Annals of Mathematics and Artificial Intelligence 9(1993)25–48 

# In the realm of agents

Nuel Belnap and Michael Perloff
*Department of Philosophy, University of Pittsburgh, Pittsburgh, PA 15260, USA*

**Abstract**

Stit theory (a logic of seeing-to-it-that) is applied to cases involving many agents. First treated are complex nestings of stits involving distinct agents. The discussion is driven by the logical impossibility of "a sees to it that b sees to it that $Q$" in the technical sense, even though that seems to make sense in everyday language. Of special utility are the concepts of "forced choice", of the creation of deontic states, and of probabilities. Second, joint agency, both plain and strict (every participant is essential) is given a rigorous treatment. A central theorem is that strict joint agency is itself agentive.

## 0. Introduction

*Stit* theory has been used to explore selected structural aspects of agency in a rigorous fashion, for example some of the agentive modalities, refraining, and when an agent "could have done otherwise". The goal of this paper is to inspect some structural aspects of *multiple and joint agency*, a task sufficiently complex to give pause to the three inseparables, Aramis, Athos and Porthos.

In other publications, we have described various parts of *stit* theory: a general introduction in Belnap and Perloff [7], some context in Belnap [2], history and pictures in [3], more formal developments in [3,4] and Belnap and Perloff [8], with further comparisons in Perloff [17]. See also Chellas [10] for a critical discussion, and Xu [23] for further studies.

We begin by reviewing the fundamental concepts and techniques of *stit*. In section 2, we apply *stit* theory to other-agent nestings. In section 3, we enrich *stit* theory to accommodate joint agency. In the final section, we look briefly at other-agent joint agentives.

Even minimal progress toward the goal of this paper has required a variety of simplifications: (i) As in *stit* theory generally, we have totally avoided the reification of actions, and (ii) we have minimized reference to intention. (iii) Of relevant notions from earlier papers, we have omitted *stits* that are based on "witness by intervals", and (iv) we have omitted *dstit* (due to von Kutschera and independently to Horty), the agency concept in which the "moment of outcome" is, so to speak, identified with the "moment of choice". In exchange, in this paper we have limited ourselves to ideas and applications that we think work equally well for either *stit* or *dstit*. (v) Of notions not yet discussed in print, we have omitted concepts requiring the notions of "*stit* strategies", and (vi) we do not consider the evident fact that agents interact in space–time.

We refer to Tuomela [21,22] for an alternate methodology that, in contrast to *stit* theory, freely permits one to (i) reify actions and (ii) refer to intentions. Those papers also provide access to some earlier studies of joint agency.

## 1. Review

*Stit* theory begins with the *stit* sentence [$\alpha$ *stit:* $Q$], an approximation to "$\alpha$ sees to it that Q", with the meaning that $Q$ is guaranteed true by a prior choice of the agent $\alpha$. One evaluates [$\alpha$ *stit:* $Q$] in a temporal structure having multiple branches open to the future but only a single route to the past.

The present restatement of the postulates, included primarily for reference, is a mathematically equivalent variant that emphasizes moments and de-emphasizes histories.[1)] This change in emphasis is more convenient for present purposes, but in fact we think it less satisfactory from a conceptual point of view. The notation is that of Xu [23], except that we use, for example, "$\alpha$" ambiguously as sometimes ranging over agents and sometimes as ranging over terms for agents.

We assume a structure $\langle$**T**, $\leq$, **Instant**, **Agent**, **Choice**$\rangle$. **T**, whose members $m$, $w$, etc. are called "*moments*", is *partially ordered* by $\leq$ and subject to *no downward branching* (incomparable moments have no upper bound in T) and *historical connection* (every pair of moments has a lower bound). **History** is the set of all histories $h$, etc., where a *history* is defined as a maximal chain of moments. **Instant** is a "same-time" partition of **T** whose members $i$ are called *instants*, subject to the *unique intersection* condition that each history $h$ intersects each instant $i$ in a unique moment $m_{(i,h)}$, and the *order-preserving* condition that $m_{(i_1,h_1)} \leq m_{(i_2,h_1)}$ iff $m_{(i_1,h_2)} \leq m_{(i_2,h_2)}$. We let $i_{(m)}$ be the instant to which $m$ belongs, and say that all of its members are *co-instantial* with $m$. The order relation is extended in a natural way to instants. For $w < i$, the *horizon at i from w* is the set of all members of $i$ above $w$. We say that $m_1$ and $m_2$ are *undivided at w* if $m_1$ and $m_2$ have a common lower bound that is properly greater than $w$. **Agent** is a nonempty set whose members $\alpha$, $\beta$, etc. are called "agents". **Choice** is a function defined on agents $\alpha$ and moments $w$, yielding as value a partition of all the moments properly greater than $w$. The elements of the partition we call "possible choices for $\alpha$ at $w$". The choice partition must satisfy the *no choice between undivided moments* condition that if $m_1$ and $m_2$ are undivided at $w$, then they belong to the same member of the choice partition. We write $m_1 \equiv_w m_2$ if $m_1$ and $m_2$ are co-instantial moments undivided at $w$, and $m_1 \equiv_w^\alpha m_2$ if $m_1$ and $m_2$ are co-instantial moments belonging to the same possible choice for $\alpha$

1) Thanks to K. Schlechta for pointing out an error in an earlier version of this restatement.

at $w$, so that in part this condition says that $m_1 \equiv_w m_2$ implies that $m_1 \equiv_w^\alpha m_2$. The choice partition must also satisfy *independence of agents*, which can be put as follows: Choose $w$ and $i$, and let $m$ be any function from Agent into the horizon from $w$ at $i$, writing its value as $m_\alpha$. There is a moment $m_0$ such that for every agent $\alpha$, $m_0 \equiv_w^\alpha m_\alpha$. (We have previously used "something happens" or "the world goes on" for alternate statements of this principle. In context, and allowing for the manifest extrinsicness of co-instantiality, the principle says that any combination of choices made by distinct agents at exactly the same moment is consistent. It is evident that "independence of agents" is more apt. There is additional discussion after each of definition 9 and fact 13.)

For the scope of this paper, and only for the sake of expository simplicity, we assume *no busy choice sequences*: Each upper-bounded sequence of non-vacuous choices has a last member, where by a "non-vacuous choice" we mean a moment for which there is an agent such that the choice partition for that agent at that moment is not the vacuous partition. Thus, there are "no busy choosers" in the sense of Belnap [3], which discusses how things have to go if there are any busy choosers. See Xu [23] for an in-depth study.

We adopt from Prior via Thomason the principle that in branching time truth must be seen as relative to moment/history pairs $m/h$, and we say that $Q$ is *settled true* [*false*] at $m$ if it is true [false] at $m/h$ for every history $h$ to which $m$ belongs. In the special case when we can *easily* see that $Q$ is settled one way or the other at $m$, we permit ourselves to say that "$Q$ is true [false] at $m$", without mentioning a history or inserting a "settled".

The truth conditions for a *stit* sentence are as follows:

1. DEFINITION

[$\alpha$ *stit:* $Q$] is true at $m/h$ just in the case there is a choice point $w$ – we call $w$ a "witness to [$\alpha$ *stit:* $Q$] at $m$" – satisfying the following. *Priority*: $w < m$. *Positive condition*: $Q$ is settled true at each $m_1$ such that $m \equiv_w^\alpha m_1$. *Negative condition*: $Q$ is not settled true at some moment – we call it a "counter" – on the horizon from $w$ at $i_{(m)}$.

Since [$\alpha$ *stit:* $Q$] is always either settled true or false at $m$, [$\alpha$ *stit:* $Q$] falls under our special case, so that it is permissible to say that [$\alpha$ *stit:* $Q$] is true [false] at $m$.

The following facts are useful.

2. FACT

*Downward monotony.* $w_1 \leq w_2$ and $m_1 \equiv_{w_2}^\alpha m_2$ imply $m_1 \equiv_{w_1}^\alpha m_2$.

3. FACT

*Witness identity lemma* (Chellas [10]). Suppose that $Q_1$ implies $Q_2$, that $m$, $w_1$, and $w_2$ are moments, and that $\alpha_1$ and $\alpha_2$ are (possibly identical, possibly distinct) agents. Suppose further that $w_1$ is a witness to $[\alpha_1\ stit\text{:}\ Q_1]$ at $m$, and that $w_2$ is a witness to $[\alpha_2\ stit\text{:}\ Q_2]$ at $m$. Then $w_2 \leq w_1$.

4. FACT

*Uniqueness of witness.* As a corollary, whenever $[\alpha\ stit\text{:}\ Q]$ is true at $m$, we may speak unambiguously of *the* witness to $[\alpha\ stit\text{:}\ Q]$ at $m$.

5. FACT

*Second witness lemma.* If $w$ is the witness to $[\alpha\ stit\text{:}\ Q]$ at $m_1$ and if $m_1 \equiv_w^\alpha m_2$, then $w$ is alo the witness to $[\alpha\ stit\text{:}\ Q]$ at $m_2$ – which is therefore settled true at $m_2$.

6. FACT

*A sufficient condition for unsettledness.* For any instant $i$, if each of $[\alpha\ stit\text{:}\ Q]$ and $\sim[\alpha\ stit\text{:}\ Q]$ is true somewhere on the horizon at $i$ from $w$, then there is a moment on the horizon at $i$ from $w$ at which $Q$ is not settled true.

One can use fact 6 as a way of seeing that the following is inconsistent: $[\alpha\ stit\text{:}\ (Q\ \&\ \sim[\alpha\ stit\text{:}\ Q])]$. The Red Duke is subtle, but not so subtle that he could see to it that both M. Bonacieux disappeared but that he, the Cardinal, did not see to that fact. (Of course, he could so act that that is what people would *say*.)

Among the central theses of *stit* theory are (i) that $Q$ is agentive iff it is equivalent to $[\alpha\ stit\text{:}\ Q]$, and (ii) the restricted complement thesis – that a variety of constructions, including imperatival and deontic, must take agentives as their complements. (See Bartha [1] for an extended discussion of agentives as the complement of deontic statements.)

Turning first to imperatives, reflect on the letter from Mme Bonacieux to d'Artagnan:

> Be in Saint-Cloud at ten o'clock tomorrow night, across the street from the bungalow at the corner of Monsieur d'Estrées' house. (1)

Although it might seem that the content of the imperative construction of (1) is a non-agentive describing d'Artagnan's whereabouts, according to the restricted complement thesis this is mere appearance. In truth, the content of (1) is well-regimented by the explicit agentive

[d'Artagnan *stit:* d'Artagnan is in Saint-Cloud . . .]. (2)

And when d'Artagnan's father says of the old yellow horse "Never sell him", the content of his imperative for each moment may appear to be a non-agentive that merely denies agency to d'Artagnan,

~[d'Artagnan *stit:* d'Artagnan sells the yellow horse]. (3)

The restricted complement thesis, however, drives us to take content as agentive:

$$[\text{d'Artagnan } \textit{stit: } {\sim}[\text{d'Artagnan } \textit{stit: } \text{d'Artagnan sells the yellow horse}]]. \quad (4)$$

That is, the father charges d'Artagnan to *deny himself* the agency – to refrain, where according to *stit* theory, refraining from seeing to it that $Q$ is always definable as $[\alpha\, \textit{stit: } {\sim}[\alpha\, \textit{stit: } Q]]$. See, in particular, Belnap and Perloff [7]. In (4) and frequently below, we pursue our policy of writing an agentive $Q$ in the form $[\alpha\, \textit{stit: } Q]$ even though – since $Q$ is agentive – the *stit* is redundant. As explained elsewhere, $[\alpha\, \textit{stit: } Q]$ is intended as a helpful normal form rather than an "analysis" of $Q$.

As for deontic statements, the restricted complement thesis requires that they have one of the following forms:

$\textit{Obligated:}[\alpha\, \textit{stit: } Q]$: $\alpha$ is obligated to see to it that $Q$

$\textit{Forbidden:}[\alpha\, \textit{stit: } Q]$: $\alpha$ is forbidden to see to it that $Q$

$\textit{Permitted:}[\alpha\, \textit{stit: } Q]$: $\alpha$ is permitted to see to it that $Q$

The restricted complement thesis does not affect the standard deontic equivalence

$$\textit{Forbidden:}[\alpha\, \textit{stit: } Q] \leftrightarrow {\sim}\textit{Permitted:}[\alpha\, \textit{stit: } Q],$$

which continues to hold. The standard forbidden/obliged equivalence, however, requires correction.

$$\textit{Forbidden:}[\alpha\, \textit{stit: } Q] \leftrightarrow \textit{Obligated:}[\alpha\, \textit{stit: } {\sim}Q]$$

is in general false (unless either $Q$ or ${\sim}Q$ is agentive for $\alpha$), and

$$\textit{Forbidden:}[\alpha\, \textit{stit: } Q] \leftrightarrow \textit{Obligated: } {\sim}[\alpha\, \textit{stit: } Q]$$

violates the restricted complement thesis because ${\sim}[\alpha\, \textit{stit: } Q]$, though the negation of an agentive, is not itself agentive.

The following equivalences seem likely to be the appropriate ones.

7. CONJECTURE

$$\textit{Forbidden:}[\alpha\, \textit{stit: } Q] \leftrightarrow \textit{Obligated: } [\alpha\, \textit{stit: } {\sim}[\alpha\, \textit{stit: } Q]]$$

$$\textit{Obligated: } [\alpha\, \textit{stit: } Q] \leftrightarrow \textit{Forbidden:}[\alpha\, \textit{stit: } {\sim}[\alpha\, \textit{stit: } Q]].$$

From these, one can easily calculate that

$$\textit{Obligated: } [\alpha\, \textit{stit: } {\sim}[\alpha\, \textit{stit: } {\sim}[\alpha\, \textit{stit: } Q]]] \leftrightarrow \textit{Obligated: } [\alpha\, \textit{stit: } Q].$$

Fortunately, this already follows from the "normality" of the obligation modality together with the equivalence in *stit* theory of the complements (given no busy choice sequences):

$$[\alpha\, stit\!: \sim[\alpha\, stit\!: \sim[\alpha\, stit\!: Q]]] \leftrightarrow [\alpha\, stit\!: Q].$$

## 2. Other-agent nested *stits*

The *stit* construction encourages nesting. Talja [20], extending Lindahl [14], has examined situations describable by *truth-functional* combinations of distinct-agent *stits*. Included are clauses involving their deontic modalizations, but without regard to the restricted complement thesis. In previous papers, we have discussed same-agent nesting. In this section, we work with *other-agent nested agentives*, where, by definition, each is an expression (i) agentive in some agent $\alpha$, and (ii) whose complement is *or contains* a sentence agentive in a distinct agent $\beta$.

We consider the following examples of other-agent nested agentives.

Queen Anne sees to it that d'Artagnan retrieves her diamond tags. (5)

Jussac sees to it that Biscarat surrenders. (6)

Cardinal Richelieu sees to it that M. Bonacieux agrees to spy on his wife. (7)

Kitty sees to it that d'Artagnan seduces her. (8)

Count de Wardes sees to it that Madame Bonacieux does not keep her rendezvous. (9)

Madame Bonacieux sees to it that d'Artagnan does not follow her. (10)

The examples fall into two groups. The difference lies in their complements. In (5)–(8), the complements appear to be agentives, whereas in (9)–(10) the complements appear to be negations of agentives. We treat the groups separately, taking first those with apparently agentive complements.

It might appear that $[\alpha\, stit\!: [\beta\, stit\!: Q]]$ is the appropriate form to represent (5)–(8).[2] Appearances, however, can be misleading. When $\alpha \neq \beta$ then, as Chellas shows, the witness identity lemma and the independence of agents together imply

[2] Obviously, when $\alpha = \beta$, $[\alpha\, stit\!: [\beta\, stit\!: Q]]$ does not involve multiple agents, and indeed is equivalent to $[\alpha\, stit\!: Q]$. Henceforth, we ignore that case.

8. FACT

*[$\alpha$ stit: [$\beta$ stit: Q]] is impossible.*

Since this fact is central to our present concerns, we give a proof.

Assume the following for *reductio*: (a) [$\alpha$ *stit:* [$\beta$ *stit: Q*]] is true at $m_1$ with $w$ as witness, and with $m_2$ a "counter" as required for the negative condition, so that (b) [$\beta$ *stit: Q*] is not settled true at $m_2$. By independence of agents, there must be an $m_3$ such that both (c) $m_1 \equiv^{\alpha}_{w} m_3$ and (d) $m_3 \equiv^{\beta}_{w} m_2$. By (a), (c), and the second witness lemma, it must be that (e) $w$ is witness for [$\alpha$ *stit:* [$\beta$ *stit: Q*]] at $m_3$. By (a) and (c) we must, by the positive condition, have (f) [$\beta$ *stit: Q*] true at $m_3$ – let $w_1$ be the witness for this. From (e), (f), and the witness identity lemma, we infer (g) $w_1 \leq w$. So (d) and (g) imply, by downward monotony, that (h) $m_3 \equiv^{\beta}_{w_1} m_2$. However, then the second witness lemma with (f) and (h) gives that [$\beta$ *stit: Q*] must be settled true at $m_2$, which contradicts (b) and completes the proof.

So [$\alpha$ *stit:* [$\beta$ *stit: Q*]] will not do as a representation of anything consistent. Before further considering (5)–(8), we turn to the examples (9)–(10), whose complements appear to be negations of agentives.

We represent (9) straightforwardly with the form [$\alpha$ *stit:* ~[$\beta$ *stit: Q*]], inasmuch as the Count makes it true by kidnapping Mme Bonacieux. We observe that although kidnapping is unusual, the form [$\alpha$ *stit:* ~[$\beta$ *stit: Q*]], in contrast to the impossible [$\alpha$ *stit:* [$\beta$ *stit: Q*]], depicts a common enough occurrence.

Although (10) appears to be similar to (9), appearances can be deceiving. We cannot represent (10) with [$\alpha$ *stit:* ~[$\beta$ *stit: Q*]]: Mme Bonacieux is in no position to *guarantee* that d'Artagnan fails to follow her.

Observe that among (5)–(10), we have so far provided a representation only for (9). In particular, we have pointed out that neither [$\alpha$ *stit:* [$\beta$ *stit: Q*]] nor [$\alpha$ *stit:* ~[$\beta$ *stit: Q*]] is the appropriate other-agent nested *stit* to represent any of the others. Does this indicate a weakness of stit theory? We think not; we think it indicates a strength. The failures of [$\alpha$ *stit:* [$\beta$ *stit: Q*]] and [$\alpha$ *stit:* ~[$\beta$ *stit: Q*]] encourage us to find more adequate other-agent nested *stits* as interpretations of (5)–(8) and (10). Chellas [10] says that it would be "bizarre to deny that an agent should be able to see to it that another agent sees to something". Our acceptance of this view *for everyday language* is exactly what drives our search. We will look at three interpretations involving other-agent nested *stits*: deontic, disjunctive and probabilistic.

## 2.1. DEONTIC READING OF OTHER-AGENT NESTED STITS

Let us suppose that the facts of (5) are as follows: The Queen calls d'Artagnan to her chambers and says

Retrieve my diamond tags. (11)

D'Artagnan retrieves the tags for the Queen. He succeeds in the task she assigned. Shall we represent the situation by [$\alpha$ *stit:* [$\beta$ *stit: Q*]] (with $\alpha$ = Queen Anne,

$\beta$ = d'Artagnan, and $Q \leftrightarrow$ d'Artagnan retrieves the diamond tags)? No, that would be a mistake. It would be a mistake because, as we have seen, [$\alpha$ *stit:* [$\beta$ *stit:* $Q$]] claims that the Queen guaranteed the truth of [$\beta$ *stit:* $Q$]. Since her imperative does not *guarantee* that d'Artagnan retrieves the diamond tags, we must be more delicate in our analysis. We begin by reminding ourselves of three things. (i) As Austin taught us, issuing an order is doing something with words; it is, as everyone now says, a speech act; (ii) speech acts are acts; and (iii) *stit* encourages us to understand action by asking what it is that an agent sees to. Let us therefore ask, what did the Queen see to when she uttered (11)? Our answer is that she created an obligation. She saw to it, by her pronouncement, that an obligation existed where there was none before. Specifically, she saw to it that d'Artagnan was *obliged* to retrieve the diamond tags.

We mean, incidentally, that really giving an order really does create an obligation, not just that the speaker so intends. In making this distinction with absolute sharpness, we separate ourselves from Searle [18], who appears indifferently to use the language of intention in his "essential condition" but not to use it in his "essential rule" for promises. Having read Hamblin [13], we also reject the claim of Searle and Vanderveken [19] that "the point of orders and commands is to try to get people to do things" (p. 14). Instead, the point of an order is to create an obligation. Nor does advice have the "causing of action" as its point. Aramis is plainly right that "as a general rule, people ask for advice only in order not to follow it; or, if they do follow it, in order to have someone to blame for giving it". What needs telling is the true story of what deontic state agents really see to when they use not only orders and commands, but also advice, requests, invitations, promises, etc., and indeed assertions and questions.

Structural features of *stit* theory accordingly lead us to the following interpretation of (5).

$$[\alpha \textit{ stit: Obligated: } [\beta \textit{ stit: } Q]]. \qquad (12)$$

Such cases are many and important. When in (6) Jussac orders Biscarat to surrender and Biscarat replies "You're my commander and I must obey you", he recognizes that his commander has seen to an obligation. [$\alpha$ *stit:* [$\beta$ *stit:* $Q$]] is not the appropriate reading here because Jussac did not guarantee Biscarat's surrender. What did Jussac accomplish with his order? Jussac saw to the creation of an obligation:

[Jussac *stit: Obligated:* [Biscarat *stit:* Biscarat surrenders]]. (13)

Regarding (10), [$\alpha$ *stit:* ~[$\beta$ *stit:* $Q$]] is not appropriate because when Madame Bonacieux sees to it that d'Artagnan does not follow her, she does not prevent him from following her; rather, she sees to it that a *prohibition* exists where none existed previously. The form

[Madame Bonacieux *stit: Forbidden:* [d'Artagnan *stit:* d'Artagnan follows]] (14) is therefore preferable. When d'Artagnan obeys, he remains agentive; (14) reflects these features of the situation.

To summarize the results of this section, we make a conceptual advance when we represent the character is agentives like (5) and (6) with the deontic form

$$[\alpha\ \textit{stit: Obligated:}\ [\beta\ \textit{stit:}\ Q]]; \tag{15}$$

and an agentive such as (10) with

$$[\alpha\ \textit{stit: Forbidden:}\ [\beta\ \textit{stit:}\ Q]]. \tag{16}$$

One final point before leaving the topic. Notice that if conjecture 7 is correct, then by substitution in the complement, (14) is equivalent to

(Madame Bonacieux *stit: Obligated:* [d'Artagnan *stit:* ~[d'Artagnan *stit:* d'Artagnan follows]]]. (17)

We think this is correct.

## 2.2. DISJUNCTIVE READINGS OF OTHER-AGENT NESTED STITS

Not all other-agent nested agentives can be accurately interpreted by using deontic complements. Some require use of a disjunctive complement. For example, while we cannot use (15) to represent (7), we can use an other-agent nested agentive *with a disjunctive complement*. That is, although the Cardinal does not obligate M Bonacieux to agree to spy on his wife, the Cardinal does see to it that

Either M Bonacieux agrees to spy on his wife
or M Bonacieux returns to his cell. (19)

The normal form

[The Cardinal *stit:* ([M Bonacieux *stit:* M Bonacieux agrees to spy on his wife] ∨ [M Bonacieux *stit:* M Bonacieux returns to his cell])] (19)

helpfully pictures the situation in which a principal agent imposes a "forced choice" on another agent. This observation, that

$$[\alpha\ \textit{stit:}\ ([\beta\ \textit{stit:}\ Q_1] \vee [\beta\ \textit{stit:}\ Q_2])], \tag{20}$$

can successfully represent certain other-agent nested agentives, whereas $[\alpha\ \textit{stit:}\ [\beta\ \textit{stit:}\ Q]]$ always fails, seems to us surprising. See fig. 1. (See previous papers, especially Belnap [4], for help in interpreting the diagrams.)

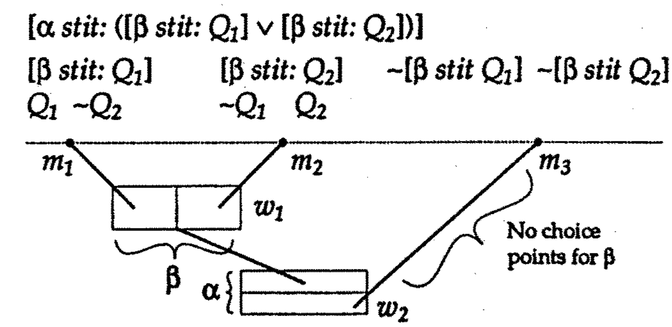


Fig. 1. Simple forced choice.

Subtly different and somewhat more complicated is the following: Suppose that Cardinal Richelieu, determined to entangle M Bonacieux more deeply in his plot, sees to it that the unfortunate draper is forced to choose whether to *put himself* into the situation of forced choice pictured above. For example, Cardinal Richelieu might see to it that M Bonacieux can avoid the forced choice between agreeing to spy on his wife and returning to his cell only by choosing to face the executioner. That is, the Cardinal is sufficiently powerful to arrange matters so M Bonacieux must himself choose between an awful alternative – facing the executioner – and putting himself in a position of forced choice – spying or returning to his cell. When looking at such complicated interactions, it is easy to lose sight of the fact that the principal agent is responsible for seeing to it that the other agent is forced into this terrible predicament. *Stit* theory has the resources to describe the dimensions of the situation in order to help us to understand the relations between the choices of different agents. See fig. 2.[3)]

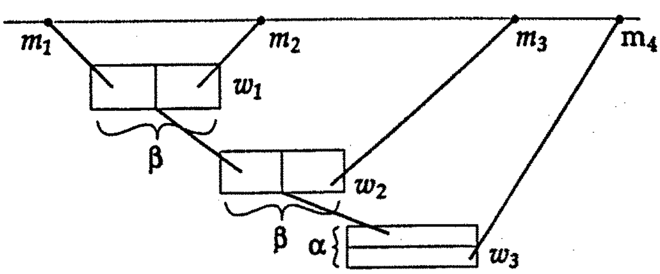


Fig. 2. Complex forced choice.

[3)] Figure 1 approximates the structure of the young man's predicament in Frank Stockton's *The lady or the tiger*. The more complex fig. 2 approximates the structure of Sophie's situation in William Styron's *Sophie's choice*.

### 2.3. PROBABILISTIC READING OF OTHER-AGENT NESTED STITS

The last reading that we shall consider for other-agent nested *stits* is probabilistic. As background, we observe the following: In this world, we seldom *guarantee* the outcomes that everyday language expresses, even when these outcomes are not agentive; this is one of the principal reasons that we think of *stit* as only an approximation to "sees to it that". That is not, however, to suggest that we never guarantee anything! One of the things we can guarantee (we think) is the *high probability* of some outcome expressed in ordinary language. When

Aramis sees to it that Lord de Winter learns of Milady's iniquity, (21)

his choice by no means *guarantees* that de Winter is informed, but the following is fine:

[Aramis *stit:* it is high probable that de Winter learns of Milady's iniquity]. (22)

In designing a language to help us to understand this matter, there are two choices: (i) permanently build the probabilistic element into the *stit* construction, so that *stit* itself indicates high probability instead of guarantee, or instead (ii) represent the idea of probability as a separate linguistic element to be combined with *stit* as wanted. We think (i) should be avoided. In practice it makes if more difficult, not easier, to analyze problems. We recommend (ii). There is a difficulty, however: The concealed double time-reference of *stit* makes it at the least confusing to think through the interactions of probabilities and *stit*. We could reduce confusion by using *dstit* (see (iv) at the beginning of the paper), but not having *dstit* available in this paper, we cannot now profitably carry out the analytical work required. Furthermore, there really is not much sense in localizing probabilities in outcomes of moments. Moments are just too big: One ought to be suspicious of the intelligibility of saying that moments have "outcomes" that might or might not be probable. (We mean to refer to objective probabilities. If the probabilities are "epistemic", then the history of analytic philosphy testifies that anything goes.) Instead, outcomes and therefore probabilities of outcomes should attach to small, local events. (A relativistic foundation for this notion is given in Belnap [6].) For these reasons, we only indicate by an example how we think (ii) applies to other-agent nested agentives.

Consider (8). Surely it is not literally true that Kitty guarantees that d'Artagnan seduces her by the provocative course of behavior she chooses.

[Kitty *stit:* [d'Artagnan *stit:* d'Artagnan seduces Kitty]] (23)

is false. However, the following, which introduces the required element of high probability, is true:

[Kitty *stit:* it is highly probable that [d'Artagnan *dstit:* d'Artagnan seduces Kitty]]. (24)

So we agree with Chellas and common sense that (8) can sometimes be true, and we add that in such cases (24) is often an appropriate reading. For example, in Lady de Winter's lying story to Felton, she does not describe her captor as causing her to seduce him when he *guaranteed* the awful outcome; she was (in her tale) drugged and deprived of choice. If in fact Kitty guarantees that something becomes true of d'Artagnan, then it cannot be that he is agentive in the matter; and if he is agentive, then Kitty can guarantee the probability of the seduction, but not the seduction itself.

## 3. Joint agency: Plain and strict

In the previous section, we explored some other-agent nested constructions with singular agents as subjects. In this section, we will study agentives with joint-agent subjects. We start with English grammar. *Constituent imperatives* (see Belnap [2]) are embedded imperatives, analogous to embedded declaratives or embedded interrogatives. Their content, like that of agentive declaratives, can always be represented by *stit* sentences. An imperative, whether stand-alone or constituent, can have a collective term in subject position, as can an agentive declarative:

M de Tréville announces: "I won't have my musketeers going to low taverns". (25)

The four friends scraped together nine or ten pistoles. (26)

Example (25) might well be taken "distributively", and as analyzable in terms of *stit* sentences with subjects taken to denote a single agent (we call these "singular *stits*"), perhaps the subjects being individual variables bound by a quantifier. On a plausible reading, M de Tréville requires *each* musketeer to see to it that he does not go to low taverns. Examples like (26), however, drive us to widen the grammar of the language of agency. Here, it is evidently the four friends "taken collectively" that succeeded in raising nine ot ten pistoles; it is not something that each of them does. We cannot usefully represent (26) with only singular *stits*. We need to add to our formal grammar of singular *stits* the category of a "joint *stit*".

Collectives can be represented by mereological sums as in Massey [15]; here, we choose to represent collectives by sets.

This choice limits applicability; the proposed apparatus cannot treat cases in which collectives change their membership over time (see Parks [16]), nor cases in which their membership is history-dependent. The limitation is for expository convenience only, and could be removed by using the language $ML^{\nu}$ of Bressan [9], taking the "cases" to be moment/history pairs. In that language, we would represent **Agent** as an absolute concept, so that **Agent**$^{e}$ would by definition be the extensionalization of **Agent**. Collectives of agents would be represented as properties $F$ (possibly extensional, possibly not; possible contingent, possibly not), such that $F \subseteq$ **Agent**$^{e}$ (this makes being a collective of agents a contingent property). Then for such $F$, [$F$ *stit:* $Q$] would be defined as equivalent to the following: [**Agent** $\cap F^{e}$ *stit:* $Q$]. The point is that for every moment/history $m/h$ there is a unique subproperty $\Gamma$ of **Agent** whose bearers are precisely those agents contingently identical at $m/h$ to some individual concept falling under $F$. We can then use $\Gamma$ in order to trace the same group of agents through the vicissitudes of time and history, since it must be a "substance concept"; see Gupta [12] for a discussion of the use of substance concepts as principles of identity.

The formal clothing of our decision is this. (a) We let $\Gamma$, etc. range over nonempty subsets of **Agent**, and (b) we count [$\Gamma$ *stit:* $Q$] as grammatical. Thus, we propose to represent (26) by

[Four friends *stit:* the four friends have nine or ten pistoles],[4) (27)

where here and below, Four friends = {Athos, Porthos, Aramis, d'Artagnan} $\subseteq$ **Agent**.

There is more than one thing to mean by [$\Gamma$ *stit:* $Q$]. First, we may mean that the bearers of $\Gamma$, without any outside help, guarantee that $Q$, on the basis of a prior simultaneous real choice by each of them. There is also a second, stronger, account. In this version, the bearers of $\Gamma$, without any outside help, *and with the essential input of each of them*, guarantee that $Q$. Each account is useful and is worth a notation of its own. Since, however, we can give only one meaning to [$\Gamma$ *stit:* $Q$], we choose the first. Later, we introduce [$\Gamma$ *sstit:* $Q$] as notation for the second account. We postpone to another occasion treatment of cases where $Q$ is best seen as due to sequential efforts of the members of $\Gamma$. Out thought is that one must first be clear on sequential choices by a *single* agent, a topic that we have only touched upon in previous publications. A consequence is that in this paper we will often treat cases that in reality represent sequential choices as if they were simultaneous, provided the sequencing seems not important and the reconstrual as simultaneous seems enlightening.

### 3.1. PLAIN JOINT STITS

The key concept is the extension of choice-equivalence to sets of agents.

### 9. DEFINITION

For $\Gamma$ a nonempty set of agents, we let $m_1 \equiv^{\Gamma}_{w} m_2$ be defined as $\forall\alpha(\alpha \in \Gamma \rightarrow m_1 \equiv^{\alpha}_{w} m_2)$.

Choice-equivalence for a set $\Gamma$ of agents at a moment $w$ is technically easy, but it is conceptually so important that we offer some further words. Let us go back

[4) We give this form for simplicity of illustration. The form ([$\Gamma$ *stit:* $P$] $\vee$ [$\Gamma$ *stit:* $Q$]) seems more apt for (26) than the form of (27), namely, [$\Gamma$ *stit:* $(P \vee Q)$]. They are certainly not equivalent, neither intuitively nor in *stit* theory.

to the idea that at bottom we are representing "possible choices" at a moment. The deepest idea of a possible choice for a *single* agent $\alpha$ at $w$ is contained in its representation as a set of histories. The deepest idea of a possible choice for a *set* $\Gamma$ of agents is also contained in its representation as a set of histories. We obtain one from the other as follows: given a possible choice for each member of $\Gamma$, we define a possible choice for $\Gamma$ as a whole to be the intersection or "combination" of all the individual possible choices. The "independence of agents" condition guarantees that such a combination always exists.

The image we have in mind in due to von Neumann. Let an outcome be dependent on the choices of two agents $\alpha$ and $\beta$. Von Neumann represents this graphically as follows. All the outcomes are arranged in a rectangular grid. Agent $\alpha$ can pick the row and agent $\beta$ can pick the column. What happens is indicated at the intersection of the row picked by $\alpha$ and the column picked by $\beta$. "Independence of agents" just says that some outcome is indicated at each intersection of a row and a column. For example, if there are three rows (choices for $\alpha$) and four columns (choices for $\beta$), then there are twelve possible outcomes for their combined choice.

With the help of the concept of choice-equivalence for sets of agents, we can state the truth conditions for [$\Gamma$ *stit: Q*]. We say that

10. DEFINITION

[$\Gamma$ *stit: Q*] is true at $m/h$ just in the case there is a choice point $w$ – a "witness to [$\Gamma$ *stit: Q*] at $m$" – satisfying the following conditions (compare definition 1): *Agency*: $\emptyset \neq \Gamma \subseteq$ **Agent**. *Priority*: $w < m$. *Positive*: $Q$ is settled true at each $m_1$ such that $m \equiv^{\Gamma}_{w} m_1$. *Negative*: $Q$ is not settled true at some moment – a "counter" – on the horizon from $w$ at $i_{(m)}$.

In parallel with the remark after definition 1, it is permissible to say of [$\Gamma$ *stit: Q*] that it is true [false] at $m$.

If we apply this definition to (27), it tells us that the raising of the pistoles was due to a simultaneous antecedent choice of the four friends. It is by so much a good approximation to (26). Furthermore, it is good logic:

11. FACT

Results or analyses concerning singular agents established without the use of the postulate of the independence of agents also hold for joint agents.

Results not transferring include those expressed by saying that $\alpha \neq \beta$ when these rely on the independence of agents. The point is that the possible choices for $\alpha$ and $\beta$ at $w$ will be independent if $\alpha \neq \beta$, but this is by no means true of the possible choices for $\Gamma_1$ and $\Gamma_2$ when $\Gamma_1 \neq \Gamma_2$. The obvious reason is that nonidentity between the two collectives does not prohibit their having members in common.

By fact 11 we mean, for example, that any implication or nonimplication that holds between singular *stit* formulas with just $\alpha$ also holds between the joint *stit* formulas that result when $\Gamma$ is substituted throughout for $\alpha$. For instance, $[\Gamma\, stit\colon [\Gamma\, stit\colon Q]]$ is equivalent to $[\Gamma\, stit\colon Q]$, and $\sim[\Gamma\, stit\colon Q]$ is not in general agentive in $\Gamma$ (i.e. it is not in general equivalent to $[\Gamma\, stit\colon \sim[\Gamma\, stit\colon Q]]$).

We can use joint agentives to express the independence of agents, provided we have the help of the following version of "ability".

12. DEFINITION

Let $Can_i[\Gamma\, stit\colon Q]$ be true at a moment/history pair $w/h$ iff there is a moment $m$ lying on the horizon from $w$ at $i$ such that $w$ witnesses the truth of $[\Gamma\, stit\colon Q]$ at $m$.

So $Can_i[\Gamma\, stit\colon Q]$ says that $\Gamma$ can see to it that $Q$ is true at $i$. Then we have the following:

13. FACT

Provided $\alpha \neq \beta$, $(Can_i[\alpha\, stit\colon P]\ \&\ Can_i[\beta\, stit\colon Q]) \rightarrow Can_i[\{\alpha, \beta\}\, stit\colon P\ \&\ Q]$. That is, if at $w$ $\alpha$ can see to it that $P$ at $i$ and $\beta$ can see to it that $Q$ at $i$, then at $w$ they can jointly see to the conjunction $P\ \&\ Q$ at $i$.

Some might think that the following is a counterexample.

Porthos can see to it that the pistoles are used to repay a debt.
Athos can see to it that the pistoles are used to purchase meals.
But the pistoles being so few, the conjunction is impossible even with their best joint effort. (28)

If one takes the situation described seriously, however, especially with regard to fixing the time references, one will find that it is impossible. Of course, if Porthos chooses first, then what Athos can see to is not independent of Porthos' choice, and vice versa. However, fix their choices as absolutely simultaneous, as required for our principle of the independence of agents, and fix the "can" not sloppily, but as Austin's all-in, no-holds-barred "can". Suppose that there are only a few pistoles. Award Porthos the ability to see to it that the debt is repaid. You have by so much restricted the power to be ascribed to Athos; there is in this situation nothing Athos can do *by his choice alone* that guarantees that the pistoles are used to purchase meals. For unless you either supply more pistoles or weaken Porthos' ability, you must allow that no matter what choice Athos makes, it is not enough *by itself* to guarantee the availability of the pistoles. Since you have given Porthos the ability to use the pistoles to repay the debt, you have described a situation in which for

Athos to use the pistoles for meals requires the de facto cooperation of Porthos. The second sentence of (28) is therefore not satisfied, so that (28) is not a counterexample to the principle of independence of agents as expressed in fact 13.

We believe that any conceivable counterexample to the principle of fact 13 will be equi-peculiar with the quantum-mechanical phenomenon discovered by Einstein–Podolsky–Rosen, for in fact it would need to have the same form: spatially separated events that are each absolutely indeterministic and perfectly correlated. Ordinary language can easily fool us about this by permitting (normally useful) waffly readings of "can"; here is a place where theory helps.

### 3.2. STRICT JOINT *stits*

However, $[\Gamma\ \textit{stit: } Q]$ still does not tell us all that we may wish to know. For instance, (27) does not imply that each of the four friends was involved. It might have been, for example, that d'Artagnan was not essential in raising the pistoles in the sense that

[The three musketeers *stit:* the four friends have nine or ten pistoles], (29)

where, as everyone knows, d'Artagnan $\notin$ the three musketeers = {Athos, Porthos, Aramis}. Here is the easy fact about $[\Gamma\ \textit{stit: } Q]$ that informs us of this possibility:

14. FACT

Given $\Gamma_1 \subseteq \Gamma_2 \subseteq$ **Agent**: if $[\Gamma_1\ \textit{stit: } Q]$, then $[\Gamma_2\ \textit{stit: } Q]$.

That is, joint *stits* are closed under "weakening" by the addition of further agents.

We need to define some related properties of agents in two versions before we can go further. The first relativizes the concepts to $\Gamma$ and $Q$. The second drops the $\Gamma$, relativizing only to $Q$. The point is to be careful as to which concept is at stake. (We remark that although the terminology to be introduced seems apt in context, one needs to be sensitive to the considerations mentioned in section 3 concerning sequential choices.)

15. DEFINITION

$\alpha$ is *essential for* $[\Gamma\ \textit{stit: } Q] \leftrightarrow ([\Gamma\ \textit{stit: } Q]\ \&\ {\sim}[(\Gamma - \{\alpha\})\ \textit{stit: } Q])$. $\alpha$ is *inessential for* $[\Gamma\ \textit{stit: } Q] \leftrightarrow [(\Gamma - \{\alpha\})\ \textit{stit: } Q]$.

16. DEFINITION

$\alpha$ is *essential* [*inessential, a mere bystander, not a mere bystander*] for $Q \leftrightarrow \exists\Gamma[\Gamma\ \textit{stit: } Q]$ & for *every* [*not all, some*] $\Gamma$ such that $[\Gamma\ \textit{stit: } Q]$, $\alpha$ is essential for $[\Gamma\ \textit{stit: } Q]$.

Thus, (29) says that d'Artagnan is inessential for (27), but (29) does not say that d'Artagnan is a mere bystander for the four friends having nine or ten pistoles.

What then about the idea that $\Gamma$ sees to it that $Q$, with the added provision that each of its members is essential? We think the intuitive concept thus described is rigorously definable (up to an approximation) just by saying that every member of $\Gamma$ is essential for $[\Gamma \textit{ stit: } Q]$ (but without requiring that every member is essential for $Q$). For preference, we adopt an equivalent way of adding that there are no inessential members: $\Gamma$ sees to it that $Q$, but *no proper subset of* $\Gamma$ *does so*:

$$[\Gamma \textit{ stit: } Q] \; \& \; \forall \Gamma_1 (\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow {\sim}[\Gamma_1 \textit{ stit: } Q]). \tag{30}$$

We will soon define "*strictly stit*" (an expression we introduce for joint agency when each of the agents is essential) by just this formula, but first we must face a difficulty: Only the first part of (30) has an agentive form; the second conjunct is instead a *denial* of agency. So the whole may not itself be agentive! The difficulty is, however, easily overcome. In fact, (30) is equivalent to each of the following:

$$[\Gamma \textit{ stit: } (Q \; \& \; \forall \Gamma_1 (\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow {\sim}[\Gamma_1 \textit{ stit: } Q]))], \tag{31}$$

$$[\Gamma \textit{ stit: } ([\Gamma \textit{ stit: } Q] \; \& \; \forall \Gamma_1 (\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow {\sim}[\Gamma_1 \textit{ stit: } Q]))]. \tag{32}$$

The equivalence of (30) and (32) establishes that (30) is agentive in $\Gamma$ in spite of the fact that a conjunct of (30) is a denial of agency. We state this as a

17. THEOREM

The formulas (30), (31), and (32) are mutually equivalent. In other words, where we let

$$\text{NIM} \leftrightarrow \forall \Gamma_1 (\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow {\sim}[\Gamma_1 \textit{ stit: } Q]),$$

the following are equivalent:

$$[\Gamma \textit{ stit: } Q] \; \& \; \text{NIM}, \tag{30}$$

$$[\Gamma \textit{ stit: } (Q \; \& \; \text{NIM})], \tag{31}$$

$$[\Gamma \textit{ stit: } ([\Gamma \textit{ stit: } Q] \; \& \; \text{NIM})]. \tag{32}$$

"NIM" is an acronym for "no inessential members". The proof of theorem 17 goes in a circle as follows:

(30) $\rightarrow$ (31). Suppose (30) true at $m_1/h$ with prior witness $w$, so that in particular ($a_1$) $[\Gamma \textit{ stit: } Q]$ and ($a_2$) NIM are each true at $m_1/h$. From ($a_1$), we have that (b) $Q$ is settled true at all $m_2$ such that $m_1 \equiv_w^\Gamma m_2$, and there is a counter $m_3$ on the horizon from $w$ at $i_{(m_1)}$ at which (c) $Q$ is not settled true. We show that the same witness

and counter will serve for (31). The part about the counter is evident, since if (c) $Q$ is not settled true at $m_3$, then neither is its conjunction with NIM. What we need to show for the positive condition, since we already have (b), is that supposing (d) $m_1 \equiv^{\Gamma}_{w} m_2$, we have (x) NIM is settled true at $m_2$. Suppose for *reductio* that (x) fails, i.e. that (e) $\varnothing \neq \Gamma_1 \subset \Gamma$ and (f) [$\Gamma_1$ *stit:* $Q$] true at $m_2$ with witness $w_1$. By ($a_1$) and (d) and the second witness lemma, we know that [$\Gamma$ *stit:* $Q$] is true at $m_2$ with witness $w$, and therefore by (f) and the witness identity lemma, we know that (g) $w = w_1$ (any two *stits* to $Q$ at $m_2$ must have the same witness). However, $m_1 \equiv^{\Gamma_1}_{w} m_2$ from (d) and (e), and hence (h) $m_1 \equiv^{\Gamma_1}_{w_1} m_2$ by (g). Now (h) with (f) and the second witness lemma puts [$\Gamma_1$ *stit:* $Q$] true at $m_1$. However, ($a_2$) and (e) imply that [$\Gamma_1$ *stit:* $Q$] is false at $m_1$, a contradiction.

(31) → (32). Suppose (a) (31) true at $m_1$ with prior witness $w$. By the positive condition, each of (b) $Q$ and (c) NIM is settled true at every $m_2$, such that $m_1 \equiv^{\Gamma}_{w} m_2$, and by the negative condition there is a counter $m_3$ such that at $m_3/h_3$ for some $h_3$ to which $m_3$ belongs, (d) ($Q$ & NIM) is false, i.e. either ($e_1$) $Q$ is false or ($e_2$) [$\Gamma_1$ *stit:* $Q$] is true, for some nonempty proper subset $\Gamma_1$ of $\Gamma$. We need to establish the positive and negative conditions for (32). The part about the negative condition is easy; since [$\Gamma$ *stit:* $Q$] implies $Q$, (d) implies that ([$\Gamma$ *stit:* $Q$] & NIM) is false at $m_3/h_3$. Choose $m_2$ such that $m_1 \equiv^{\Gamma}_{w} m_2$. We need to show that ([$\Gamma$ *stit:* $Q$] & NIM) is settled true there. Now (c) already tells us that NIM is settled true at $m_2$, and indeed from (b) we know that $Q$ is settled true at every moment $m_{2'}$, such that $m_2 \equiv^{\Gamma}_{w} m_{2'}$, which gives us the positive condition for $w$ to witness [$\Gamma$ *stit:* $Q$] at $m_2$. We therefore are missing only the negative condition for $w$ to witness [$\Gamma$ *stit:* $Q$] at $m_2$, namely, (x) there is a moment on the horizon from $w$ at $i_{(m_2)}$ at which $Q$ is not settled true. In case ($e_1$), we obviously have that (x); we need to show that case ($e_2$) also implies (x). However, this is a consequence of ($e_2$) with (c) and fact 6 (a sufficient condition for unsettledness).

(32) → (30). This is trivial, having the form that [$\Gamma$ *stit:* $P$] implies $P$.

We therefore enter the following definition, where "*sstit*" is to be read "strictly sees to it that", and connotes the absence of inessential members.

18. DEFINITION

[$\Gamma$ *sstit:* $Q$] $\leftrightarrow$ ([$\Gamma$ *stit:* $Q$] & $\forall\Gamma_1(\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow \sim$[$\Gamma_1$ *stit:* $Q$])).

19. DEFINITION

$Q$ is strictly agentive in $\Gamma$ iff $Q \leftrightarrow$ [$\Gamma$ *sstit:* $Q$].

In order to show that [$\Gamma$ *sstit:* $Q$] is strictly agentive in $\Gamma$, which one would certainly expect, we enter the following:

20. LEMMA

$\forall\Gamma_1(\varnothing \neq \Gamma_1 \subset \Gamma \rightarrow \sim[\Gamma_1\ \mathit{stit}\colon [\Gamma\ \mathit{sstit}\colon Q]])$.

For proof, suppose, for *reductio*, that $\varnothing \neq \Gamma_1 \subset \Gamma$ and $[\Gamma_1\ \mathit{stit}\colon [\Gamma\ \mathit{sstit}\colon Q]]$ at $m_1/h$ with witness $w$ and counter $m_3$. Choose any $m_2$ such that $m_1 \equiv_w^{\Gamma_1} m_2$; then both $[\Gamma\ \mathit{stit}\colon Q]$ and NIM are true at $m_2$. Hence, $Q$ is settled true at all such $m_2$, and we have the positive condition for $[\Gamma_1\ \mathit{stit}\colon Q]$ at $m_1$ to be witnessed by $w$. If, then, $Q$ is not settled true everywhere on the horizon from $w$ at $i_{(m_1)}$, we shall have the negative condition as well, and $w$ will witness the truth of $[\Gamma_1\ \mathit{stit}\colon Q]$ at $m_1$, contrary to the truth of NIM there. We obtain the desired unsettledness of $Q$ from the counter at $m_3$ as follows. We know that $[\Gamma\ \mathit{sstit}\colon Q]$ is not settled true at $m_3$, so that either $\sim[\Gamma\ \mathit{stit}\colon Q]$ or $\sim$NIM is true at $m_3$. Since both $[\Gamma\ \mathit{stit}\colon Q]$ and NIM are supposed true at $m_1$, in either case we can use fact 6 to infer that there is a moment on the horizon from $w$ at $i_{(m_1)}$ on which $Q$ is not settled true, as required.

The following is then an easy calculation.

21. FACT

$[\Gamma\ \mathit{sstit}\colon Q]$ is strictly agentive in $\Gamma$, i.e. is equivalent to $[\Gamma\ \mathit{sstit}\colon [\Gamma\ \mathit{sstit}\colon Q]]$.

One direction comes from the fact that quite generally $[\Gamma\ \mathit{sstit}\colon Q]$ implies $Q$. The other direction is a consequence of theorem 17 and lemma 20.

The "S4" property that we just proved of *sstit* does not give us copious information about the behavior of strict seeing to it that; although it is doubtless a beginning, there is much that we do now know.

22. QUESTION

Suppose we treat *sstit* as a modal operator. What illuminating properties does it have? What about its modal interactions with plain *stit*? Et cetera.

### 3.3. APPLICATIONS OF $[\Gamma\ \mathit{stit}\colon Q]$ AND $[\Gamma\ \mathit{sstit}\colon Q]$

We now turn to applications of the distinction between $[\Gamma\ \mathit{stit}\colon Q]$ and $[\Gamma\ \mathit{sstit}\colon Q]$. A plausible hypothesis is that $[\Gamma\ \mathit{stit}\colon Q]$ is not of much use, and that only $[\Gamma\ \mathit{sstit}\colon Q]$ has application. We have come to think this misses the mark. Although $[\Gamma\ \mathit{sstit}\colon Q]$ is sometimes exactly correct, often $[\Gamma\ \mathit{stit}\colon Q]$ is or should be intended. One should keep in mind two quite different contexts: stand-alone agentive declaratives used descriptively, and agentives in their role as complements. First the stand-alone agentive.

> The Queen's ladies-in-waiting brought the Queen's diamond tags from the Louvre to the ball, (33)

where the Queen's ladies-in-waiting = {Mme de Guitaut, Mme de Sablé, Mme de Montbazon, and Mme de Guéménée} $\subseteq$ **Agent**. Does (33) in its normal use imply that *all* of the ladies were involved, so that it should be awarded the form [$\Gamma$ *sstit: Q*], or does it report only something having the plain form [$\Gamma$ *stit: Q*]? We really have no fixed opinion, and we recognize that "conversational implicature" might be at work, but we do think it worth entering our own "intuition": It is consistent with (33) that Mme de Sablé was an inessential lady. If this "intuition" is correct, then only the weaker [$\Gamma$ *stit: Q*] is appropriate. If not, then the stronger [$\Gamma$ *sstit: Q*] is wanted. In either case, the statement (33) is agentive.

Embedding a construction with the same content as (33), however, changes our "intuitions". The weaker reading is then much more plausible. Consider even the truth-functional case of negation:

> The Queen's ladies-in-waiting failed to bring the
> Queen's diamond tags from the Louvre to the ball. (34)

It would seem to us at least misleading to use (34) to describe the situation in which Mme de Sablé alone was inessential.

It is, however, when agentives are complements of deontics or imperatives that we are most struck with the appropriateness of using the plain [$\Gamma$ *stit: Q*] form.

> The Queen sent her ladies-in-waiting to bring
> the diamond tags from the Louvre to the ball. (35)

Presumably, this royal order lays a *joint obligation* on the ladies-in-waiting. What is the content of that joint obligation? Which of the following is correct?

> [The Queen *stit: Obligated:* [the ladies *stit:* the ladies
> bring the diamond tags from the Louvre to the ball]], (36)

or

> [The Queen *stit: Obligated:* [the ladies *sstit:* the ladies
> bring the diamond tags from the Louvre to the ball]]. (37)

Although the matter is uncertain, surely it is plausible that the content of the Queen's order has only the plain form [$\Gamma$ *stit: Q*], so that is is quite consistent with the content of that order that Mme de Sablé should be inessential. If Ann of Austria really wants all her ladies to be involved, she should explicitly say so, using something with the content [$\Gamma$ *sstit: Q*]. By theorem 17, she will have the satisfaction of knowing that the content of her order is indeed agentive.

We think that the outcome is the same for other deontics with joint subjects such as permissions and for prohibitions: although there is no logical reason not to permit or forbid a collective to see to it that *Q* in the strict sense, often the plain sense is more likely to catch what is wanted. For example, Although de Tréville does not forbid the four friends to spend a total in excess of 6000 livres on their equipment for the seige of La Rochelle, he advises them not to do so. (38)

This example describes a prohibition and some advice. By the restricted complement thesis, each should have an agentive complement. The content of the advice is that the friends should refrain from spending more than 6000 livres, i.e. that

The four friends *see to it that* it is false that the four friends *see to it that* the four friends spend in excess of 6000 livres. (39)

It would follow from conjecture 7 that the content of the prohibition (the one never issued by Tréville), when reconstrued as an obligation, is exactly the same as the content of the advice, i.e. (39). Perhaps this logical parallelism is why (38) sounds so eminently intelligible.

So now the question is, with what sort of *stit* should we approximate the *see-to-it-thats* that occur in (39)? Let $\Gamma$ = the four friends, and let $Q \leftrightarrow$ the four friends spend in excess of 6000 livres. Do we want (i) $[\Gamma\, stit{:} \sim[\Gamma\, stit{:}\, Q]]$, or (ii) $[\Gamma\, sstit{:} \sim[\Gamma\, stit{:}\, Q]]$, or (iii) $[\Gamma\, stit{:} \sim[\Gamma\, sstit{:}\, Q]]$, or (iv) $[\Gamma\, sstit{:} \sim[\Gamma\, sstit{:}\, Q]]$? One can use lemma 20 to show that (iv) is equivalent to (iii); and neither is tolerable. Suppose Porthos in his vanity chooses to spend over 6000 livres, and thus alone guarantees the truth of the complement, so that he alone guarantees that the other three friends are inessential. It seems clear that this behavior counts as *not* following Tréville's advice to refrain, so that (39) is false on that story; but the candidate (iii) is true and so cannot be an accurate representation of (39). We are left with (i) and (ii). The latter of course mixes plain and strict *stits*, but in the absence of a more thorough investigation, both logical and conceptual, we ought not say more.

Our last example concerns a permission.

The four friends allowed their servants Planchet, Grimaud, Mousqueton and Bazin to finish the Beaugency wine. (40)

It seems implausible that the content of this permission should be represented by a strict *stit*. Instead,

[The four friends *stit: Permitted:* [The servants *stit:* the servants finish the Beaugency]] (41)

seems a more likely normal form. In this version, the four friends permit that the Beaugency is finished by the choice of Planchet, Grimaud, and Mousqueton alone, Bazin having antecedently gone off to study his theology.

### 3.4. OVERDETERMINATION AND "FREE RIDERS"

Definitions 15 and 16 distinguished relativized and unrelativized notions of essentiality. Because of overdetermination, they do not come to the same thing. The simplest example involves singular *stits* with $\alpha \neq \beta$: [$\alpha$ *stit:* $Q$] and [$\beta$ *stit:* $Q$] can both be true. By fact 3, the witness for the two *stits* at $m$ will have to be the same moment $w$, but there is no reason that $Q$, while satisfying the negative condition, cannot be true both at all $m_1$ such that $m \equiv_w^\alpha m_1$ and at all $m_1$ such that $m \equiv_w^\beta m_1$ It may be that at a certain moment, Bois-Robert makes a real choice that guarantees that Richelieu knows of Buckingham's meeting with the Queen, and that the Marquis de Beautru quite independently makes an equally real choice guranteeing the same thing.

So this case can be represented by the truth of [$\alpha$ *stit:* $Q$], [$\beta$ *stit:* $Q$], and [{$\alpha$, $\beta$} *stit:* $Q$], and the failure of [{$\alpha$, $\beta$} *sstit:* $Q$]. Observe that each of $\alpha$ and $\beta$ are inessential for [{$\alpha$, $\beta$} *sstit:* $Q$], and therefore inessential for $Q$, but that neither is a mere bystander for $Q$. For this reason, we think it would be incorrect to describe either Bois-Robert or de Beautru as a "free rider" even though each is inessential. Only mere bystanders should be called "free riders".

The following point to the need for further work. (i) There is the statement NMB that $\Gamma$ contains no mere bystander for $Q$: $\forall\alpha((\alpha \in \Gamma) \rightarrow \exists\Gamma_1((\alpha \in \Gamma_1)$ & [$\Gamma_1$ *stit:* $Q$])). Evidently, NIM $\rightarrow$ NMB, but not conversely. Thus, the proposition that ([$\Gamma$ *stit:* $Q$] & NMB) stands as follows: [$\Gamma$ *sstit:* $Q$] $\rightarrow$ ([$\Gamma$ *stit:* $Q$] & NMB) $\rightarrow$ [$\Gamma$ *stit:* $Q$]. (ii) There is the statement OMB that outside of $\Gamma$ there are only mere bystanders for $Q$. The proposition ([$\Gamma$ *sstit:* $Q$] & OMB) says that $\Gamma$ is the one and only joint agent for $Q$; it is evidently not agentive (in the sense of *stit*). It is, however, something that could be seen to.

## 4. Other-agent nested joint *stits*

It is evident that the investigations of section 2 on other-agent nested *stits* and section 3 on joint *stits* need to be combined. In this area, there is much to be considered. Here, we offer only a single illustration, which is that the apparatus developed can distinguish the content of the following in an illuminating way:

> The four friends required of one of Planchet and Fourreau that *he* see to it that Brisement has a proper burial. (42)

> The four friends required of Planchet and Fourreau that *they* see to it that Brisement has a proper burial. (43)

> The four friends required of Planchet and Fourreau that *one of them* see to it that Brisement has a proper burial. (44)

Letting $\alpha$ = Planchet, $\beta$ = Fourreau, and $Q \leftrightarrow$ Brisement has a proper burial, the normal forms are, respectively,

$$\text{[The four friends } \textit{stit: } (\textit{Obligated: } [\alpha \textit{ stit: } Q] \vee \textit{Obligated: } [\beta \textit{ stit: } Q])], \quad (45)$$

$$\text{[The four friends } \textit{stit: Obligated: } [\{\alpha, \beta\} \textit{ stit: } Q]], \quad (46)$$

and

$$\text{[The four friends } \textit{stit: Obligated: } [\{\alpha, \beta\} \textit{ stit: } ([\alpha \textit{ stit: } Q] \vee [\beta \textit{ stit: } Q])]]. \quad (47)$$

In (44), the obligation is jointly on Planchet and Fourreau as a pair, but the execution is supposed to be by one of them as an individual. This complex content, so subtly different from that of (42) and (43), can be clearly expressed by an other-agent nested joint *stit* as in (47).

The lesson is that deontic logic needs other-agent nested joint agentives; and it therefore needs to include a theory with at least the expressive power of joint *stits*.

## References


[1] P. Bartha, Conditional obligation, deontic paradoxes, and the logic of agency, this issue, Ann. of Math. and AI 9(1993)1–23.

[2] N. Belnap, Declaratives are not enough, Philos. Studies 59(1990)1–30.

[3] N. Belnap, Before refraining: concepts for agency, Erkenntnis 34(1991)137–169.

[4] N. Belnap, Backwards and forwards in the modal logic of agency, Philos. Phenomen. Res. 51(1991) 777–807.

[5] N. Belnap, Agents in branching time, unpublished manuscript, Department of Philosophy, University of Pittsburgh (1991). Forthcoming in: *Logic and Reason. Essays in Pure and Applied Logic, in Memory of Arthur Prior*, ed. J. Copeland (Oxford University Press, Oxford).

[6] N. Belnap, Branching space–time, Synthese 92(1992)385–434.

[7] N. Belnap and M. Perloff, Seeing to it that: a canonical form for agentives, Theoria 54(1988) 175–199. Corrected version in: *Knowledge Representation and Defeasible Reasoning*, ed. H.E. Kyburg, Jr., R.P. Loui and G.N. Carlson, Studies in Cognitive Systems, Vol. 5 (Kluwer, Dordrecht, Boston, London) pp. 167–190.

[8] N. Belnap and M. Perloff, The way of the agent, Studia Logica 51(1992)463–484.

[9] A. Bressan, *A General Interpreted Modal Calculus* (Yale University Press, New Haven, 1972).

[10] B.F. Chellas, Time and modality in the logic of agency, Studia Logica 51(1992)485–518.

[11] D. Gabbay and G. Guenthner (eds.), *Handbook of Philosophical Logic*, Vol. II: *Extensions of Classical Logic*, Synthese Library, Studies in Epistemology, Vol. 165 (Reidel, Dordrecht, 1984).

[12] A. Gupta, *The Logic of Common Nouns: An Investigation in Quatified Modal Logic* (Yale University Press, New Haven, 1980).

[13] C.L. Hamblin, *Imperatives* (Basil Blackwell, Oxford and New York, 1987).

[14] L. Lindahl, *Position and Change: A Study in Law and Logic* (Reidel, Dordrecht, 1977).

[15] G.J. Massey, Tom, Dick, and Harry, and all the king's men, Amer. Philos. Quart. 13(1976)89–107.

[16] Z. Parks, Classes and change, J. Philos. Logic 1(1972)162–169.

[17] M. Perloff, *Stit* and the language of agency, Synthese 86(1991)379–408.

[18] J.R. Searle, What is a speech act?, in: *Philosophy in America* (Allen and Unwin, London, 1965) pp. 221–239. Reprinted in: *The Philosophy of Language*, ed. J.R. Searle, Oxford Readings in Philosophy (Oxford University Press, 1971) pp. 39–53.

[19] J.R. Searle and D. Vanderveken, *Foundations of Illocutionary Logic* (Cambridge University Press, Cambridge, UK, 1985).
[20] J. Talja, A technical note on Lars Lindahl's *Position and change*, J. Philos. Logic 9(1980)167–183.
[21] R. Tuomela, Collective action, supervenience, and constitution, Synthese 80(1989)243–266.
[22] R. Tuomela, Actions by collectives, Philos. Perspectives 3(1989)471–496.
[23] M. Xu, Refraining formulas and busy choosers, unpublished manuscript, Department of Philosophy, University of Pittsburgh (April 1991).