On the Formal Structure of Continuous Action

Thomas Müller

ABSTRACT. Analytical investigations of agency are mostly concerned with a description *ex post acto*. However, continuous action (being doing something) needs to be considered as well. The paper shows that while the modal-logical treatment of agency in branching time-based stit theory is currently unable to handle continuous action, the stit framework can be extended such as to handle these cases as well. Our new operator, *istit*, provides for an adequate expression of the notion of being doing something, and we present a simple axiomatisation. In our extended framework, agency, ability, and refraining are linked to an agent's current strategy.

1 Introduction

Since the 1980ies, a number of agency-related concepts have been explored using the resources of modal logic. The key idea, sometimes dubbed the "Anselmian approach" since there is textual evidence for the analysis in some writings of St. Anselm's, is that acting is best described in terms of an agent's bringing about some state of affairs. Thus, the concept of agency is seen to give rise to a family of (agent-indexed) modal operators. A natural reading for these modalities is "α sees to it that ϕ", abbreviated as "$\alpha\, stit : \phi$". The stit-modalities have been given a formally rigorous semantics in the framework of branching time; the approach is laid out and well argued for by Belnap *et al.* in their recent book, *Facing the Future* [3].

The present paper is an attempt at extending that formally rigorous modal-logical treatment of agency to cover the case of continuous action, which present stit theory does not cover. We will present two related, but differently motivated challenges, one action-theoretic, the other linguistic. From these challenges, which point out a striking feature of the phenomenology of continuous action, we will derive conditions of adequacy for formal analysis. While current stit theory does not meet these conditions, we argue that they can be met by adding a new stit modality.

The paper is structured as follows: In section 2, we introduce the stit framework, giving both philosophical motivation and the key formal defini-

Advances in Modal Logic, Volume 5.

tions. In section 3, we present the two mentioned challenges, derive formal conditions of adequacy, and argue that the existing stit framework does not meet these conditions. In section 4, we enrich the growing zoo of stit modalities by introducing our new operator, *istit* ("... is seeing to it that"), and we show that this operator meets the conditions of adequacy. The new operator involves a new index of evaluation for the formal language, s (for "strategy"), which provides a link between formal modeling on the one hand and our everyday mentalistic vocabulary on the other hand. Finally, in section 5 we present an axiomatisation for the new *istit* modality.

2 A brief outline of *stit* theory

2.1 A metaphysical presupposition

How should one set out to analyse agency, or to defend some specific analysis? Questions of methodology always loom large in philosophy, and in understanding agency, some of these questions are crucial. Before considering the phenomenon of continuous action, it is therefore important to spell out a metaphysical presupposition of this paper: We assume that agency presupposes indeterminism. This means that there can be no agency if the future is not open, if it does not contain more than one possible course of events. To some this may seem like a bold assumption, taking sides in the endless debate about compatibilism. A full defence is certainly out of place here. However, the assumption is inevitable from the methodological standpoint of descriptive metaphysics that we adopt. Descriptive metaphysics, so-called by Strawson [11], means that in deciding metaphysical questions (such as the one about determinism or indeterminism), one needs to stick to the conceptual scheme that one is actually using, as shown by a broad range of linguistic, cultural, legal and other practices. Surely a metaphysical notion of an open future is deeply entrenched in all our agency-related concepts, from deciding to attributing responsibility, praise, and blame. Thus, our formal analysis needs to honour the concept of an open future. The framework, which is based on the indeterministic theory of branching time, is therefore the most natural starting point.[1]

2.2 The formalities of *stit*

A structure is an indeterministic branching time model with agents and choices, $\mathfrak{S} = \langle M, \leq, Agents, Choice\rangle$. The substructure $\langle M, \leq\rangle$ is a branching time structure, i.e., M is a nonempty set of moments partially ordered by $\leq$, which ordering is tree-like, so it satisfies the axiom of "no backward

[1] Cf. [4] for a proposed extension to branching space-times.

branching",

$$\forall x \forall y \forall z\, ((x < z \wedge y < z) \rightarrow (x \leq y \vee y \leq x)),$$

and any two moments have a common lower bound:

$$\forall x \forall y \exists z\, (z \leq x \wedge z \leq y).$$

Further postulates are discussed in [3, Chap. 7].[2] Maximal linear subsets of $\mathfrak{S}$ are called *histories*. In $\mathfrak{S}$, branching occurs where histories diverge. At each moment $m \in M$, the set $H_{(m)}$ of histories containing m is partitioned via the equivalence relation $\equiv_m$ of being undivided at m, where $h_1 \equiv_m h_2$ iff $m \in h_1 \cap h_2$ and there is $m' \in h_1 \cap h_2$ such that $m < m'$. Π_m is the respective partition of $H_{(m)}$. If Π_m has more than one element, then at m there is (indeterministic) splitting of histories.

The set *Agents* describes which agents there are in the model. In this paper, we consider a single agent for simplicity only, so the set $Agents = \{\alpha\}$. The function $Choice^{\alpha}$ determines the choices open for agent α at any moment. This information is needed since the metaphysical basis for agency is not just indeterminism, as encoded by Π_m, but agent-related indeterminism. Thus, $Choice^{\alpha}_m$ is a partition of $H_{(m)}$ that describes the set of choices open for agent α at moment m. The partition $Choice^{\alpha}_m$ may be more coarse-grained, but not more fine-grained, than the partition Π_m: An agent has at most as fine a "control" over what will happen as nature's indeterminism allows. If $Choice^{\alpha}_m$ has only the one element $H_{(m)}$, this means that at moment m, agent α has no choice. Metaphysically, the structure $\mathfrak{S}$ is taken to be a formal picture of the ontology. This means that the partitions $Choice^{\alpha}_m$ are given by nature and cannot be changed by the agent.[3]

As we consider a propositional formal language based on stit structures only, a stit model $\mathfrak{M} = \langle \mathfrak{S}, V \rangle$ consists of a stit structure $\mathfrak{S}$ and a valuation V that maps moment-history pairs to subsets of the set of propositional letters. In accord with standard two-dimensional semantics, sentences are evaluated at an index of evaluation, which usually consists of a context of utterance and some more indexes. In one-agent stit theory, the context is taken to be a moment of utterance m_c, and formulae are evaluated additionally at a moment-history pair m, h, where $m \in h$, to allow for standard Prior-Thomason tense operators P ("it was the case that") and F ("it will

[2]Besides in-depth coverage of the formalities of , the book [3] also presents further uses of the framework and gives historical notes as well as an extensive bibliography.

[3]If you wish, you may read an existentialist note into this: According to stit theory, our freedom is not just given, but forced upon us by the way the world is.

be the case that"). The semantics of the weak Occamist future tense operator F is the following:

$$\mathfrak{M}, m_c, m, h \models F\phi \quad \text{iff there is } m' \in h \text{ with } m < m' \text{ s.t. } \mathfrak{M}, m_c, m', h \models \phi.$$

The weak past tense operator P ("it was the case that") employs the mirror image of this, exchanging "$m' < m$" for "$m < m'$". The corresponding strong operators are denoted G ("it is always going to be the case that") and H ("it has always been the case that"). Apart from these two pairs of modal operators for shifting the m parameter, there is also a pair of so-called historical modalities, $[h]$ and $\langle h \rangle$, that shift the h parameter. The semantic clause for the strong operator $[h]$, called "historical necessity" or "settled truth", is:

$$\mathfrak{M}, m_c, m, h \models [h]\phi \quad \text{iff for all } h' \in H_{(m)}, \quad \mathfrak{M}, m_c, m, h' \models \phi.$$

The dual weak operator $\langle h \rangle$ is called "historical possibility". Note that the truth of $F\phi$ depends on the history of evaluation index h, whereas the truth of $P\phi$ depends only on the moment of evaluation m—if $P\phi$ is true at m, h, then so is $[h]P\phi$. Accordingly, operators like P are called *moment-determinate*, while F is called *moment-indeterminate*.

There is a choice of two stit operators available: the "deliberative stit", *dstit*, due to von Kutschera [14], and the "achievement stit", *astit*, due to Belnap and Perloff [2], which is based on extended stit structures.

dstit To start with the simpler of the two, the semantics for *dstit* is as follows:

$\mathfrak{M}, m_c, m, h \models \alpha\, dstit : \phi$ iff

1. for all $h' \in Choice^{\alpha}_{m}(h)$ we have $\mathfrak{M}, m_c, m, h' \models \phi$ and
2. there is a "counter" $h' \in H_{(m)}$ such that $\mathfrak{M}, m_c, m, h' \not\models \phi$.

The first clause is positive: It states that the current choice of α (singled out from $Choice^{\alpha}_{m}$ through the history h in the index of evaluation) secures the truth of ϕ. The negative second clause, on the other hand, excludes those ϕ that are true under any circumstances: nobody sees to it that $2 + 2 = 4$.

Sometimes a simplified version of *dstit* called *cstit* is considered: The semantics for *cstit* employs only the first clause of the *dstit* semantics (cf., e.g., [8]). Note that both *cstit* and *dstit* support the inference from $\alpha\, stit : \phi$ to ϕ (as $h \in Choice^{\alpha}_{m}(h)$).

astit The *dstit* operator considers a current choice that brings about ϕ. The *astit* operator, on the other hand, expresses the idea that, as compared to simultaneous possibilities in other histories, a previous choice of α (or a previous chain of such choices) secured some present truth. In order to express the notion of simultaneous possibilities across histories, *astit* requires an equivalence relation of "same clock time" across histories. Without going into details of the semantics (for which cf. [3, chap. 2]), we note that *astit* also supports the inference from $\alpha\, astit : \phi$ to ϕ. Thus, all the existing stit operators satisfy that inference:

$$\text{If } \mathfrak{M}, m_c, m, h \models \alpha\, stit : \phi, \text{ then also } \mathfrak{M}, m_c, m, h \models \phi.$$

3 Continuous action: Two challenges

3.1 Action theory: Two kinds of examples of agency

In analyses of agency, the examples used point to two differing approaches.[4] Consider the following examples from Davidson's "The logical form of action sentences" [6]:

- Jones buttered the toast in the bathroom with a knife at midnight. (107)
- The doctor removed the patient's appendix. (111)
- Amundsen flew to the north pole. (115)

In *Intention* [1], Anscombe mostly uses examples of the following kind:

- I'm pumping. (38)
- He is replenishing the water-supply. (39)
- She is making tea. (40)

The first difference that one may notice is one of tense: Davison's examples are in the past tense, whereas Anscombe's are in the present tense. Still more importantly, the examples differ with respect to verbal *aspect*: Davidson considers actions in the perfective aspect (from a point of view after the action is finished). Anscombe uses the imperfective aspect (from a point of view while the action is occurring), as marked by the present continuous. Analytical investigations of agency have mostly been concerned with Davidson-type examples. Stit theory is no exception in this respect: the examples considered are usually of the Davidsonian variety. Certainly these examples are important, but a full account of agency needs to consider Anscombe-type examples of continuous action, too. The two classes of examples differ with respect to what may be called their "phenomenology".

[4] This point is made by Thompson [12].

Davidson: When an action is finished, the following account will be adequate for many purposes: First a certain initial state of affairs obtained. Due to the agent's action, some outcome state of affairs obtained later. The agent saw to it that a transition from initial to outcome occurred. Consider apple peeling.[5] *Ex post acto*, the situation seems to be simple enough: First the apple wasn't peeled, then it was peeled, and the agent did it.[6] Stit theory is able to handle many cases of that kind smoothly, e.g., via the *dstit* operator introduced above.

Anscombe: While an agent is acting, the phenomenology is different. E.g., the ongoing action of apple peeling may be described like this: Some time in the past, the agent decided to peel, and from then until the present moment, she chose (when she had a choice) in such a way as to continue peeling. Before the peeling will be finished, the agent will have more choices. In particular, if the agent is acting freely, we need to allow for "dropping out", i.e., not finishing. The later choices need to be made when they are due; they cannot be made now ("no choice before its time"). But still, when the agent is really peeling, there will be "defaults" for the later choices. We may say that the agent is now committed to the future defaults, but still she cannot *choose* these defaults now. Thus, the following points are crucial for continuous action:

1. If an agent is doing something, she has defaults set for her future choices.

2. The future choices are real choices nonetheless: When the time comes, the default is not forced upon the agent.

3. An agent may be truly said to be doing something even though it may turn out later that she didn't finish.

If stit theory is to provide for a general analysis of agency, it must allow for such cases.

3.2 Linguistics: The "imperfective paradox"

As we have just pointed out, action theory has mostly ignored the phenomenon of continuous action, at least as far as formal analysis is concerned.

---

[5] Besides having the advantage of not involving manslaughter in the way most action-theoretic examples (including St. Anselm's!) do, apples have a venerable tradition in practical philosophy, as witnessed, e.g., by Kant [9] and, more recently, by Segerberg [10] and Xu [16].

[6] Certainly a transition from apple not peeled to peeled involves an agent's continuous action in any case. However, *ex post acto* it is often feasible to ignore this as an unimportant detail.

In linguistics, the situation is different: The progressive tense, which is employed to describe continuous action, is so pervasive in English that it could not be ignored for long. Soon after tense logic had been established as a modal-logical framework for analysing tenses, attempts were made to tackle the progressive in a formally rigorous, modal-logical way as well.[7] It was soon pointed out that the phrase "progressive tense" is misleading, as "the progressive is not simply a temporal operator, but a kind of mixed modal-temporal operator" [7, p. 146]. As mentioned above, the distinguishing feature of the progressive is the imperfective verbal aspect.

One main linguistic observation about the imperfective aspect takes up a point that was noted above. The so-called imperfective paradox consists of the observation that a sentence like

(1) Carlos is building a house.

can be true even though, in fact, Carlos does not finish, and no house comes into existence.[8] There are too many instances of this phenomenon to attempt to "explain it away". One might be inclined to argue that if no house comes into existence, that will show that the sentence (1) was uttered inappropriately, but this is not in accord with well-established usage and thus contrary to linguistic methodology. We understand perfectly well what goes on in such cases. In fact we can further differentiate, as cases of not finishing can be due to "external" or "internal" factors. Consider the following sentences:

(2) John was crossing the street when a bus hit him.

(3) God was creating a unicorn, when he changed his mind.

In sentence (2), there is outside intervention: We picture a situation in which John did everything he could be expected to do in order to be crossing the street, but still a bus hit him (unexpectedly).[9] In contrast, the situation pictured in sentence (3) is such that it is entirely up to the agent whether the result will be achieved or not. Note that this is not something exclusive to god: We all perform tasks from which we may drop out, and we all show

[7] Dowty [7, p. 145] points to early work by Montague and Scott, among others.

[8] Our examples are taken from [13], but they have been around in the linguistics literature for longer ([13] gives detailed references).

[9] The way the situation is described already shows that context will play an important role in assessing outside influence: What may count as crossing under normal, lax standards will not count as crossing (but as "running into the street" or some other action) when standards are raised. Cf. the next note.

what we call in our case "weakness of the will" every so often.[10]

Dowty's early and influential analysis of the "imperfective paradox" [7, chap. 3] stays within the modal-logical framework of tense logic, introducing the additional concept of so-called "inertia worlds" to account for the modal implications of the progressive. This addition has been quite controversial; we will discuss it in section 4.3 below. A different reaction to the phenomenon has been to abandon the modal-logical approach in favour of some event-based calculus. A sophisticated framework of that kind has recently been proposed by van Lambalgen and Hamm [13]. We wish to show that it is possible to stick to a simple and transparent modal-logical approach, but this requires doing justice to the "imperfective paradox".

3.3 Common conditions of adequacy

The action-theoretic and the linguistic challenge can be taken together under a single heading: A modal-logical analysis of agency must (a) be able to account for Anscombe-type examples and (b) offer a solution to the "imperfective paradox". In formal terms, this means that we wish to express "α is ϕ-ing" via some modal operator (or combination of operators) $O_\alpha\,\phi$ such that (a) O_α makes intuitive sense as a "progressive operator" and (b) it is possible to have a model $\mathfrak{M}$ and index of evaluation I s.t.

$$\mathfrak{M}, I \models (P\,O_\alpha\,\phi) \;\wedge\; (\neg\phi \wedge \neg P\phi \wedge \neg F\phi),$$

i.e., while it was true that α was ϕ-ing, α never finished, nor will finish. Figure 1 illustrates such a case. Read ϕ as "the apple will be peeled": At m, you may say that earlier (at m') you were peeling the apple, but now, it will not be peeled after all.

3.4 Why current stit theory fails the test of adequacy

As was pointed out at the end of section 2, the existing stit operators support the inference from $\alpha\,stit : \phi$ to ϕ. Thus in particular, for any stit model $\mathfrak{M}$ and all of the three stit operators, *dstit*, *cstit*, and *astit*, we have

$$\mathfrak{M}, m_c, m, h \models P\,\alpha\,stit : \phi \rightarrow P\phi.$$

This means that the existing stit operators fail the test of adequacy in every model. Current stit theory simply does not have the resources for expressing defeasible defaults.

[10] Context plays a much weaker role in assessing those cases.—While discussions of agency often assume that "weakness of the will" is a specially problematic phenomenon, in the light of our formal analysis these cases are simpler than cases of "outside influence", since the latter imply that the agent somehow misrepresented the situation (albeit in a way that is held to be excusable in the given context). Cf. note 14 below for some further remarks.

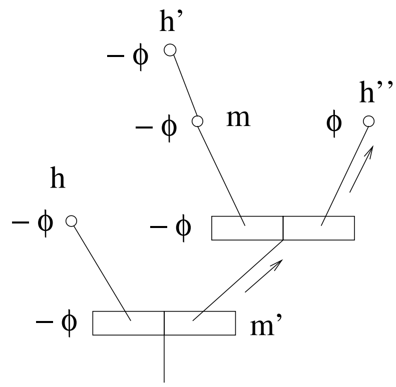

Figure 1. A picture of the conditions for adequacy: The framework needs to allow for models like the one shown, where at m, it was true (namely, at m') that α is ϕ-ing (arrows indicating defaults), but still, at m, ϕ was, is, and will be false. Histories are denoted h, h', and h''.

4 A proposed solution

How should a positive formal account of continuous action look like? In the previous section it was argued that the stit framework so far does not have the necessary resources. Thus there are two options—either to abandon the stit framework or to extend it. In our view, the second option is to be preferred, since the branching time basis of stit theory gives a formally perspicious and well-understood picture of the metaphysical basis of agency. It would be a good thing to stay within that framework. In what follows we argue that it can be done.

4.1 The *istit* operator

The target of analysis is a modal operator that we will call *istit* ("is seeing to it that"), which needs to do justice to the requirements laid down in section 3.3. Thus, we need to be able to affix "defaults" to future choices while preserving their status as genuine choices, i.e., without changing the underlying branching structure. A single element of a partition $Choice^{\alpha}_{m}$ (or of Π_m) at m is not fine-grained enough to describe future defaults—it only carries one a single step, so to speak. On the other hand, specifying a single history is too fine-grained, since an agent cannot normally guarantee a single history—there are usually other sources of indeterminism, be they nature or other agents. This means that the resources of the current stit framework do not allow one to specify sets of histories of the right level of granularity to express future defaults. More resources are needed.

Our proposed solution is to add a new parameter s to the index of evaluation. That new index is to stand for the agent's strategy with respect to which a sentence is to be evaluated.[11] The concept of a strategy is firmly grounded in the framework of branching time.[12] Roughly, a strategy s is a partial function from moments m to subsets of $H_{(m)}$ that are closed under $Choice^{\alpha}_{m}$, and $s(m)$ may be interpreted as "what the strategy advises to choose at m". For simplicity's sake, we will assume that s is total (defined on all of M) and strict (so that $s(m) \in Choice^{\alpha}_{m}$, i.e., s always gives the most specific kind of advice). It follows that s is consistent, i.e., for all $m \in M$, we have $s(m) \neq \emptyset$.

The concept of a strategy thus incorporates the required notion of "default choice" without altering the sets $Choice^{\alpha}_{m}$ (which, as we assume, are given by nature). A strategy s *admits* a history h iff by following the advice of s, history h may be reached,[13] and it admits a history h from moment m on iff h may be reached by following the advice of s starting at m. The set of histories admitted by s from m on is denoted $Admh(s, m)$:

$$h \in Admh(s, m) \quad \text{iff} \quad \text{for all } m' \text{ for which } m < m',\ h \in s(m').$$

This set of histories captures the required level of granularity.

An interesting question is whether for a total, consistent strategy, we always have $Admh(s, m) \neq \emptyset$, as one might expect. The answer in fact depends on a fine point that we have not mentioned before, namely, on the existence of "busy choosers". A busy chooser is an agent who faces infinitely many non-trivial choices in a finite time. Since we wish to study simple cases first, we explicitly exclude busy choosers from consideration. It can then be shown [3, p. 374] that in the absence of busy choosers, $Admh(s, m) \neq \emptyset$ for all $m \in M$.

Based on the new index s, we now define a new pair of modal operators, $[h_s]$ and $\langle h_s \rangle$. The strong operator $[h_s]$, the strategic necessity operator, has the semantics

$$\mathfrak{M}, m_c, m, h, s \models [h_s]\phi \text{ iff for all } h' \in Admh(s, m),\ \mathfrak{M}, m_c, m, h', s \models \phi.$$

[11] In the case of multiple agents, there will be a strategy parameter s_α for each agent α. In what follows, we will use "s" to stand for "s_α", α being the only agent under consideration.

[12] For a detailed introduction that describes various concepts of strategies, cf. [3, chap. 13].

[13] Normally a strategy will not be able to *guarantee* a single history—apart from the indeterminism over which α has control, there is usually also other indeterminism in a model, as mentioned above.

Thus $[h_s]\phi$ is true iff by following the advice of s from m on, the agent can secure the truth of ϕ, no matter how other sources of indeterminism play out. Note that this does not imply that ϕ is historically necessary (even though the inference from $[h]\phi$ to $[h_s]\phi$ is valid). Nor does the truth of $[h_s]\phi$ imply that the agent will in fact be following her strategy. Thus, the truth of $[h_s]\phi$ captures the sought-for notion of defeasible defaults.[14]

Now taking up the idea that a stit operator should contain a positive clause that describes the agent's securing an outcome and a negative clause excluding trivial cases, our semantics for the *istit* operator ("is seeing to it that") is the following:

$\mathfrak{M}, m_c, m, h, s \models \alpha\, istit : \phi$ iff

1. $\mathfrak{M}, m_c, m, h, s \models [h_s]\phi$ and
2. there is a "counter" $h' \in H_{(m)}$ such that $\mathfrak{M}, m_c, m, h', s \not\models \phi$.

Alternatively, we can introduce "$\alpha\, istit : \phi$" as an abbreviation for "$[h_s]\phi \wedge \neg[h]\phi$". Our new operator supports the inference from $\alpha\, istit : \phi$ to $\langle h\rangle\phi$ and to $\langle h\rangle\neg\phi$: If α is ϕ-ing, then the outcome is possible, but not necessary. However, in contrast to the other stit operators, the inference from $\alpha\, istit : \phi$ to ϕ is invalid—Figure 1 provides a countermodel (at index m', h', considering the formula $F\phi$). This model shows that *istit* meets the conditions of adequacy laid out in section 3.3 above. Note that *istit* is moment-determinate: If $\alpha\, istit : \phi$ is true, then so is $[h]\alpha\, istit : \phi$; the history of evaluation index plays no role.

4.2 Status and initialization of the strategy parameter s

In the usual treatment of Prior-Thomason tense logic, there is a difference between the indexes m and h: If a stand-alone sentence is to be evaluated, m

[14] We need to mention a certain complication that arises from the strictness of the semantic clause for $[h_s]$, especially in connection with "outside influence" (cf. notes 9 and 10 above): For $[h_s]$ to be true, α has to have a strategy that guarantees the truth of ϕ no matter how the other sources of indeterminism described in $\mathfrak{M}$ play out. If $\mathfrak{M}$ is taken to be a picture of the real world, we have to concede that perhaps nobody ever has a strategy that really guarantees any non-trivial ϕ—one can almost always cook up a story that would prevent ϕ, and not all of those stories have to involve alien abduction. E.g., if you are peeling an apple you normally disregard a number of real, but very remote possibilities—the knife breaking apart, earthquakes, or other emergencies. As was noted above, what counts as a salient possibility often depends on context. We believe that the best way of dealing with this complication is to follow van Lambalgen and Hamm [13] in taking models not to be mirror images of the real world, but some type of minimal models dependent on some decription of a situation. Spelling all this out in detail must be deferred to a separate paper.—It is for these reasons that "internal influences", i.e., just dropping out of something by one's own free decision, appears to be the structurally simpler case, since it is even possible in cases where $\mathfrak{M}$ is the real world, and no extra epistemic layer allowing for an agent's misrepresentation of a situation is needed.

is assigned an initial value from the context via the moment of utterance m_c, but there is no context-initialization for the history parameter, h: Assuming a "history of the context" would mean falling prey to the myth of "the real future".[15] What is the status of the s parameter?

In the *istit* picture, s functions as an interface between our everyday mentalistic vocabulary that describes an agent's plans and intentions and the formal branching time framework. We talk about an agent's current plans or intentions as something objective (they can play a role in court, for example), and this is what s should capture. Thus, s is an initialized parameter like m, not an uninitialized one like h. Paralleling the treatment of m means that we should add a "current strategy" index s_c to the context.

4.3 Comparison with "inertia worlds"

Dowty [7, pp. 151ff.] gives a reading of the progressive in a branching time framework that is similar to our treatment here. His fundamental notion is that of an "inertia world", or "inertia future". Formally, Dowty singles out a set of histories as "inertia futures", $Inr(m) \subsetneq H_{(m)}$, that are supposed to represent "what is happening *now*, what is the outcome of events as they could be expected to transpire without [...] interference" [7, p. 149]. This concept aims at providing a set of histories at exactly the same level of granularity as our set $Admh(s, m)$. There are two less important differences bewteen Dowty's treatment and *istit*: His semantics is interval-based, and he does not consider a negative clause to exclude trivial cases of "is ϕ-ing". The substantial difference between his approach and ours, which we see as the main advantage of *istit*, is that while *istit* grounds the set $Admh(s, m)$ on the formally rigorous and intuitively clear notion of a strategy, Dowty does not provide further analysis of his concept of an "inertia world" or an "inertia future". In the quote above, he may be interpreted as saying in effect that "an agent is now doing what an agent is doing now", and the inertia approach has consequently been accused of emptyness, or even of incoherence in view of the unclear notion of "intervention". By spelling out "inertia" in terms of the defeasible defaults set by strategies, our analysis goes one crucial step further. Furthermore, our analysis meshes well with the already established stit framework.

4.4 Further extensions

So far we have introduced the bare minimum for handling continuous action: The modal operator *istit* meets the adequacy conditions spelled out in section 3.3. In the extended framework that includes the strategy index, a number of other modalities and agency-related concepts can be defined,

[15] For a decisive criticism of that notion (under the name of "the thin red line"), cf. [3, chap. 6].

showing the flexibility of the framework. We briefly mention three such extensions.

Strategic modalities: In two-dimensional semantics, to each index that is not part of the context there corresponds a number of one-place modal operators for shifting that parameter. E.g., the (weak) tense operators P and F (and their strong duals H and G) shift the moment parameter m backward or forward in time. For the history parameter h there are the corresponding "historical" modalities $\langle h \rangle$ (weak) and dually $[h]$ (strong). Along the same lines one can define "strategic" modalities $\langle s \rangle$.

We call a strategy s' an *alternative* to s at m iff for all $m' \in M$ for which $m \not\leq m'$, $s'(m') = s(m')$; i.e., the two strategies have to agree everywhere outside the future of possibilities of m. The clause for the weak strategic modality $\langle s \rangle$ then reads:

$$\mathfrak{M}, m_c, m, h, s \models \langle s \rangle \phi \quad \text{iff there is } s' \text{ s.t.} \quad \mathfrak{M}, m_c, m, h, s' \models \phi,$$

where s' is a total, strict strategy that is an alternative to s at m.

Ability: The weak strategic modality $\langle s \rangle$ allows one to express the concept of ability: If $\langle s \rangle [h_s] \phi$ is true at m, h, s, then this means that the agent could change her current strategy s such as to guarantee strategically the outcome ϕ. If ϕ is not guaranteed anyway, $\langle s \rangle \alpha\, istit : \phi$ is true, meaning that the agent could be ϕ-ing.

Refraining: Refraining means not doing something while one could do it (cf., e.g., von Wright's analysis [15]). Thus, we may say that α refrains from ϕ-ing iff α is not ϕ-ing ($\neg \alpha\, istit : \phi$) and still could be ϕ-ing ($\langle s \rangle \alpha\, istit : \phi$), so (with "$iref$" the companion "refraining" operator for *istit*)

$$\alpha\, iref : \phi \;\Leftrightarrow_{df}\; \langle s \rangle [\alpha\, istit : \phi] \wedge \neg [\alpha\, istit : \phi].$$

There is a difference between our treatment of refraining and the usual stit analysis: The notion of refraining that corresponds to the other stit operators is itself agentive, i.e., can itself be phrased as a stit sentence:

$$\alpha\, ref : \phi \;\Leftrightarrow_{df}\; [\alpha\, stit : \neg [\alpha\, stit : \phi]].$$

For *dstit* this can be shown to be equivalent to our analysis in terms of ability and not doing [3, p. 438]:

$$\alpha\, ref : \phi \;\Leftrightarrow\; \langle h \rangle [\alpha\, dstit : \phi] \wedge \neg [\alpha\, dstit : \phi].$$

This equivalence does not hold for *istit*.

The crucial question is which of the two analyses is more fundamental. While we do not have a final answer to this question, it appears that the analysis championed by von Wright, which the *istit* framework is able to capture, is the more basic one. It is true that refrainings, or omissions, are generally agentive—in some countries, one can even be legally liable, e.g., for refraining from helping somebody. However, it is rare to speak of countinuous refraining,[16] so there seems to be no complelling reason to analyse "α refrains from ϕ-ing" in terms of "α is ψ-ing". Thus, the non-agentive *istit* analysis of refraining should not count against the theory.[17]

5 Axiomatising *istit*

In this final section we take up the task of axiomatising the new stit modality.

5.1 The semantics

We assume a single agent α who is not a busy chooser; α's strategy will be denoted s. (The agent α will not be mentioned in what follows.)[18] At a moment m, we have to consider the sets of all histories containing m, $H = H_{(m)}$, and the set of histories admitted by s at m, $A = Admh(s, m)$; we have

$$A \subseteq H; \quad A \neq \emptyset.$$

In these terms, the semantic clause for *istit* reads:

$\mathfrak{M}, m_c, m, h, s \models \alpha\, istit : \phi$ iff

1. for all $h' \in A$, $\mathfrak{M}, m_c, m, h', s \models \phi$ and
2. there is $h' \in H$ s.t. $\mathfrak{M}, m_c, m, h', s \not\models \phi$.

For the purpose of axiomatisation, the rest of the language that we are using (including tense operators and so forth) only plays a role by giving rise to formulae that are true at some index m, h, s and false at some other index m, h', s. (The ϕ in the above definitions is treated as a "black box".) Thus the problem of axiomatising *istit* amounts to axiomatising what we will call *set-inclusion structures* $\langle H, A\rangle$, where $A \subseteq H$ and $A \neq \emptyset$. A *set-inclusion model* is such a structure together with a valuation V. We thus consider a simple propositional language L_1 with just two modalities $\Box_H$

---

[16] In conversation, Michael Perloff came up with the example "I was actively refraining from hitting the guy", where the progressive is used for emphasis, stressing how hard it was to refrain.

[17] Certainly the question of continuous refraining merits further study.

[18] A multi-agent extension will have to include assumptions about the independence of simultaneous choices by different agents; the technical side of that will follow the *dstit* analysis presented in [3, chap. 17].

and $\Box_A$, and the index of evaluation is $h \in H$. The semantic clauses for the modalities are

$$\mathfrak{M}, h \models \Box_H \phi \text{ iff for all } h' \in H, \mathfrak{M}, h' \models \phi,$$

$$\mathfrak{M}, h \models \Box_A \phi \text{ iff for all } h' \in A, \mathfrak{M}, h' \models \phi.$$

Here, the modality $\Box_H$ corresponds to $[h]$ above, $\Box_A$ corresponds to $[h_s]$, and $\alpha\, istit : \phi$ is treated as an abbreviation for $\Box_A \phi \wedge \neg \Box_H \phi$. The task of axiomatising *istit* is thus reduced to the task of axiomatising the logic of set-inclusion structures.

5.2 The axiomatisation

It is obvious that by the semantic clause, $\Box_H$ is an S5 modality. The only thing that needs to be investigated is the logic of $\Box_A$ and the interrelation between these two modalities. We proceed by first finding a description of set-inclusion models in the familiar framework of relational Kripke models, employing a language L_2 with two modal operators $\Box_1$ and $\Box_2$ that have the standard Kripke semantics in terms of two relations R_1 and R_2 on a set of worlds W, $w \in W$ being the index of evaluation:

$$\mathfrak{M}, w \models \Box_1 \phi \text{ iff for all } w' \in W \text{ for which } R_1(w, w'), \mathfrak{M}, w' \models \phi,$$

$$\mathfrak{M}, w \models \Box_2 \phi \text{ iff for all } w' \in W \text{ for which } R_2(w, w'), \mathfrak{M}, w' \models \phi.$$

The following Lemma describes the interrelation between the two semantic approaches.

Lemma 1
For every set-inclusion model $\mathfrak{M}_1 = \langle H, A, V \rangle$ ($\emptyset \neq A \subseteq H$), there is an equivalent relational model $\mathfrak{M}_2 = \langle U, R_1, R_2, V' \rangle$, where $U = H$ and

1. *R_1 is an equivalence relation,*
2. *R_2 is serial, transitive, and Euclidean,*
3. *for all $u, u' \in U$, if $R_2(u, u')$, then also $R_1(u, u')$, and*
4. *for all $u, u', u'' \in U$, if $R_2(u, u'')$ and $R_1(u, u')$, then also $R_2(u', u'')$.*

In the other direction, given a relational model as just described, for every point-generated submodel there is an equivalent set-inclusion model.[19]

[19] A point-generated submodel of a relational model $\mathfrak{M}_2 = \langle U, R_1, R_2, V' \rangle$ is a submodel all of whose points can be reached from one point $u \in U$ by following R_1. This ensures that R_1 will be the universal relation on U.

By "equivalent" we mean that for ϕ a L_1-wff and ϕ' the L_2-wff derived from ϕ by replacing $\Box_H$ with $\Box_1$ and $\Box_A$ with $\Box_2$ everywhere, we have

$$\mathfrak{M}_1, w \models \phi \quad \textit{iff} \quad \mathfrak{M}_2, w \models \phi'.$$

Proof: "$\Rightarrow$": Let $\mathfrak{M}_1$ be given. Set $U = H$, $V' = V$, $R_1 = H \times H$ and $R_2 = H \times A$. (1) is immediate, since R_1 is the universal relation on H. For (2), seriality follows from $A \neq \emptyset$, and the other two conditions are also immediate (if $R_2(u, u'')$, then $R_2(u', u'')$ for any u'). (3) follows from $A \subseteq H$, and (4) is verified analogously to (2). The equivalence of the two models follows directly from the translation employed.

"$\Leftarrow$": Let $\mathfrak{M}_2$ be a point-generated submodel of a given relational model. Set $H = \{w' \in U \mid \exists w \in U\, R_1(w, w')\}$, $A = \{w' \in U \mid \exists w \in U\, R_2(w, w')\}$. We need to show that (a) $A \neq \emptyset$ and (b) $H \subseteq W$. Furthermore, if we can show that (c) $H = U$, (d) $R_1 = H \times H$ and (e) $R_2 = H \times A$, we can set $V = V'$, and the equivalence of the two models will follow directly, finishing the proof.

Condition (a) follows from the seriality of R_2. The fact that $\mathfrak{M}_2$ is a generated submodel ensures that R_1, being an equivalence relation, is the universal relation on U, from which we have (c) and (d). Condition (b) then follows from clause (3) above. To establish (e), let $s \in A$, so that by definition there is $w \in H$ s.t. $R_2(w, s)$. Let $w' \in H$. By (d), we have $R_1(w, w')$, and now (4) above gives us indeed $R_2(w', s)$. $\Box$

We now proceed to axiomatise the logic of set-inclusion models by axiomatising the logic of the corresponding relational models. We start with a normal bimodal logic with boxes $\Box_1$ and $\Box_2$ (i.e., propositional logic, the K-axioms for the boxes, and modus ponens, substitution and necessitation). Furthermore, we posit the following axioms for the logic L_{istit}:[20]

T_1 $\Box_1\phi \to \phi$

4_1 $\Box_1\phi \to \Box_1\Box_1\phi$

5_1 $\Diamond_1\phi \to \Box_1\Diamond_1\phi$

D_2 $\Box_2\phi \to \Diamond_2\phi$

4_2 $\Box_2\phi \to \Box_2\Box_2\phi$

5_2 $\Diamond_2\phi \to \Box_2\Diamond_1\phi$

I1 $\Diamond_2\phi \to \Diamond_1\phi$

[20] Thanks to Yuko Murakami for helpful discussions of the axiomatisation.

I2 $\Diamond_2\phi \rightarrow \Box_1\Diamond_2\phi$

$\alpha\, istit : \phi$ is an abbreviation for $\Box_2\phi \wedge \neg\Box_1\phi$.

Here, the first three axioms make $\Box_1$ an S5 modality, the next three make $\Box_2$ a KD45 modality, and the axioms I1 and I2 describe the interaction between the modalities. Note that *istit* itself is not a normal modal operator, even though it is defined from two normal modal operators.

5.3 Soundness and completeness

We can now establish the sought-for result:

Theorem 1
L_{istit} is sound and complete w.r.t the class of set-inclusion structures.

Proof: The soudness of the axioms w.r.t. the set-inclusion structures can be verified immediately from the semantic clauses (I1 corresponds to $A \subseteq H$, while I2 captures the semantic clause for $\Box_A$).

In order to establish completeness, we first show that L_{istit} is complete w.r.t. the class of frames described by clauses (1)–(4) above. Since satisfiability is invariant under generated submodels, Lemma 1 then yields the desired result.

Completeness follows easily from the fact that all the axioms are Sahlqvist formulae. Since the correspondence of the first six axioms to clauses (1) and (2) above is so well known, we only treat the "interaction" axioms $I1$ and $I2$ briefly. From the Sahlqvist-van Benthem algorithm [5, chap. 3.6] we get the following:

I1 The corresponding second order formula in the variable x is
$\forall P((\exists y\,(R_2(x,y) \wedge P(y)) \rightarrow \exists z\,(R_1(x,z) \wedge P(z)))$.
Pulling out the first $\exists$ and instantiating $P(u)$ as $\lambda u.u = y$ gives
$\forall y\,((R_2(x,y) \wedge y = y) \rightarrow \exists z\,(R_1(x,z) \wedge z = y))$,
which is equivalent to clause (3) above.

I2 The corresponding second order formula in the variable x is
$\forall P((\exists y\,(R_2(x,y) \wedge P(y)) \rightarrow \forall z\,(R_1(x,z) \rightarrow \exists z'\,(R_2(z,z') \wedge P(z')))$.
Pulling out the first $\exists$ and instantiating $P(u)$ as $\lambda u.u = y$ gives
$\forall y\,((R_2(x,y) \wedge y = y) \rightarrow \forall z\,(R_1(x,z) \rightarrow \exists z'\,(R_2(z,z') \wedge z' = y))$,
which is equivalent to clause (4) above.

The Sahlqvist completeness theorem [5, chap. 4.3] then establishes the completeness of our axiomatisation. $\Box$

In order to bind this completeness result back to the propositional fragment of the language of *istit* described in section 4, we can use the following construction: Let ϕ be a non-theorem of L_{istit} ($L_{istit} \not\vdash \phi$), so that by Theorem 1 there is a set-inclusion model $\mathfrak{M} = \langle H, A, V \rangle$ and $h^* \in H$ s.t. $\mathfrak{M}, h^* \not\models \phi$. We wish to find a stit model $\mathfrak{M}' = \langle \mathfrak{S}, V' \rangle$ and an index m, h, s s.t. $\mathfrak{M}', m, h, s \not\models \phi$.[21] Recall that a stit structure $\mathfrak{S}$ consists of a branching time structure $\langle M, \leq \rangle$ and the two sets *Agents* and *Choices*. We set

$$M = \{m_0, m_1, m_2\} \cup \{m_h \mid h \in H\},$$

and for the partial ordering $\leq$ we take the reflexive and transitive closure of the following:

$$m_0 < m_1;\ m_0 < m_2;\ m_1 < m_h \text{ for } h \in A;\ m_2 < m_h \text{ for } h \in H - A.$$

Thus, m_0 is a minimal node with two direct successors m_1 and m_2, and the histories are in one-to one correspondence with the elements of H (as the $h \in H$ correspond to maximal elements m_h); we will denote the histories by h as well. We now set $Agents = \{\alpha\}$ and $Choice^{\alpha}_{m_0} = \Pi_{m_0}$, i.e., at m_0, the agent α has two choices corresponding to the two direct successors, m_1 and m_2. For the valuation, we set $V'(m_0, h) = V(h)$. For the strategy s, we set $s(m_0) = \Pi_{m_0}\langle m_1 \rangle$: at m_0, the strategy advises to choose the bundle of histories going through m_1, which corresponds to the set A in the set-inclusion model. It follows that

$$\mathfrak{M}', m_0, h^*, s \not\models \phi.$$

6 Conclusion

The phenomenon of continuous action poses a challenge for formal theories of agency. The branching time-based stit framework in its present form, comprising the operators *astit* and *dstit*, is unable to capture that phenomenon. However, by incorporating an agent's current strategy as a new index of evaluation it is possible to extend the stit framework such as to capture continuous action. The new operator, *istit*, forms the basis of a rich structure of agency-related modalities, and it can be axiomatised in a straightforward way.

[21] We ignore the moment of utterance index m_c, which plays no role for the semantics here.

Acknowledgements

I wish to thank audiences at the Center for the Philosophy of Science, University of Pittsburgh (lunchtime talk, 19 September 2003) and at Manchester (AiML 5, 10 September 2004) for helpful discussions and suggestions. Special thanks are due to Nuel Belnap, Yuko Murakami, Michael Perloff, and to two anonymous referees for AiML 5.

BIBLIOGRAPHY

[1] Anscombe, G.E.M. (1963). *Intention*, 2nd ed. Oxford.
[2] Belnap, N. and Perloff, M. (1988). Seeing to it that: A canonical form for agentives. *Theoria* 54:175–199.
[3] Belnap, N., Perloff, M., and Xu, M. (2001). *Facing the Future*. Oxford.
[4] Belnap, N. (2002). Branching histories approach to indeterminism and free will. Preprint at the Pittsburgh Philosophy of Science Archive, URL=http://philsci-archive.pitt.edu/archive/00000890/index.html.
[5] Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge.
[6] Davidson, D. (1967). The Logical Form of Action Sentences. In his *Essays on Actions and Events*, Oxford 1980, 105–122.
[7] Dowty, D.R. (1979). *Word Meaning and Montague Grammar*. Dordrecht.
[8] Horty, J.F. (2001). *Agency and Deontic Logic*. Oxford.
[9] Kant, I. (1797). *Metaphysische Anfangsgründe der Rechtslehre*. Ed. B. Ludwig. Hamburg ^{2}1998.
[10] Segerberg, K. (1989). Bringing it about. *Journal of Philosophical Logic* 18:327–347.
[11] Strawson, P.F. (1959). *Individuals*. London.
[12] Thompson, M. (2004). *Life and Action*. Cambridge, MA, *to appear*.
[13] van Lambalgen, M. and Hamm, F. (2005). *The Proper Treatment of Events*. Oxford.
[14] von Kutschera, F. (1986). Bewirken. *Erkenntnis* 24:253–281.
[15] von Wright, G.H. (1963). *Norm and Action: A Logical Inquiry*. London.
[16] Xu, M. (1996). *An Investigation into the Logics of Seeing-to-it-that*. Unpublished Dissertation, University of Pittsburgh, Pittsburgh, PA.

Thomas Müller

Philosophisches Seminar, LFB III, University of Bonn,
Lennéstr. 39, 53113 Bonn, Germany

Thomas.Mueller@uni-bonn.de