Connecting Logics of Choice and Change

Johan van Benthem and Eric Pacuit

Abstract This chapter is an attempt at clarifying the current scene of sometimes competing action logics, looking for compatibilities and convergences. Current paradigms for deliberate action fall into two broad families: dynamic logics of events, and STIT logics of achieving specified effects. We compare the two frameworks, and show how they can be related technically by embedding basic STIT into a modal logic of matrix games. Amongst various things, this analysis shows how the attractive principle of independence of agents' actions in STIT might actually be a source of high complexity in the total action logic. Our main point, however, is the compatibility of dynamic logics with explicit events and STIT logics based on a notion that we call 'control'—and we present a new system of dynamic-epistemic logic with control that has both. Finally, we discuss how dynamic logic and STIT face similar issues when including further crucial aspects of agency such as knowledge, preference, strategic behavior, and explicit acts of choice and deliberation.

1 Introduction: Logical Frameworks for Agency

The STIT logic of Belnap et al. (2001) and its variants have proven fruitful tools to help philosophers and computer scientists explore their intuitions about agency and social interaction. These logics provide a framework to reason about choices,

We thank Roberto Ciuni and Thomas Müller for helpful comments on this chaper.

J. van Benthem (✉)
Institute for Logic, Language and Computation (ILLC), University of Amsterdam,
P. O. Box 94242, 1090 GE Amsterdam, The Netherlands
e-mail: J.vanBenthem@uva.nl

E. Pacuit
Department of Philosophy, Skinner Building, University of Maryland,
College Park, MD 20742-7505, USA
e-mail: epacuit@umd.edu

T. Müller (ed.), *Nuel Belnap on Indeterminism and Free Action*,
Outstanding Contributions to Logic 2, DOI: 10.1007/978-3-319-01754-9_14,

abilities and actions of agents, all placed in a temporal setting. And further issues lie just below the surface: what agents know or believe at the time of choice, how they act based on preferences, and engage in deliberate strategic interaction (cf. Horty 2001).

But STIT is not the only game in town. Many logical paradigms are active in the above territory, and they often show clear similarities. This calls for analysis and reflection. For instance, van Benthem and Pacuit (2006) relate the major varieties of epistemic temporal logics, coming from mathematical logic, computational logic, and studies of agency. Continuing in this line, van Benthem et al. (2009) prove representation theorems linking dynamic-epistemic models with epistemic-temporal ones, making it possible to enlist ideas from one logic in the service of the other. In the case of STIT, too, much has been done to clarify its connections with other frameworks. In fact, Belnap et al. (2001) already pointed out links with earlier work of Chellas (1992), to which one can add the neighborhood logics of ability in (Brown 1988, 1992). Moreover, connections have been found with coalition logic (Broersen et al. 2006b) and alternating-time temporal logic (Broersen et al. 2006a), while Lorini and Schwarzentruber (2010) relates STIT to logics for strategic and extensive games—a line that we will continue in this chapter (cf. also Herzig and Lorini 2010). Finally, Ciuni and Zanardo (2010) shows how STIT extends well-known logics of branching time.

Our aim in this chapter is to continue in the latter vein, and connect STIT models further with modal models for action from the realm of propositional dynamic logic (PDL), modal game logics (see van Benthem 2014), and dynamic-epistemic logic DEL (van Benthem 2011). We start by addressing an initial barrier to making any comparison between these different logical frameworks.

STIT logics are primarily intended as logics of *ontic* freedom and indeterminacy while the logical systems we discuss in this chapter are focused on *epistemic* uncertainty (i.e., knowledge about what will happen next). The heart of our comparison is the simple observation that the basic STIT modality turns out to be precisely the "knowledge" modality found in many epistemically-oriented logical systems. Importantly, however, we are *not* suggesting that all discussion about "agency" and agents making choices in an indeterministic world can or should be replaced with an analysis of what the agents know about their own choices and the consequences of their actions in an indeterministic world, or vice versa. Our point is simply that similar logical frameworks are open to different interpretations. The goal is not to argue for the primacy of any single interpretation, but rather to demonstrate how two different perspectives on modeling rational agency can lead to similar insights. This is in line with a broader goal. The arena of logics for agency appears to be moving from an initial stage of a "Battle of the Sects" to a more detached understanding of both similarities and relative advantages of different paradigms, leading to a more unified sense of purpose and methodology.

2 Preliminaries: The STIT Framework

In this section, we introduce the basic STIT framework. We will be very brief, only touching on the key notions we need later in this chapter. For more information, the reader is invited to consult (Horty 2001; Belnap et al. 2001; Horty and Belnap 1995; Balbiani et al. 2008).

STIT structures STIT models are based on *branching-time frames*, structures $\langle T, <\rangle$ where T is a nonempty set of "moments", and $<$ is a strict partial order on T without backwards branching: for all m, m', m'', if $m' < m$ and $m'' < m$, then either $m' \leq m''$ or $m'' \leq m'$ (where $x \leq y$ iff $x < y$ or $x = y$). A *history* is a maximal linearly ordered subset of T. Let Hist denote the set of all histories and for $t \in T$, $H_t = \{h \in \mathsf{Hist} \mid t \in h\}$ is the set of histories containing moment t.

At each moment, there is a choice available to the agent. Let $\mathcal{A}$ be the set of agents. Formally, the choices available to agent i at moment t are represented by a partition $Choice_i^t$ on the set H_t of histories containing t. Let $Choice_i^t(h)$ denote the cell containing h. Since $Choice_i^t$ is a partition, we have for each $i \in \mathcal{A}$ and $t \in T$, $Choice_i^t \neq \emptyset$ and $\emptyset \notin Choice_i^t$. In addition, the choice partitions of the agents must satisfy one additional condition:

Independence For all $t \in T$ and all $s_t : \mathcal{A} \to \wp(H_t)$ with $s_t(i) \in Choice_i^t$, $\bigcap_{i \in \mathcal{A}} s_t(i) \neq \emptyset$.

Now we define a *STIT model* as a tuple $\langle T, <, \mathcal{A}, Choice, V\rangle$, where $\langle T, <\rangle$ is a branching-time frame, $\mathcal{A}$ is a finite set of agents, $Choice$ is a function assigning to each $i \in \mathcal{A}$ and $t \in T$ a partition on H_t satisfying Independence, and V is a function assigning to each atomic proposition a set of history/moment pairs ($V : \mathsf{At} \to \wp(T \times \mathsf{Hist})$).

STIT language Let At be a set of atomic propositions. The STIT language is the smallest set of formulas generated by the following grammar

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i\ \mathsf{stit}]\varphi \mid \Box\varphi$$

where $p \in \mathsf{At}$ and $i \in \mathcal{A}$. Additional boolean connectives ($\vee, \to, \leftrightarrow$) are defined as usual. Further, $\langle i\ \mathrm{stit}\rangle\varphi$ is the dual modality $\neg[i\ \mathsf{stit}]\neg\varphi$ and $\Diamond$ the dual $\neg\Box\neg\varphi$. The interpretation of $[i\ \mathsf{stit}]\varphi$ is that "agent i *sees to it that* φ is true" and the historic necessity $\Box\varphi$ means that "φ is true at all alternative histories".

STIT Semantics Let $\mathcal{M} = \langle T, <, \mathcal{A}, Choice, V\rangle$ be a STIT model. Truth of a STIT formula φ is defined inductively as follows, at pairs t/h of histories h and moments t on them:

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$
- $\mathcal{M}, t/h \models \varphi \wedge \psi$ iff $\mathcal{M}, t/h \models \varphi$ and $\mathcal{M}, t/h \models \psi$
- $\mathcal{M}, t/h \models \Box\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in H_t$
- $\mathcal{M}, t/h \models [i\ \mathsf{stit}]\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in Choice_i^t(h)$

In addition, one sometimes defines an additional STIT operator (the so-called "deliberative STIT"):

- $\mathcal{M}, t/h \models [i \text{ dstit}]\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in Choice_i^t(h)$ and there is a $h'' \in H_t$ such that $\mathcal{M}, t/h'' \models \neg\varphi$

This modality is definable in the basic language: $[i \text{ dstit}]\varphi \ := \ [i \text{ stit}]\varphi \wedge \Diamond\neg\varphi$. A number of other STIT-operators can be found in the literature. For example, the "achievement STIT operator" (see Horty and Belnap 1995, Sect. 2.2 for a definition and discussion) and the "next time STIT operator" (Broersen 2011) both make use of the underlying past and future time structure.

Logic and axiomatics The models and language are one major aspect of current uses of STIT, as a style of representing action semantically. However, there is also the issue of syntactic proof rules for reasoning about action. The following axiomatization was proven sound and complete for the class of all STIT models in (Xu 1995; Balbiani et al. 2008):

- The **S5** axioms for $\Box$ and $[i \text{ stit}]$: $\bigcirc(\varphi \to \psi) \to (\bigcirc\varphi \to \bigcirc\psi)$, $\bigcirc\varphi \to \varphi$, $\bigcirc\varphi \to \bigcirc\bigcirc\varphi$, $\neg\bigcirc\varphi \to \bigcirc\neg\bigcirc\varphi$, for $\bigcirc \in \{\Box, [i \text{ stit}]\}$
- $\Box\varphi \to [i \text{ stit}]\varphi$
- $\left(\bigwedge_{i\in\mathcal{A}} \Diamond[i \text{ stit}]\varphi_i\right) \to \Diamond\left(\bigwedge_{i\in\mathcal{A}}[i \text{ stit}]\varphi_i\right)$
- Modus Ponens and Necessitation for $\Box$.

It will be clear that these axioms do not reflect, let alone enforce, any particular view of time, whether branching or linear. This is no accident. The basic ideas of STIT seem compatible with about every major temporal logic that is on the market.

Now that we have all major components of STIT on the table, we will discuss its semantics and axiomatics in relation to other approaches for studying agency coming from the "dynamic logic family". We will not define these other frameworks in any detail, but refer the reader to the literature on dynamic logic, game logics, and dynamic-epistemic logics cited in this chapter.

3 Modeling Choice Situations

3.1 The Modal Heart of Choice

Abstracting from the temporal component that could come from any existing framework, the heart of STIT-style choice is a very simple **S5** logic.[1] A **STIT choice scenario** for a set of agents $\mathcal{A}$ is a tuple $\mathcal{M} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, V\rangle$, where W is a nonempty set, for each $i \in \mathcal{A}$, $\sim_i$ is an equivalence relation on W (we write $[w]_i$

[1] An earlier modal analysis of STIT scenarios can be found in Herzig and Schwarzentruber (2010), Balbiani et al. (2008) and follow-up literature—but in this chapter, we will eventually choose a path of our own.

for the equivalence class of w under $\sim_i$) and V is a valuation function. We focus on two agents ($\mathcal{A} = \{1, 2\}$) for convenience in what follows. STIT choice scenarios are standard multi-agent **S5** models, and so a simple modal language describes them: for each $i \in \mathcal{A}$, use '$[\,i\,]$' for the modality matching the relation $\sim_i$ and 'E' for the existential modality.[2] The Independence assumption above corresponds to the validity of the following *product axiom*:

$$(E[\,1\,]\varphi \wedge E[\,2\,]\psi) \rightarrow E(\varphi \wedge \psi)$$

By standard frame correspondence, this says that any pair of choices for the two agents overlap.

The key idea of STIT in these models may be called *control*: the equivalence relations represent the extent to which agents control outcomes by their choices. The product axiom says that no agent can prevent any other agent from making any of her choices. There is more to this condition than meets the eye. For instance, assume that agent 1 has a singleton choice somewhere. Since 2's choices must always overlap with this singleton, and different choices are disjoint, it follows that 2 has only one choice set.

The logic of these models is many-agent **S5** plus the product axiom. In this basic system, we can derive interesting facts, such as

$$[\,1\,][\,2\,]\varphi \leftrightarrow [\,2\,][\,1\,]\varphi \leftrightarrow U\varphi$$

where U is the universal modality dual to E.[3] In slightly extended modal languages, more can be proved. For instance, the previous comment about singleton choices amounts to the validity of $([\,1\,](\varphi \wedge \neg D\varphi) \wedge E[\,2\,]\varphi) \rightarrow U\varphi$, where $D\varphi$ is the *difference modality* true at a world w if there is a $v \neq w$ such that $\mathcal{M}, v \models \varphi$. Thus, the product axiom packs a lot of punch.

So, basic STIT logic is a nice simple multi-**S5**-extension. This first natural connection with modal logic shows that we are at least generally in the same world as modal logics of action.[4]

---

[2] Truth for these operators is defined as usual: $\mathcal{M}, w \models [\,i\,]\varphi$ iff for all $v \in W$, if $w \sim_i v$ then $\mathcal{M}, v \models \varphi$, and $\mathcal{M}, w \models E\varphi$ iff for some $v \in W$, $\mathcal{M}, v \models \varphi$.

[3] As observed in Balbiani et al. (2008), this principle can also function as a product axiom by itself. Also inter-derivable with our version of the product axiom is the stronger-looking $(E[\,1\,]\varphi \wedge E[\,2\,]\psi) \rightarrow E([\,1\,]\varphi \wedge [\,2\,]\psi)$., for which Roberto Ciuni has proposed an interesting epistemic interpretation.

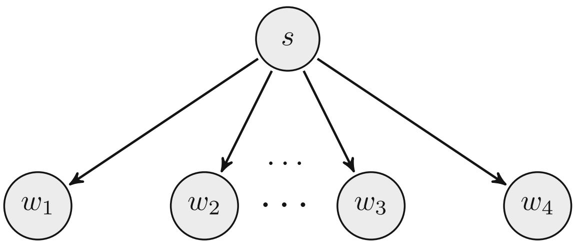
[4] These simple modal equivalence models show up when studying many aspects of rational agency: they work for specifying ranges of knowledge, issues in the logic of questions, etc.

3.2 An Initial Comparison with Modal Logics of Action

Broadly speaking, there are two general views about how to model the actions available to an agent. The first is the view found in STIT as presented in Sects. 2 and 3.1 above. Let us now consider the second view, that of modal and dynamic logics of actions (see Harel et al. 2000 for a discussion), which is also the main model of action in Situation Calculus (Reiter 2001), Automata Theory, Decision Theory and Game Theory. Its general idea is to think of actions as *transitions* moving between different "states of the system". This happens again in standard modal models $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Act}}, V, s\rangle$, with worlds in W viewed as states of some process (s is the initial state), and labeled transition relations $R_a \subseteq W \times W$ for each action label $a \in \mathsf{Act}$. Each relation R_a indicates the possible executions of the basic action a. Modal languages over these models then describe possible effects of actions, while real dynamic logics also have an explicit language for speaking about complex actions defined by means of sequential composition, conditional choice, or iteration.[5] We use the phrase *PDL scenarios* for this family of paradigms.

At a first glance, these are very different views. While both perspectives acknowledge variety in possible outcomes of actions, they also have structure that the other lacks. In action-labeled approaches, the primary emphasis is on actions or events themselves and their properties, of which a description of outcome states seems only one. For instance, dancing a tango involves many features in addition to its end state: we would trivialize the process by just having an end state of 'having danced a tango'. On the other hand, many daily actions expressed in natural language are largely defined by just post-conditions on their outcomes, witness 'opening the door' or 'posting a letter'. In that sense, STIT's approach to describing actions is very natural.

We now proceed to a more technical comparison of the two styles. But to do so, we need some further touches. For a start, our simple modal picture of STIT choice situations takes out all of temporal structure. However, for a comparison with PDL, it seems more concrete to view the above 'worlds' as steps emanating from some root toward next states in a tree, a snap-shot of an ongoing decision process. The actual world is then the actual transition from the root to some next state:



[5] This is just a first intuitive pass. We will have occasion to spell out things further later on.

What this suggests is introducing a richer modal language for basic STIT, referring also to the two stages: 'now' and 'next'. This motivates the NEXT-STIT of Broersen (2011), and we will also encounter this setting in the DEL-style logic of Sect. 6. But right now, we continue in a semantic mode with worlds viewed intuitively as transitions.

Likewise, in order to compare STIT with PDL, we must also clarify the intuitive interpretation of PDL-style models. In particular, there are two broad views in the literature. One is that of transition models as abstract processes or machines, the other as unraveled temporal executions. On the *process view*, worlds are states in a process, and the relations indicate possible transitions. On this view, the model is a sort of automaton, perhaps in a very compact form, where many different transition relations can go from one state to the same next state. By contrast, the second view of PDL-models is one of *unraveled temporal execution*. Intuitively, once a process starts working, it produces a temporal universe of *executions*, being histories of successive admissible actions (cf. Clarke et al. 2000; Clarke and Emerson 1981 for this view). For the usual modal languages of action, the difference between the two views does not matter, since the execution tree is just a bisimilar unraveling of the process. And vice versa, we can think of a process as a sort of bisimulation-contracted essence of what can happen in the execution tree. But in our present setting, comparing with STIT seems to favor the temporal execution view.[6]

We therefore continue with the temporal view, where for simplicity, all event labels are taken to be unique.[7] Like with the above basic STIT, we will not take the full temporal models here, but just the snapshots of a one-step action. A **PDL action scenario** is a set of labeled transitions from some initial state s, each leading to a different successor state. This can be viewed as an obvious special "one-shot" case of the earlier-mentioned transition models.

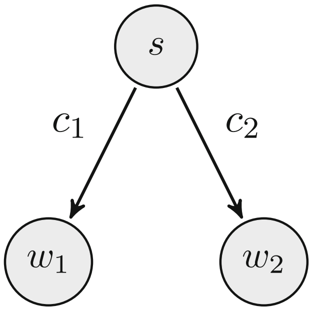
3.3 Merging the Two Perspectives on Action

Our goal in this chapter is not to reduce STIT models to PDL models, or vice versa. We find it more rewarding to show connections between the two perspectives leading to merged systems.

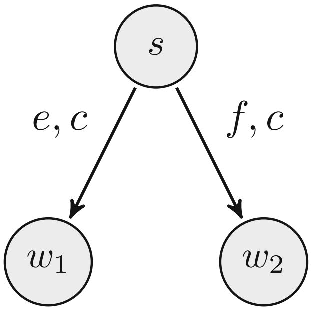
For better focus, we start with the single-agent case. Consider a simple STIT choice situation with two states $W = \{w_1, w_2\}$ and two equivalence classes: $[w_1] = \{w_1\}$ and $[w_2] = \{w_2\}$. Thus, there are two choices for the agent, which we label c_1 and c_2, respectively. A simple corresponding PDL action scenario has two transitions from the root state s labeled by c_1 and c_2, respectively:

[6] However, the process view of PDL may be closer to the dynamics of agents making choices and performing actions. We do not claim that our take in this chapter is the only way to go.

[7] This uniqueness is standard modeling practice in many temporal formalisms: if histories differ at a point, then there should be a difference in the next event.



This seems straightforward, but we have not yet found the real structure that we need. To get at this, consider a STIT choice situation with the same two states, but now only one equivalence class $[w_1] = [w_2] = \{w_1, w_2\}$. Now the agent only has one choice c. We cannot label the two transitions by c now, since that gives a PDL model with the same event, and it is unclear how this would fit the scenario. Here the difficulty is not that we cannot label the transitions: We can introduce different events for them, say e and f. In fact, this makes sense even in STIT, since histories consist of events, and as we said earlier, if two histories are different, this is because different events take place on them. But this still does not address the matter of the choice structure, and crucially related to this: how we interpret the branching in our PDL model.[8] What emerges here is an ambiguity in the usual talk about PDL models. In particular, *what do branchings mean*? Sometimes, people talk as if these are conscious choices a process or an agent can make, sometimes as if they are variations that cannot be predicted. What we need to distinguish the two senses is precisely the notion provided by STIT, that of *control*. In our first scenario, the two labels c_1 and c_2, when added to events, divide them into two control equivalence classes. In the second scenario, adding the label c to both e and f indicates how the events belong to the same control class. The agent cannot choose between the events.



[8] We could view the branching as "non-determinism" in PDL, but this does not clarify the issues very much. Non-determinism usually means that a process has several options, 'ways of doing c', but that is not the situation in the STIT model: it is not up to the agent to non-deterministically chose one or the other transition.

Action as 'events under control' To us, the preceding discussion suggests that it make sense to pool ideas. PDL has labeled events, and this makes sense, if we want to describe what happens on histories regardless of agents' choices. But STIT adds the notion of control, which also makes sense as a key feature of agency, and this helps remove a potential ambiguity in thinking about PDL models. The resulting view of actions is this:

$$Action = events + control$$

In line with this, it makes sense to merge the basic ideas of PDL and STIT into a logic with both features. Its models can have pair labels: $(event, choice)$ for transitions, thinking of equivalence relations of control on either whole transition relations, or on concrete state transitions. Such structures support a joint language with PDL event modalities $[\,e\,]$ and STIT modalities $\langle[i]\rangle$. We will not pursue technical details here, since we will discuss concrete systems of this kind later in this chapter in the modal game logics of Sect. 4, and the dynamic epistemic logic of Sect. 6. For the moment, it suffices to note that it is quite possible to have the best of both worlds, in combined logics that might be called "eventful STIT" or "controlled PDL".

We end with two more general comments about this encounter.

More on interpretations of PDL The confrontation with STIT leads to some useful clarification. We already mentioned the two main views of models as representing 'process' structure versus 'execution space'. We also discussed a major ambiguity in how one interprets branching. As a final point, we mention the issue of *events versus actions*. There is a lot of loose talk in the PDL literature about action and choice. For instance, back-and-forth clauses in bisimulation are justified by looking at 'internal choices' that a process has, and one often talks about events in PDL models as actions performed by agents.[9] But really, PDL talks about arbitrary events, all further meaning for these has to be supplied additionally in different settings. In particular, actions by agents are events with special further structure, and if it matters, these need to be made explicit. The above case of 'control' is one clear instance.[10]

A caveat about framework comparison In this section, we have engaged in high-level framework comparison. But rarefied air can exaggerate ideological differences, and it is important to also think of applied experience. In modeling practice, framework differences often prove much less dramatic than expected, as is well-known from the fact that the same real process can often be specified very happily in quite different computational paradigms. For instance, in our setting, dealing with concrete scenarios of choices and actions requires an explicit *modeler's decision* as to individuating states and actions: formal frameworks themselves do not tell us how

[9] The same is true in the dynamic epistemic logic literature: notice the terminology 'action models' versus 'event models' for its core update rules.

[10] By itself our point is not new. Adding internal structure is crucial when modeling *simultaneous* action, where one endows PDL events with internal vector structure, as in the 'interpreted systems' of Fagin et al. (1995) or the parallel games of van Benthem et al. (2008).

to do that. But then, differences between STIT and PDL tools may just amount to different legitimate decisions on how one individuates actions.[11] The problem of individuating actions has been discussed extensively in philosophy (see footnote 3 on pg. 588 of Horty and Belnap (1995) for a concise explanation), and it also shows in modeling practice in computer science.

This brings us to an important philosophical issue which we have thus far swept under the rug. In PDL models, actions are labels of transitions and this basic sorting of transitions by their labels seems to suggest a particular ontology of actions and events. In STIT models, there is no such sorting and, indeed, the only way to characterize an action is by reference to the outcomes. This raises an important question for the philosophical logician: Does adopting PDL as a logic of *actions* force one to take sides in philosophical debates about the ontology of events and actions? Our response is to bracket this question since we feel that both STIT and PDL models are open to a wide range of philosophical interpretations, regardless of the original intended interpretation of these logical frameworks. However, we certainly admit that this rather mathematical "formal modeling" view is itself controversial and we welcome (and enjoy) debates on this issue. Nonetheless, we hope that the comparative points we are making in this chapter still make sense.

4 A Merged System: Matrix Game Logic

Now, we want to make our comparisons and merges more concrete by looking at a concrete modal logic that already existed independently, and that turns out to shed some additional light on the semantic and axiomatic aspects of STIT meeting PDL.

Choices and pair events Let us return to the STIT choice situation for two agents. There is an actual world with the choices that were actually made. It makes sense to think of the worlds here as pairs of actions chosen. Note that each world w can be mapped to a unique pair of equivalence classes containing it, one for each agent, and by the product axiom, this map to pairs of equivalence classes is surjective. What we do not know is whether the map is injective, and indeed it may not be, unless we modify the product axiom to require that *different choices for all the agents have singleton intersections*. The latter constraint says that all slack in choices has been explained by introducing enough agents—perhaps including the 'environment' to take up all remaining slack. There is some simple arithmetic involved here. Assume that our model is finite. The product axiom with the singleton clause forces all

[11] As a concrete example, suppose there are two histories h, h' where an agent refrains from choosing either. Presumably, refraining means she could have made a choice for h or for h'. One way of viewing this involves three actions: choosing h, choosing h', or 'leaving things be': h, h'. This would violate the disjointness constraint of STIT. But we can also individuate events differently, with four histories: one where h is chosen, one where h' is chosen, and two copies of these except for the fact that no choice was made.

equivalence classes for agent 1 to have the same size n, as they need room for representatives of all choices of 2. The total size will be $n \times k$, with k the fixed size for 2 that exists similarly. But this suggests a viewpoint in terms of "matrix models" for joint actions that is well-known from logics of games in strategic form (cf. Osborne and Rubinstein 1994). We will develop this analogy here, using a logic proposed in (van Benthem 2007) that provides a particularly apt comparison for STIT, while also doing full justice to the PDL perspective.[12]

4.1 Modal logic of matrix games

Games induce natural models for epistemic, doxastic and preference logics, as well as conditional logics and temporal logics of action. See van der Hoek and Pauly (2006) for an overview of many such systems. Our discussion just takes a small slice.

Recall the definition of a *strategic game* for a set of players N: (1) a set A_i of actions for each $i \in N$, and (2) a utility function or preference ordering on the set of outcomes. For simplicity, one often identifies the outcomes with the set $S = \Pi_{i \in N} A_i$ of *strategy profiles*. Given a strategy profile $\sigma \in S$ with $\sigma = (a_1, \ldots, a_n)$, σ_i is the ith projection (*i.e.*, $\sigma_i = a_i$) and σ_{-i} lists the choices of all agents except agent i: $\sigma_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$.

Now, from a logical perspective, it is natural to treat the set S of strategy profiles as a universe of "possible worlds".[13] Following (van Benthem et al. 2011) for the rest of this subsection, two natural relations can be defined on these worlds. For each $\sigma, \sigma' \in S$, set for each player $i \in N$:

- $\sigma \sim_i \sigma'$ iff $\sigma_i = \sigma'_i$: this epistemic relation represents player i's "view of the game" at the *ex interim* stage where i's choice is fixed but the choices of the other players' are unknown,
- $\sigma \approx_i \sigma'$ iff $\sigma_{-i} = \sigma'_{-i}$: this relation of "action freedom" (a term taken from Seligman (2010)) gives the alternative choices for player i when the other players' choices are fixed.

Control can be freedom Our earlier discussion of STIT was in terms of *control*, including the lack of it inside players' equivalence classes. But in a multi-agent perspective, one person's lack of control is another person's freedom, and labels can switch easily.

This can all be packaged in a standard relational structure

$$\mathcal{M} = \langle S, \{\sim_i\}_{i \in N}, \{\approx_i\}_{i \in N} \rangle$$

[12] What follows here has strong resemblances to earlier work by a number of authors, including (Herzig and Lorini 2010; Balbiani et al. 2008; Lorini 2010; Lorini et al. 2009).

[13] One can also have more abstract worlds in so-called 'models of games', as is usual in epistemic game theory, see (Aumann 1999)—but this generality is not needed in what follows.

with S the set of strategy profiles and the relations just defined. Adding a valuation function interpreting a set At of atomic propositions that represent basic facts about strategy profiles (physical, or game-internal), we get standard multi-modal models.[14]

Such game models support many logical languages, from simple modal formalisms to 'hybrid modal logics', first-order logic, or even non-first-order fixed-point logics. Cf. van Benthem (2010) and Blackburn et al. (2002) on the balance of expressive power and computational complexity that arises in such design choices, a topic that will return below. However, the simplest system will do for us here. In particular, here are the key modalities for a modal logic of strategic games:

- $\sigma \models [\sim_i]\varphi$ iff for all σ', if $\sigma \sim_i \sigma'$ then $\sigma' \models \varphi$.
- $\sigma \models [\approx_i]\varphi$ iff for all σ', if $\sigma \approx_i \sigma'$ then $\sigma' \models \varphi$.

The first modality expresses the knowledge a player has once her choice is made, and given her uncertainty about what others will do, the second modality refers to her freedom of choice. As is well-known, combining the two modalities makes φ true in each world of a matrix game model: $[\sim_i][\approx_i]\varphi$ acts as a universal modality U.[15] This reflects an earlier observation about STIT—and that is no coincidence, witness the observations in Sect. 4.2 below.

What is the deductive power of the basic modal logic of strategic games? As before, we restrict attention to two-player games. First, given the nature of our relations, the separate logics are standard modal **S5** for epistemic outlook and action freedom. In addition, the interaction of these modalities validates further laws. In particular, the above fact about the universal modality is reflected in the following law:

the equivalence $[\sim_i][\approx_i]\varphi \leftrightarrow [\approx_i][\sim_i]\varphi$ is valid in all matrix game models.

This validity depends on, and in fact it expresses, the geometrical "grid property" of game matrices that, if one can go on a path $x \sim_i y \approx_i z$, then there also exists a point u with $x \approx_i u \sim_i z$. We will discuss what this feature means in some more detail in Sect. 4.3.

This concludes our brief introduction to the modal logic of matrix games. For details and further issues, the reader is referred to (van Benthem 2014).

4.2 STIT in Modal Matrix Logic

Given our discussion in Sect. 3, it will be evident how to translate the basic STIT operators into our modal language of matrix games:

---

[14] For example, a proposition p_i^a might say "agent i plays action a in the current profile"—but atomic propositions could also encode utility values for players.

[15] As noted in (van Benthem 2007), another interesting feature of our models is that 'distributed knowledge' $D_G\varphi$ for a group of players accesses those profiles where only players outside the group still have options.

$$[i \text{ stit}]\varphi := [\sim_i]\varphi, \quad \Box\varphi := [\sim_i][\approx_i]\varphi$$

This connection gives just the right combination of what we have called freedom plus knowledge.

Fact 4.1. *Our translation embeds STIT logic faithfully into the modal logic of full matrix games.*

Proof. First consider the direction from STIT theoremhood to modal game logic. Our translation validates the earlier STIT axioms, where the action modality refers to all consequences of the choice actually made, while the freedom modality looks at all alternative histories passing through the current profile. In particular, the quantifier combination employed in the Freedom axiom now becomes derivable through the theorems that are derivable for the STIT modality plus the existential modality E defined as $\langle\approx_1\rangle\langle\approx_2\rangle$:

Fact 4.2. *The formula* $(E[\sim_1]\varphi \wedge E[\sim_2]\psi) \rightarrow E(\varphi \wedge \psi)$ *is derivable in multi-***S5** *plus the commutation law for the two modalities.*

Conversely, to prove that the embedding is faithful, we need to refute each non-valid STIT law in our matrix models. To do so, take any STIT temporal counter-model in the sense of Sect. 2, and note that it suffices to look at the current moment and the next moments only (recall, that our STIT language does not contain temporal modalities). Furthermore, without loss of generality, we assume that this model is finite. More precisely, as in Sect. 3.1, we can abstract a finite two-agent basic STIT **S5**-model out of the temporal structure by letting histories be worlds, and defining agent's equivalence relations respecting their choice partitions. Now, the historic necessity operator is the universal modality while the two STIT modalities are the modalities for the equivalence relations. The last step is to show that we can transform this model into a matrix model.

If the intersections of the equivalence classes, one from each agent, are singletons, then we are done. Otherwise, we proceed as follows. A **cell** is an intersection of the agents' equivalence classes (i.e., $C = [w]_1 \cap [v]_2$ for some states w and v). Since the model is finite, there are finitely many cells and each cell has only finitely many states in them. Furthermore, by the independence assumption, each cell is non-empty. Let m be the number of elements in the largest cell. Without loss of generality, we can assume that all cells contain exactly m states (this may require adding copies of states to the model).

Organize the cells so that they form a matrix where each row contains all the cells making up a 1-equivalence class and each column contains all cells making up a 2-equivalence class. Label each cell by its position in the matrix (so, the pair (x, y) corresponds to the cell in row x and column y). There may be more than one way to organize the cells so that the rows correspond to a 1-equivalence class and the columns correspond to a 2-equivalence class. Our construction does not depend on the choice of labeling. For the remainder of the proof, fix such a $r \times c$ matrix.

Now, construct an $m \times m$ matrix for each cell. Fix a cell C labelled with (x, y) containing states $w_1, \ldots, w_m$. Worlds in the new model with be 4-tuples (i, j, x, y) where (i, j) denotes the position in the matrix and (x, y) denotes the cell containing the world. Formally, let (i, j, x, y) be a copy of $w_{i+j-1 \mod m}$. So, for example, if $m = 3$, then the world $(2, 3, x, y)$ is a copy of w_1. Note that each row and each column contains a copy of all the worlds in C.

The model is $\mathcal{M} = \langle W', \sim_1', \sim_2', V' \rangle$ where $W' = \{(i, j, x, y) \mid i, j \leq m, x \leq r, y \leq c\}$ (where r and c are the number of rows and columns respectively in the outer matrix). We define the uncertainty relations for the agents as follows:

- $(i, j, x, y) \sim_1^0 (i, j', x, y)$ for all $j, j' \leq m$
- $(i, j, x, y) \sim_2^0 (i', j, x, y)$ for all $i, i' \leq m$

So $\sim_1'$ runs along the rows of each inner matrix, and $\sim_2'$ runs along the columns. We extend this relation as follows:

- $(i, j, x, y) \sim_1^0 (i, 0, x, y + 1)$, where the addition is taken modulo m
- $(i, j, x, y) \sim_2^0 (0, j, x + 1, y)$, where the addition is taken modulo m

Let $\sim_1'$ and $\sim_2'$ be the reflexive and transitive closure of $\sim_1^0$ and $\sim_2^0$, respectively. Finally, the valuation V' is copied from the original valuation in the obvious way. We note the following two facts about the construction:

1. If $(i, j, x, y) \sim_1' (i', j', x', y')$, then $i' = i$ and $x' = x$. If $y' = y$, then $w_{i+(j-1) \mod m}$ and $w_{i+(j'-1) \mod m}$ are both in the cell labeled by (x, y), and so $w_{i+(j-1) \mod m} \sim_1 w_{i+(j'-1) \mod m}$. If $y' \neq y$, then $w_{i+(j-1) \mod m}$ and $v_{i+(j'-1) \mod m}$ are in different cells. However, we still have $w_{i+(j-1) \mod m} \sim_1 v_{i+(j'-1) \mod m}$ since we assume that cells in the same row are in the same 1-equivalence class.
2. If $(i, j, x, y) \sim_2' (i', j', x', y')$ then $j' = j$ and $y' = y$. If $x' = x$, then $w_{i+(j-1) \mod m}$ and $w_{i+(j'-1) \mod m}$ are both in the cell labeled by (x, y), and so $w_{i+(j-1) \mod m} \sim_2 w_{i+(j'-1) \mod m}$. If $x' \neq x$, then $w_{i+(j-1) \mod m}$ and $v_{i+(j'-1) \mod m}$ are in different cells. However, we still have $w_{i+(j-1) \mod m} \sim_2 v_{i+(j'-1) \mod m}$ since we assume that cells in the same column are in the same 2-equivalence class.

These observations show immediately that the newly constructed model is bisimilar to the original STIT model. Hence, they satisfy the same formulas in our language.

The last thing we need to check is that the intersection of agents' equivalence classes are singletons. Suppose that $(i_0, j_0, x_0, y_0) \sim_1' (i', j', x', y')$, $(i_0, j_0, x_0, y_0) \sim_1' (i'', j'', x'', y'')$, $(i_1, j_1, x_1, y_1) \sim_2' (i', j', x', y')$ and $(i_1, j_1, x_1, y_1) \sim_2' (i'', j'', x'', y'')$. Then, by construction, $i' = i'' = i_0$, $x' = x'' = x_0$ and $j' = j'' = j_1$ and $y' = y'' = y_1$. Hence, $(i', j', x', y') = (i'', j'', x'', y'')$, as desired.

We have shown that our translation is both correct and faithful.[16] QED

---

[16] Note that the construction given in this proof is only needed because the *singleton intersection property* (the intersection of all the agents equivalence classes are singletons) is not definable in our

This proof exploits the fact that matrix game models are close to the multi-**S5** models for basic STIT defined in Sect. 3.1. Still, the geometrical matrix perspective is useful, since it links up with a body of existing results. We will see a number of examples as we proceed.

4.3 Complexity and Correlation

While the preceding embedding makes sense, it does embed STIT in a system whose behavior is potentially complex. Richer modal logics of matrix games may well be unaxiomatizable and worse. The reason is the above commutation law for the two equivalence relations. While this may look like a pleasant structural feature of matrices, its logical effects are delicate. It is well-known that the general logic of bi-modal languages plus a universal modality on 'grid models' with two immediate successor relations is not decidable, and not even axiomatizable: indeed, it is "Π^1_1-complete" (cf. Halpern and Vardi 1989; Marx 2007; Gabbay et al. 2003; van Benthem and Pacuit 2006). The reason is that grid structure can be exploited to encode computations of Turing machines on successive rows, or geometrical "tiling problems" of known high complexity.

Now, it is not clear whether our most basic modal game logic falls into this trap, since our models only have two *equivalence relations*, one horizontal and one vertical. Indeed, its closeness to STIT may suggest that it remains decidable–even though this does not follow from our earlier embedding result, that went in the opposite direction. Still, Halpern and Vardi (1989) and Spaan (1990) show high complexity of modal logics on grid models with reflexive transitive relations, using an encoding trick with alternating proposition letters.[17]

This potential high complexity, while not directly threatening to STIT, does raise an interesting issue in modeling action. A standard way of defusing high complexity results is by allowing *more models*. In the present setting, the resulting structures are *general game models* where certain strategy profiles may be absent. Then general modal game logic becomes much simpler, being just multi-agent modal **S5** without any connecting axioms (van Benthem 1997).[18]

(Footnote 16 continued)
language. However, if the language contains *group* STIT operators $[G\ \mathsf{stit}]\varphi$ meaning that the group G can see to it that φ is true, then the singleton intersection property is definable via the formula $[\mathcal{A}\ \mathsf{stit}](\varphi \vee \psi) \rightarrow [\mathcal{A}\ \mathsf{stit}]\varphi \vee [\mathcal{A}\ \mathsf{stit}]\psi$, where $\mathcal{A}$ is the set of all agents. Furthermore, the argument would be very different once we consider STIT formulas with temporal operators shifting moments along histories, as is suggested in Sect. 7.

[17] Such encodings also work with two equivalence relations and common knowledge in one dimension of the grid model, while time provides the other dimension. See van Benthem and Pacuit (2006) for an extensive survey.

[18] For a concrete counter-example, note that the formula in Fact 4.2 is not valid on such models. Suppose that $\approx_i$ and $\sim_i$ are arbitrary equivalence relations for each i. Consider a model where $w \approx_1 v$ and $w \approx_2 v'$ with $v \neq v'$, and both v and v' are dead-end states (i.e., we only have $v \sim_1 v$ and $v' \sim_2 v'$). Suppose that φ is true at v only

Now this is not just a technical move: "profile gaps" encode something interesting, namely *correlations* between behavior of agents. In a general game model, if player i changes her move, then the only available profiles for this may now be ones where some other player j has changed his move as well. Game theorists have studied correlations extensively: cf. (Aumann 1987; Brandenburger and Friedenberg 2008). But the same notion has come up in logic, since correlations provide "information channels" where the behavior of one agent can carry information about that of another (Barwise and Seligman 1997). And more recently, generalized forms of such dependencies have become the focus of attention in "dependence logics" (Väänänen 2007). In other words, independence may be costly, and the Product Axiom that seemed the pride of STIT may eventually stand in the way, being just an extreme case of a more sophisticated theory of agent behavior.[19]

In the rest of this chapter, we look at extensions of the current framework with features that seem essential to rational agency, and that have been the subject of study in dynamic logics.

5 The Roles of Knowledge

Our connection between STIT and matrix games introduced a notion of knowledge, of agents that have decided, but do not know yet what the others have chosen. Knowledge is not mentioned explicitly in STIT framework, but it seems to be lurking behind the scenes here. In fact, it is present in more than one way: choice and action naturally come with *varieties of knowledge*. Here is how this can happen, even in the simple setting that we have considered.

Consider a one-step action. Before I have made my choice, I only know that one of the available future histories will occur: and in that sense, the STIT tree modality already acts as a form of knowledge about how the whole future can unfold. This knowledge can be significant, since the tree encodes the "protocol" of all possible runs of the current process.

Next, right after I have chosen my action, I know what I am going to do, but I still do not know what the others will do, and this was the sort of knowledge based on personal decisions that was made explicit in the matrix models for games of Sect. 4.

Finally, once both our actions have actually taken place, agents do know what was chosen, if we assume that they observe these actions publicly. Knowledge from observation of events is a major source of information in a temporal world. It is is often encoded in epistemic uncertainty relations between moments of time that are used to model information-driven processes, such as games with imperfect information (cf. Binmore 2009; Parikh and Ramanujam 2003; Fagin et al. 1995). As for its driving

(Footnote 18 continued)
and ψ is true at v' only. Then the antecedent is true, but the consequence is not.

[19] Also relevant to the issue of generalized "profile models" is recent work by Roberto Ciuni on connections between generalized STIT models and notions of effectivity in games, and on actions whose effects are only given probabilistically: Ciuni (2013).

forces, updating knowledge from public observation or more private sources is the key topic in dynamic-epistemic logics (van Benthem 2011).

It is natural to add epistemic operators of all these sorts to logics of decision and action, and in fact, this is happening in logics of games (cf. van Benthem 2014). Many kinds of knowledge relevant to action scenarios are local, having to do with what agents know temporarily as they make a choice. But more global "procedural knowledge" about the future of the process is essential, too, and then the trees of STIT may lose their grip. If I know something about your space of possible strategies, the informational situation will need "STIT forests" rather than trees to distinguish the alternatives (cf. van Benthem et al. 2009). The same complication arises in genuine multi-agent scenarios. One cannot assume that agents know everything about others, and to cope with this variation, again, models have to be complicated beyond the basic STIT format.

Pursuing these matters is beyond the scope of this chaper, but explicit modeling of knowledge seems inescapable in a serious theory of choice and action. For a discussion along these lines, see Pacuit and Simon (2011) for a logical system that merges ideas from STIT and PDL while explicitly representing the agents' knowledge. We see it as one virtue of our linking up STIT and PDL that experiences in the latter area can then be enlisted for the former. Our next section will present a case study, of one particular dynamic epistemic logic with added STIT features.

6 Dynamic Epistemic Logic Meets STIT

Temporal trees with epistemic features may be viewed as a record of actions unfolding over time, while marking local uncertainties (or information) that agents had. If we want to understand the dynamics that gives rise to such a record, we need an account of information update in a temporal universe. A typical system where PDL-style events and knowledge come together is *dynamic epistemic logic* (DEL). We assume the reader is familiar with its basics, and so, we only give the key definitions here (see van Benthem 2011 for more details and motivation).

The basics of DEL update The basic structures are **epistemic models**, tuples $\langle W, \{R_i\}_{i\in I}, V\rangle$ with W a (finite) set of worlds, $R_i \subseteq W \times W$ an equivalence relation, and $V : \mathsf{At} \to \wp(W)$ a valuation function marking at which worlds the atomic propositions in At are true. Over these models the basic language of epistemic logic $\mathcal{L}_{EL}$ can be interpreted, including universal modalities $K_i\varphi$ for "agent i knows that φ. This much is completely standard.

The central idea of *dynamic* epistemic logic is now to describe social interaction, including agents' uncertainty about the events they witness, in so-called **event models**. These are tuples $\mathbf{E} = \langle E, \{S_i\}_{i\in I}, \mathsf{pre}\rangle$ with E a (finite) set of basic events, $S_i \subseteq E \times E$ is an uncertainty relation, and $\mathsf{pre} : E \to \mathcal{L}_{EL}$ assigns to each event $e \in E$ a formula that serves as a **precondition** for that event.

Now, dynamic changes in agents' information can be described by means of *product update* transforming a current pointed[20] epistemic model $\mathcal{M}$ using the event model $\mathbf{E}$. The **product model** $\mathcal{M} \oplus \mathbf{E} = \langle W', \{R'_i\}_{i \in I}, V' \rangle$ is defined as follows:

- $W' = \{(w, e) \in W \times E \mid \mathcal{M}, w \models \mathsf{pre}(e)\}$;
- $(w, e)R'_i(w', e')$ iff wR_iw' and eS_ie'; and
- $(w, e) \in V'(p)$ iff $w \in V(p)$

More precisely, the understanding is that $\mathcal{M}$ has an actual world w, while $\mathbf{E}$ has an actual event e. Product update works for many epistemic scenarios, while it has also been extended to deal with belief and preference change. The language of DEL then adds dynamic modalities $\langle \mathbf{E}, e\rangle\varphi$ that describe at worlds w in $\mathcal{M}$ what is true one step later in the product model with $\mathbf{E}$ and actual event e. The resulting logic of informational events can be axiomatized completely by a compositional technique of 'recursion axioms' analyzing compounds $\langle \mathbf{E}, e\rangle K_i\varphi$ in terms of conditional knowledge that agents had before the update. The details of this are beyond our needs here, but see van Ditmarsch et al. (2007), van Benthem (2011) for more extensive analysis.

Our aim in this section is just to show how, in line with our analysis of Sect. 3, STIT ideas of control fit quite well in this PDL stronghold.

Extending DEL with control A first easy task is adding the earlier control relations for different agents to event models, which just requires adding equivalence relations.[21] Now we can set up a calculus of reasoning. Our dynamic-epistemic language still has its basic event modalities $\langle \mathbf{E}, e\rangle\varphi$, but now we can also introduce a STIT operator

$$\langle \mathbf{E}, e, i\rangle\varphi$$

saying in $\mathcal{M}, w$ that φ is true in all product models $(\mathcal{M}, w) \oplus (\mathbf{E}, f)$ for all events f that are control equivalent to e for agent i. This is formally quite similar to an operator that would already make sense in DEL as it stands, namely, stating the 'observational knowledge' that an agent has acquired after product update with the current event model $\mathbf{E}$.

The complete dynamic logic of this expanded system lies embedded in the base logic of DEL in an obvious manner. Its laws for the new control operator will be essentially those of STIT. But what we obtain in this way is a much richer logic of one-step information flow plus an explicit account of agents' choices of actions where relevant. However, it should be noted that this logic still runs on the usual analysis of DEL's standard dynamic modality.

One crucial feature is that, unlike standard DEL logics, this new system does not have a modality reflecting its dynamic control relations in the static epistemic

[20] A pointed epistemic model is a model with a distinguished state intended to represent the "actual" state of the world.

[21] This may have to be modified when we want some events to just happen without agency. Also, there are problems of intuitive interpretation for control in private-information scenarios, but we ignore these here.

base models $\mathcal{M}$.[22] One might think of this negatively as limiting the logical status of control, reflecting its ephemeral nature. Our own more positive view is that this feature makes event models really come into their own, as carrying crucial information that is *sui generis*.

While our proposal for merging enriches DEL with STIT ideas, what good does it do in the opposite direction? One effect is that we now have a logic that describes both steps in the fork models of Sect. 3, before and after the choice. Thus, it is a logic of choosing and moving ahead, like the NEXT-STIT system of (Broersen, 2011). But the main virtue is that, given the long experience in DEL, our merged system plugs STIT into the world of private versus public information, imperfect information games, and much more.

Dynamifying STIT But the DEL perspective also suggests a more radical move, affecting our view of the scenarios that motivated STIT in the first place. STIT is a logic of deliberate choice and action, but remarkably, it does not analyze any of these activities explicitly, recording only their outcomes.[23] By contrast, the DEL methodology follows a main principle of Logical Dynamics:

Where there is a change, there is an event.

Taking this line, can we 'dynamify' STIT in DEL style? What are the main events that take place in a choice scenario? Here are the main stages as we see them:

deliberation, decision, action, and observation

In a first *deliberation stage*, we analyze our options, and find optimal choices. Next, at the *decision stage*, we make up our mind and choose an action of our own. Then at the *action stage*, everyone acts publicly, and this gets observed, something that we can also model as a separate *observation stage*, though things happen simultaneously.

All these stages can be analyzed using DEL-style models. Perhaps the easiest is the final stage, where an event model will do with all possible events, marking the actual one, and giving agents the right amount of observational powers: totally public in STIT, perhaps more mixed in other settings. But the intermediate stage, too, invites event models. We can have pair events with control relations, as we just introduced in Sect. 6, and then get the matrix models of Sect. 4 as an output. Finally, modeling the initial deliberation stage is more complex, since many factors can weigh in here that are not represented in basic STIT models, such as agents' *preferences* over outcomes. Still, there is a growing body of work on deliberation analyzed in terms of DEL-style updates (van Benthem 2007; Pacuit and Roy 2011), and this might inform an account of deliberation that seems a natural companion to any logic of "deliberate action".

---

[22] This makes our DEL system with control different, e.g., from DEL logics of questions with issue relations: see van Benthem and Minică (2012).

[23] This output orientation on choice is of course precisely the official STIT view of actions.

7 Further Directions

There are many follow-up topics to our analysis, of which we mention three.

The first is the addition of agents' *preferences*. Clearly, this further structure is crucial to the game logics of Sect. 4, and existing modal systems do incorporate preferences in order to define and reason about notions like 'best response', Nash equilibrium, and rational behavior generally (cf. van Benthem 2011, 2014 for such notions analyzed in DEL). In particular, the interplay of actions and preferences has already been studied in the matrix logics of Sect. 4, using techniques from (Liu 2012; van Benthem et al. 2009).[24]

Adding preferences seems a necessity for STIT as well, since rational agency is about *best actions* rather than just any actions, and agents may also prefer ensuring φ rather than ψ for other reasons, including deontic norms. This need not be a simple matter, since best action is not just a matter of, say, finding Pareto-optimal simultaneous choices for all agents. As we know from game theory, more complex deliberation methods are needed, such as iterated removal of dominated actions.[25] All this seems a happy marriage with STIT, and indeed, many of the relevant issues are addressed in Horty's book (Horty 2001).

The next extension would be the study of long-term *temporal evolution*. Our logics so far described single steps in a larger process, but it has long been acknowledged that the proper stage for studying agency is that of a linear- or branching-time temporal logic (Fagin et al. 1995; Parikh and Ramanujam 2003). The same is true for STIT, and one question that seems of interest is whether our one-step event models with control relations can be related systematically to epistemic temporal universes via representation theorems extending those of van Benthem et al. (2009).

Our final topic is *strategic interactive behavior*. We started our presentation of STIT with its basic properties for agents' choices: for each agent, these formed a *partition* of all possible outcomes (call this the Partition property), and also, any two choices for different agents have to overlap. This level of stating constraints is similar to that of representation theorems for games characterizing players' strategic *powers*, forcing the game to end in certain sets of outcomes by playing one of their strategies against any counterplay of the opponent. The latter type of result, however, usually refers to powers in a longer extensive game that can take many individual steps. For instance, van Benthem (2001) characterizes players' powers in finite determined two-player games in terms of three constraints: *Monotonicity* (powers are upward closed), *Consistency* (any two powers of different players overlap), and *Determinacy* (if a set of outcomes is not a power for one of the players, then its complement is a power for the other player). Of these three, Determinacy is typically lost in the

24 Many interesting new problems arise in this area. One is finding a formalization of basic game-theoretic reasoning that makes sense for rational action generally: as initiated in (van Benthem 2007). Another unresolved issue is whether introducing preference structure increases the computational complexity of the modal logic of action, an issue known as the "price of rationality".

25 Thus one might first iteratively prune a given choice situation in this way, and only follow the standard STIT-style format once an equilibrium has been reached.

STIT setting of simultaneous action. Nevertheless, it seems significant that there are extended representation results for players' powers in extensive games with *imperfect information* that require only Monotonicity and the typical STIT constraint of Consistency (cf. again van Benthem 2001).[26]

We end with just one simple observation. What happens to the key STIT constraints when we consider iterated simultaneous action? Most importantly, the crucial property of Partition disappears, and the reason is very instructive. When we make consecutive choices, our available strategies get enriched. In a one-step scenario, agents could only choose one of their actions *ab initio*. But now, they can have strategies letting their next action depend on the observed behavior of the other agents. A standard example of this is the famous strategy *Tit for Tat* in evolutionary game theory: one copies the opponent's preceding move. Hence, the strategies available at the second level do not just consist of choosing an action uniformly, they can depend on the behavior of others. It is easy to see that the disjointness property for sets of outcomes (i.e., the powers matching these strategies) are no longer disjoint.[27] On the other hand, this richer set of strategies does depend crucially on a special feature of the STIT scenario, namely the public observation of everyone's moves. If there were no such observation, then players' could not make their choices dependent on what others have done, and we would get a simple product model of two consecutive actions that does satisfy the Partition condition. Put differently, one-step simultaneous action does not allow for sequential *dependence* of actions, though it may allow for *correlation* as we saw in Sect. 4. But it is precisely the observation feature built into STIT that does make more sophisticated dependent behavior possible as actions get repeated.

8 Conclusion

In this chapter, we have lightly compared the STIT approach to choice and action with that offered by dynamic logic, broadly conceived (including dynamic-epistemic logics). We found that, despite differences in style and presentation, these frameworks are much more congenial than is often thought. Indeed, key ideas from STIT about actions and control merged well with modal logics of games, and in particular, they led to natural dynamic-epistemic logics of information and events that incorporate the crucial STIT notion of *control*. We have only proposed a few such bridges here, without any sustained development, suggesting how ideas might flow across, and further directions pursued. Even so, we hope to have put to rest some views about vast chasms separating STIT and PDL that are sometimes found in the literature.

[26] There is also a literature with more sophisticated representation results that are significant here, of which we mention Bonanno (1992), Pauly (2001) and Goranko et al. (2013).

[27] It is an interesting problem whether some special properties remain for STIT powers. In particular, the temporal logic of Ciuni and Zanardo (2010) seems relevant to analyzing these matters, including the special constraints imposed on STIT models if we do insists on the above properties of powers for single agents, or groups of these: cf. Zanardo (2013).

We are by no means the first to have observed the compatibility of STIT and ideas from the world of PDL and DEL. Notably, Horty articulates many of the idea sketched in this chapter in his important book (Horty 2001). Also, Xu (2010, 2012) are interesting examples of STIT systems that have borrowed notions of action and strategy from the PDL tradition to form richer frameworks for strategic agency. We see our analysis as making a small push in the same direction.

Finally, we recall an earlier point made at the start of our analysis. A paradigm is not just a set of definitions of structures and axioms for reasoning. It is also a belt of applications, in the terminology of Kuhn (1962), a growing family of successful "exemplars". This makes frameworks harder to compare and merge, since their success does not just depend on their formal backbone, but also on the "art of modeling" that has been invested by skilled practitioners. In a practical setting, choices between paradigms may just be choices of taste and life-style, and these of course will not be affected much by theoretical analysis. Still, tastes can at least be diversified—and we hope to have contributed at least to what is on the menu in the logical study of deliberate action.

References

Aumann, R.J. 1987. Correlated equilibrium as an expression of bayesian rationality. *Econometrica* 55: 1–18.

Aumann, R. 1999. Interactive epistemology I: Knowledge. *International Journal of Game Theory* 28: 263–300.

Balbiani, P., A. Herzig, and N. Troquard. 2008. Alternative axiomatics and complexity of deliberative STIT theories. *Journal of Philosophical Logic* 37(4): 387–406.

Barwise, J. and J. Seligman. 1997. *Information flow: The logic of distributed systems*. Cambridge: Cambridge University Press.

Belnap, N., M. Perloff, and M. Xu. 2001. *Facing the future: Agents and choice in our indeterminist world*. Oxford: Oxford University Press.

Binmore, K. 2009. *Rational decisions*. Princeton: Princeton University Press.

Blackburn, P., M. de Rijke, and Y. Venema. 2002. *Modal logic*. Cambridge: Cambridge University Press.

Bonanno, G. 1992. Set-theoretic equivalence of extensive-form games. *International Journal of Game Theory* 20: 429–447.

Brandenburger, A., and A. Friedenberg. 2008. Intrinsic correlations in games. *Journal of Economic Theory* 141(1): 28–67.

Broersen, J., A. Herzig, and N. Troquard. 2006a. Embedding alternating-time temporal logic in strategic STIT logic of agency. *Journal of Logic and Computations* 16(5): 559–578.

Broersen, J., A. Herzig, and N. Troquard. 2006b. From coalition logic to stit. *Electronic Notes Theoretical Computer Science* 157(4): 23–35.

Broersen, J. 2011. Deontic epistemic STIT logic distinguishing modes of mens rea. *Journal of Applied Logic* 9(2): 127–152.

Brown, M.A. 1988. On the logic of ability. *Journal of Philosophical Logic* 17(1): 1–26.

Brown, M.A. 1992. Normal bimodal logics of ability and action. *Studia Logica* 51(3/4): 519–532.
Chellas, B.F. 1992. Time and modality in the logic of agency. *Studia Logica* 51(3/4): 485–518.
Ciuni, R. 2013. Ought implies can, omission, and probabilistic deliberative STIT Lecture at the Institute for Logic, Language and Computation, LIRa Session, 30 March 2013.
Ciuni, R., and A. Zanardo. 2010. Completeness of a branching-time logic with possible choices. *Studia Logica* 96(3): 393–420.
Clarke, E.M., and E.A. Emerson. 1981. Design and synthesis of synchronization skeletons using branching-time temporal logic. In *Logic of programs*, ed. D. Kozen, 52–71. Berlin: Springer.
Clarke, E.M., O. Grumberg, and D.A. Peled. 2000. *Model checking*. Cambridge: The MIT Press.
Fagin, R., J. Halpern, Y. Moses, and M. Vardi. 1995. *Reasoning about knowledge*. Cambridge: The MIT Press.
Gabbay, D., A. Kurucz, F. Wolter, and M. Zakharyaschev. 2003. *Many-dimensional modal logics: Theory and applications*. Amsterdam: Elsevier.
Goranko, V., W. Jamroga, and P. Turrini. 2013. Strategic games and truly playable effectivity functions. *Autonomous Agents and Multi-Agent Systems* 26(2): 288–314.
Halpern, J., and M. Vardi. 1989. The complexity of reasoning about knowledge and time. *Journal of Computer and System Sciences* 38: 195–237.
Harel, D., D. Kozen, and J. Tiuryn. 2000. *Dynamic logic*. Cambridge: The MIT Press.
Herzig, A., and F. Schwarzentruber. 2010. Properties of logics of individual and group agency. In *Proceedings of advances in modal logic*, vol. 7, eds. C. Areces, and R. Goldblatt, 133–149. London: College Publications.
Herzig, A., and E. Lorini. 2010. A dynamic logic of agency I: STIT, capabilities and powers. *Journal of Logic, Language and Information* 19(1): 89–121.
Horty, J. 2001. *Agency and deontic logic*. Oxford: Oxford University Press.
Horty, J.F., and N. Belnap. 1995. The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic* 24(6): 583–644.
Kuhn, T. 1962. *The structure of scientific revolution*. IL: The University of Chicago Press.
Liu, F. 2012. *Reasoning about preference dynamics*. Synthese library. Berlin: Springer.
Lorini, E., F. Schwarzentruber, and A. Herzig. 2009. Epistemic games in modal logic: Joint actions, knowledge and preferences all together. In *LORI-II Workshop on Logic, Rationality and Interaction, Chongqing, China*, eds. X. He, J. Horty, and E. Pacuit, 212–226. Berlin: Springer.
Lorini, E. 2010. A dynamic logic of agency II: Deterministic $\mathcal{DLA}$, coalition logic, and game theory. *Journal of Logic, Language and Information* 19(3): 327–351.
Lorini, E., and F. Schwarzentruber. 2010. A modal logic of epistemic games. *Games* 1(4): 478–526.
Marx, M. 2007. Complexity of modal logic. In *Handbook of modal logic*, eds. P. Blackburn, J. van Benthem, and F. Wolter, 139–189. New York: Elsevier.
Osborne, M., and A. Rubinstein. 1994. *A course in game theory*. Cambridge: The MIT Press.
Pacuit, E., and O. Roy. 2011. A dynamic analysis of interactive rationality. In *LORI, Volume 6953 of lecture notes in computer science*, eds. H.P. van Ditmarsch, J. Lang, and S. Ju, 244–257. Berlin: Springer.
Pacuit, E., and S. Simon. 2011. Reasoning with protocols under imperfect information. *The Review of Symbolic Logic* 4(3): 412–444.
Parikh, R., and R. Ramanujam. 2003. A knowledge based semantics of messages. *Journal of Logic, Language and Information* 12: 453–467.
Pauly, M. 2001. *Logic for social software*. Ph. D. thesis, ILLC University of Amsterdam (Dissertation Series 2001–10).
Reiter, R. 2001. *Knowledge in action: Logical foundations for specifying and implementing dynamic systems*. Cambridge: The MIT Press.
Seligman, J. 2010. Hybrid logic for analyzing games. Lecture at the Workshop Door to Logic, Beijing, May 2010.
Spaan, E. 1990. Nexttime is not necessary. In *Proceedings of TARK*, ed. R. Parikh, 241–256.
Väänänen, J. 2007. *Dependence logic*. Cambridge: Cambridge University Press.

van Benthem, J., and Ştefan Minică. 2012. Toward a dynamic logic of questions. *Journal of Philosophical Logic* 41(4): 633–669.

van Benthem, J., and E. Pacuit. 2006. The tree of knowledge in action: Towards a common perspective. In *Proceedings of Advances in Modal Logic*, vol. 6, eds. G. Governatori, I. Hodkinson, and Y. Venema, 87–106. London: King's College Press.

van Benthem, J., J. Gerbrandy, T. Hoshi, and E. Pacuit. 2009a. Merging frameworks for interaction. *Journal of Philosophical Logic* 38(5): 491–526.

van Benthem, J., S. Ghosh, and F. Liu. 2008. Modelling simultaneous games in dynamic logic. *Synthese* 165(2): 247–268.

van Benthem, J., P. Girard, and O. Roy. 2009b. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic* 38: 83–125.

van Benthem, J., E. Pacuit, and O. Board. 2011. Toward a theory of play: A logical perspective on games and interaction. *Games* 2(1): 52–86.

van Benthem, J. 1997. Modal foundations for predicate logic. *Logic Journal of the IGPL* 5(2): 259–286.

van Benthem, J. 2001. Games in dynamic epistemic logic. *Bulletin of Economic Research* 53(4): 219–248.

van Benthem, J. 2007. Rational dynamics and epistemic logic in games. *International Journal of Game Theory Review* 9(1): 13–45.

van Benthem, J. 2010. *Modal logic for open minds*. Stanford: CSLI Publications.

van Benthem, J. 2011. *Logical dynamics of information and interaction*. Cambridge: Cambridge University Press.

van Benthem, J. 2014. *Logic and games*, manuscript, ILLC, University of Amsterdam. To appear with The MIT Press, Cambridge MA.

van der Hoek, W., and M. Pauly. 2006. Modal logic for games and information. In *Handbook of modal logic: Studies in logic*, vol. 3, eds. P. Blackburn, J. van Benthem, and F. Wolter, 1077–1148. New york: Elsevier.

van Ditmarsch, H., W. van der Hoek, and B. Kooi. 2007. *Dynamic Epistemic Logic*. Synthese library. Berlin: Springer.

Xu, M. 1995. On the basic logic of STIT with a single agent. *Journal of Symbolic Logic* 60(2): 459–483.

Xu, M. 2010. Combinations of stit and actions. *Journal of Logic, Language and Information* 19(4): 485–503.

Xu, M. 2012. Events and actions. *Journal of Philosophical Logic* 41(4): 765–809.

Zanardo, A. 2013. Indistinguishability, choices, and logics of agency. *Studia Logica on-line* 1–22.