment-determinate*.

It is useful to review the work done in [11] in light of the current research. Recall from the introduction that the logic they propose has two operators that involve conditional strategies, namely $[\psi \mapsto a]^i$ and $\pi \mapsto \sigma$, which are interpreted as 'player i chooses move a whenever p holds' and 'player i sticks to the specification given by σ if on the history of play, all moves made by $\bar{i}$ conform to π' respectively. The ψ in the first formula is of a special nature, namely it is a boolean combination of propositional letters, and the second formula represents the condition that the other players *have acted* conform π.

We see that the conditions they allow for conditional strategies are in fact moment-determinate. We recall that (contrary to our formal framework) they assign to each atomic proposition a set of states at which the proposition holds. Therefore, any boolean combination of atomic proposition will be moment-determinate. Furthermore, it is quite natural to assume that the history of play is settled, i.e. moment-determinate. So it intuitively follows that both types of conditional strategies proposed in [11] involve a moment-determinate condition.

We do not know whether the authors are aware of possible challenges arising when one takes a more liberal view on the conditions involved.

We believe that the G.STRAT framework is an adequate formal setting to investigate conditional strategies with a clear mind. Additionally, it seems that this framework can be fruitfully applied to forms of strategic reasoning by zooming in on the structure of the strategies.

A Appendix: Proofs of propositions

We prove the propositions using the following two lemmas:

Lemma 9. *For all formulas* ψ, χ *we have that* $([i\ \mathtt{sstit}](\psi \to [i\ \mathtt{sstit}]\chi)) \leftrightarrow (\langle i\ \mathtt{sstit}\rangle\psi \to [i\ \mathtt{sstit}]\chi)$ *is valid.*

Proof. $\Rightarrow$: Suppose $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}](\psi \to [i\ \mathtt{sstit}]\chi) \wedge \langle i\ \mathtt{sstit}\rangle\psi$. Then for all profiles $\langle m, h', \vec{\beta}\rangle$ with $\beta_i = \alpha_i$ we have $\mathcal{M}, \langle m', h', \vec{\beta}\rangle \vDash \psi \to [i\ \mathtt{sstit}]\chi$. In addition, there exists a profile $\langle m, h', \vec{\beta'}\rangle$ such that $\mathcal{M}, \langle m, h', \vec{\beta'}\rangle \vDash \psi$. Hence for this profile we have $\mathcal{M}, \langle m, h', \vec{\beta'}\rangle \vDash [i\ \mathtt{sstit}]\chi$, which implies $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\chi$, since $\beta'_i = \alpha_i$.

$\Leftarrow$: Suppose $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash \langle i\ \mathtt{sstit}\rangle\psi \to [i\ \mathtt{sstit}]\chi$. Let $\langle m, h', \vec{\beta}\rangle$ be any profile such that $\beta_i = \alpha_i$ and $\mathcal{M}, \langle m, h', \vec{\beta}\rangle \vDash \psi$. We see that this implies $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash \langle i\ \mathtt{sstit}\rangle\psi$. Hence $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\chi$. It follows that $\mathcal{M}, \langle m, h', \vec{\beta}\rangle \vDash [i\ \mathtt{sstit}]\chi$, so we conclude that $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}](\psi \to [i\ \mathtt{sstit}]\chi)$. □

Lemma 10. *For all formulas* ψ *we have that* $[i\ \mathtt{sstit}]\mathtt{G}[i\ \mathtt{sstit}]\psi \leftrightarrow [i\ \mathtt{sstit}]\mathtt{G}\psi$ *is valid.*

Proof. The left-to-right is immediate from the factivity of $[i\ \mathtt{sstit}]$. To prove the right-to-left implication we assume $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\mathtt{G}\psi$. This means that for all profiles $\langle m', h', \vec{\beta}\rangle$ with $\beta_i = \alpha_i$ and $m' \geq_{h'} m$ we have $\mathcal{M}, \langle m', h', \vec{\beta}\rangle \vDash \psi$. Any profile $\langle m', h'', \vec{\gamma}\rangle$ with $\gamma_i = \beta_i$ also satisfies $\gamma_i = \alpha_i$, so the previous observation implies that for these we also have $\mathcal{M}, \langle m', h'', \vec{\gamma}\rangle \vDash \psi$. Hence $\mathcal{M}, \langle m', h', \vec{\beta}\rangle \vDash [i\ \mathtt{sstit}]\psi$. This proves the claim. □

Now we prove the Propositions using these lemmas:

Proof. In all propositions we assume that agent i performs a strategy ensuring φ conditional on c, i.e. $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\mathtt{G}(c \to [i\ \mathtt{sstit}]\varphi)$.

Proposition 6 Lemma 10 implies that we have $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\ \mathtt{G}\ [i\ \mathtt{sstit}](c \to [i\ \mathtt{sstit}]\varphi)$. Applying Lemma 9 by noting that we can replace with equivalents within a $[i\ \mathtt{sstit}]$- and a $\mathtt{G}$-modality, gives us $\mathcal{M}, \langle m, h, \vec{\alpha}\rangle \vDash [i\ \mathtt{sstit}]\mathtt{G}(\langle i\ \mathtt{sstit}\rangle c \to [i\ \mathtt{sstit}]\varphi)$.

Proposition 7 We assume that $c = [j\ \mathtt{sstit}]\psi$. This proposition now follows from the previous proposition and the fact that $\langle i\ \mathtt{sstit}\rangle[j\ \mathtt{sstit}]\psi \leftrightarrow \Diamond[j\ \mathtt{sstit}]\psi$.

Proposition 8 Under the assumption that $\mathcal{M} \vDash c \to \Box c$ we have $\mathcal{M} \vDash c \leftrightarrow \Box c$. Hence, $\mathcal{M} \vDash (c \to [i\ \mathtt{sstit}]\varphi) \leftrightarrow (\Box c \to [i\ \mathtt{sstit}]\varphi)$. This gives us $\mathcal{M} \vDash [i\ \mathtt{sstit}]\mathtt{G}(c \to [i\ \mathtt{sstit}]\varphi) \leftrightarrow [i\ \mathtt{sstit}]\mathtt{G}(\Box c \to [i\ \mathtt{sstit}]\varphi)$, which completes the proof of this proposition.

$\Box$

References

1. Rajeev Alur, Thomas A. Henzinger, and Orna Kupferman. Alternating-time temporal logic. *Journal of the ACM (JACM)*, 49(5):672–713, 2002.
2. Nuel Belnap, Michael Perloff, and Ming Xu. *Facing the Future. Agents and Choices in Our Indeterminist World. 2001.* Oxford University Press, 2001.
3. Jan Broersen. A stit-Logic for Extensive Form Group Strategies. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03*, pages 484–487. IEEE Computer Society, 2009.
4. Jan Broersen. CTL. STIT: enhancing ATL to express important multi-agent system verification properties. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 683–690. International Foundation for Autonomous Agents and Multiagent Systems, 2010.
5. Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Strategy logic. In *CONCUR 2007–Concurrency Theory*, pages 59–73. Springer, 2007.
6. James William Forrester. Gentle Murder, or the Adverbial Samaritan. *The Journal of Philosophy*, 81(4):193–197, 1984.
7. John F. Horty. *Agency and deontic logic.* Oxford: Oxford University Press, 2001.
8. Fabio Mogavero. Reasoning About Strategies. In *Logics in Computer Science*, number 3, pages 85–116. Atlantis Press, 2013.
9. Rohit Parikh. The logic of games and its applications. *North-Holland Mathematics Studies*, 102:111–139, 1985.
10. Marc Pauly and Rohit Parikh. Game logic-an overview. *Studia Logica*, 75(2):165–182, 2003.
11. Ramaswamy Ramanujam and Sunil Easaw Simon. Dynamic Logic on Games with Structured Strategies. In *KR*, pages 49–58, 2008.
12. Johan van Benthem. Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4):219–248, 2001.
13. Johan van Benthem. Extensive games as process models. *Journal of logic, language and information*, 11(3):289–313, 2002.
14. Wiebe van der Hoek, Wojciech Jamroga, and Michael Wooldridge. A logic for strategic reasoning. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 157–164. ACM, 2005.
15. Dongmo Zhang and Michael Thielscher. Representing and Reasoning about Game Strategies. *Journal of Philosophical Logic*, 44(2):203–236, 2014.