A Reduction of Doxastic Logic to Action Logic

Heinrich Wansing
Dresden University of Technology
Institute of Philosophy
D-01062 Dresden
Germany

Personal preface

Knowledge is in flux. To a very large extent, we all owe a better understanding of what this may mean to Peter Gärdenfors. But not only knowledge is in flux; also people fluctuate from time to time. I first met Peter more than a decade ago, when I was a PhD student of David Pearce and our Berlin group working on problems of knowledge representation within philosophy was looking for ties to other similar groups. The outcome were most interesting mutual visits in Lund and Berlin. As a postdoc in Hamburg, I missed Peter and his group when they were establishing links with Christopher Habel and I was moving to Leipzig. However, a Swedish-German project on belief revision coordinated by Sven-Ove Hansson and Hans Rott renewed my Swedish connections. As a result I was able to establish and keep scientific contact with several colleagues in Sweden, and I am particularly happy to be able to contribute to the present Festschrift for Peter Gärdenfors. Happy Birthday, Peter!

1 Introduction

One of the main problems in doxastic logic is to define a notion of belief that is not closed under logical consequence. In Fagin's and Halpern's (1988) logic of general awareness this failure of 'logical omniscience' is enforced by (i) stipulating a function that assigns to every agent α and possible world w the set of formulas α is 'aware of' at w and (ii) the requirement that being aware of A at w is a necessary condition for α to explicitly believe that A at w. This approach is known to be equivalent with Rantala's (1982a, 1982b) non-normal worlds semantics (see (Wansing 1989, 1990), (Pearce and Wansing

1989), (Thijsse 1992), (Thijsse and Wansing 1995)). At a non-normal world the truth-conditions of a formula need not be recursively defined.

In the present paper, a semantic analysis of belief ascriptions is suggested in terms of belief acquisition and belief abandonment. This analysis is inspired by a variant of a certain epistemological position, namely *doxastic voluntarism*. According to this version of doxastic voluntarism, belief acquisitions are concrete actions. Developing the semantics of belief ascriptions from the point of view of doxastic voluntarism will avoid closure of belief under logical consequence without stipulating devices like awareness functions and non-normal worlds. The semantics is such that it satisfies the following constraint:

(*) 'α believes that A' implies that at an earlier moment, α voluntarily acquired the belief that A.

This reduction of epistemic logic to action logic is, however, not a reduction in the precise sense of faithfully embedding a logical system S_1 into another system S_2 under a translation τ such that a formula A in the language of S_1 is provable in S_1 if and only if (iff) its translation $\tau(A)$ in the language of S_2 is provable in S_2. Although the reduction to be suggested can be presented as a translation of belief ascriptions to formulas of a certain temporal modal logic of agency, the problem is that there seems to be no standard system of epistemic logic *neither* affected by closure under logical consequence *nor* employing devices like awareness functions or non-normal possible worlds. Therefore, since a system S_1 is not given, the suggested reduction is to be understood in the sense of reformulating epistemic logic as a kind of temporal action logic.

The general idea of the reduction is as follows: 'α believes that A' is true at a moment m with respect to a certain history h passing through m iff there is an earlier moment m' such that at (m', h), α voluntarily acquires the belief that A and there is no moment m'' in-between m' and m (including m and m') such that at (m'', h), α gives up the belief that A. For this reduction to make sense, one obviously needs a clear understanding of what it means to voluntarily acquire and give up a belief. A semantic representation of ascriptions of voluntary belief acquisition and abandonment will be given in a variant of the seeing-to-it-that (stit-) theory of agency put forward by Belnap, Perloff, and Xu (see, for instance, (Belnap 1991), (Belnap and Perloff 1988), (Belnap, Perloff, and Xu 1996), (Horty and Belnap 1995), (Xu 1998)). Agentive sentences will be represented using the deliberative-stit operator introduced by von Kutschera (1986) and Horty (1989). It will

be assumed that concrete actions involve choices of agents and, moreover, that genuine choices require alternatives. In the literature on epistemology, however, there is considerable controversy about whether belief formation is in fact a matter of choice. Whereas doxastic voluntarists hold that it is possible to acquire certain beliefs at will, several anti-voluntaristic philosophers have claimed that for conceptual reasons deciding to believe is impossible. Therefore, first of all it is appropriate to address the dispute about doxastic voluntarism. In Section 2, various versions of doxastic voluntarism are identified and two alleged refutations of doxastic voluntarism are dealt with. In Section 3, structures for the deliberative-stit operator are presented as models for a variant of doxastic voluntarism. A semantic analysis of voluntary belief ascriptions based on these models is suggested as a reduction of doxastic logic to the logic of agency.

2 Doxastic voluntarism

Usually in philosophy, beliefs are thought of as psychological states of doxastic and epistemic subjects. Every belief has a content, and the content of a belief is a proposition. Acquiring the belief that A can be understood as entering a certain psychological state whose content is the proposition expressed by A, whereas abandoning the belief that A may be regarded as leaving a psychological state with a content expressed by A. Doxastic subjects take the contents of their beliefs to be true. While thinking of beliefs as psychological states is maybe the predominant view, there is also a tradition according to which beliefs are dispositions. After distinguishing different versions of the voluntaristic claim, in the present section two prominent anti-voluntaristic arguments will be reconsidered. The first has been suggested by Williams (1973), the second by Bennett (1991). In the first argument beliefs are assumed to be psychological states; in the second beliefs are taken to be dispositions.

2.1 Variants of doxastic voluntarism

Certain philosophers hold that it is conceptually impossible to acquire a belief at will. Moreover, these philosophers often claim that perceptions directly induce beliefs without any mediation by an act of will and that in general, belief acquisition is something passive that just happens to a doxastic subject. Pojman (1985, p. 40), for example, claims that "[a]cquiring a belief is a happening in which the world forces itself upon a subject." But

3

what exactly do the voluntarists claim? There are at least the following six different readings of the voluntaristic thesis.[1]

1. It is possible that one voluntarily acquires arbitrary beliefs in full consciousness.[2] (Universal possibilistic voluntarism)

2. It is possible that one voluntarily acquires some beliefs in full consciousness. (Existential possibilistic voluntarism)

3.1 For all beliefs one acquires it holds true that one voluntarily acquires these beliefs. (Universal weak factual voluntarism)

3.2 For all beliefs one acquires it holds true that one voluntarily acquires these beliefs in full consciousness. (Universal strong factual voluntarism)

4.1 For some beliefs one acquires it holds true that one voluntarily acquires these beliefs. (Existential weak factual voluntarism)

4.2 For some beliefs one acquires it holds true that one voluntarily acquires these beliefs in full consciousness. (Existential strong factual voluntarism)

Whereas the possibilistic voluntarists claim that belief acquisition at will is possible; defenders of factual voluntarism claim of already acquired beliefs that they have been acquired at will. Possibilistic voluntarism comes with the assumption that the content of beliefs is submitted to a doxastic subject, who deliberatively acquires or discards the beliefs in question or not. Since in such cases the doxastic subject is aware of the fact that it is in a choice situation, distinguishing between voluntary belief acquisition simpliciter and voluntary belief acquisition in full consciousness (of making a decision) does not make sense. For factual voluntarism this distinction does make sense and leads to differentiating between weak and strong factual voluntarism.

The thesis of universal possibilistic voluntarism (1.) is not a serious variant of voluntarism and is explicitly rejected by James (1896). For James, the possibility of voluntarily acquiring beliefs in full consciousness pertains only to belief contents which are alive for the doxastic subject and which are an alternative in the context of an inescapabe choice, where, moreover, choosing this alternative is an irreversible, unique opportunity. However, not every proposition is vivid enough to play a role in forced and momentous options. The anti-voluntaristic arguments to be dealt with are directed

against *existential* possibilistic voluntarism (2.). This version of voluntarism follows from universal strong factual voluntarism (i.e. 2. follows from 3.2) if we assume that every doxastic subject has in fact acquired some beliefs (or that we are restricting our attention to such doxastic subjects). Obviously, also 4.2 follows from 3.2, i.e. existential strong factual voluntarism follows from universal strong factual voluntarism. Moreover, 3.1 follows from 3.2, and 4.1 is derivable from 4.2.

The models to be presented in Section 3 are models of universal weak factual voluntarism (3.1). This variant of voluntarism seems to be the strongest version of voluntarism for which one may expect *semantical* models for ascriptions of voluntary belief acquisition. If a belief is acquired in full consciousness, it seems unlikely that this feature of belief acquisition can be explicated without reference to pragmatic parameters like intentions. The aim is to come up with general models allowing ascriptions of voluntary belief acquisition to be interpreted, even if the beliefs in question are not acquired in full consciousness. Models of universal strong factual voluntarism (3.2) will then have to be obtained by imposing additional pragmatic conditions. Although every concrete action is voluntary insofar as it requires choices of the agents involved, the agents are not always conscious of a choice when they choose; they just act. As Hoyler (1983, 275 f.) explains, "[u]nconscious choices are certainly an expression of the will and we should surely come to a distorted view of human agency (the will) and human responsibility if we ignored them." Voluntarily acquiring a belief means "that there are ways in which our beliefs could be different as a direct result of our own agency."

2.2   Beliefs as psychological states

Whereas universal possibilistic voluntarism is too strong a position to be plausible, Williams claims that already the weaker claim (2.) is inconsistent. In (Williams 1973, p. 148) he argues as follows:

> If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of mine, i.e. something I take to be true, and also know that I acquired it at will. With regard to

no belief could I know - or, if all this is to be done in full consciousness, even
suspect - that I had acquired it at will. But if I can acquire beliefs at will, I
must know that I am able to do this; and could I know that I was capable
of this feat, if with regard to every feat of this kind which I had performed
I necessarily had to believe that it had not taken place?

In this paragraph, the following sound argument can be identified:

(1) (Assumption) It is impossible that one takes the content of a belief to
be true and knows that one can voluntarily acquire this belief.

(2) (Assumption) If it is possible that one voluntarily acquires a belief,
then one knows that one can voluntarily acquire this belief.

(3) (Consequence of (1)) If one knows that one can voluntarily acquire a
belief, then one does not take the content of this belief to be true.

(4) (Consequence of (2) and (3)) If it is possible that one voluntarily ac-
quires a belief, then one does not take the content of this belief to be
true.

(5) (Consequence of (4)) If one takes the content of a belief to be true,
then it is impossible that one voluntarily acquires this belief.

Since beliefs are taken to be true, the anti-voluntaristic claim follows from
(5).

Assumption (2) is a reformulation of the claim that "if I can acquire
beliefs at will, I must know that I am able to do this". The succedent of
assumption (1) modifies "since I could not then, in full consciousness, regard
this as a belief of mine, i.e. something I take to be true, and also know
that I acquired it at will". The modification amounts to understanding
"know that I acquired it at will" as "know that I can acquire it at will" in
order to obtain a valid argument. Although the argument is indeed valid,
it is nevertheless inconvincing. Assumption (2) is unproblematic, because
the argument is directed against existential possibilistic voluntarism and
therefore full consciousness is assumed anyway. The negation of existential
possibilistic voluntarism thus follows from assumption (1). However, it is
also the case that assumption (1) follows from the negation of existential
possibilistic voluntarism. If assumption (1) is untrue, then it is possible that
one takes the content of a belief to be true and, since knowledge implies
truth, that one can voluntarily acquire this belief. In other words, it is

possible to acquire certain beliefs at will. Represented in this way, the argument is an instance of a *petitio principii* and fails to refute doxastic voluntarism.[3]

2.3   Beliefs as dispositions

The anti-voluntarists might hope to find a refutation of voluntarism by conceiving of beliefs as dispositions, namely as functions from intentions of a doxastic subject together with other beliefs of the subject to actions of the subject. This idea is considered by Bennett (1991). He argues as follows:

(1) (Assumption) Having a belief means having a certain disposition to act.

(2) (Assumption) Dispositions supervene on categorial, non-dispositional states.

(3) (Consequence of (2)) A disposition can be acquired only by changing a subvenient categorial state.

(4) (Consequence of (1) and (3)) Belief acquisition requires changing a subvenient categorial state of a doxastic subject.

(5) (Consequence of (4)) A belief cannot be acquired immediately by an act of the will (but only mediately by changing a subvenient categorial state).

Bennett himself, however, concedes that the argument is unacceptable, and his rejection of the argument takes the form of a reductio ad absurdum. It seems plausible to assume that if someone intends to perform an action of a certain type, then necessarily the person's intention is to some extent abstract. Not every detail of the concrete realization of the action type in question can be represented in the intention to perform an action of that type. Therefore, every performance of a generic action supervenes on the performance of a more specific generic action. In other words, a concrete action of a certain type can be performed only as a subvenient performance of a more specific action type. But if this is so, we cannot immediately perform an action of any type, which is, of course, absurd. Moreover, if the subvenient state to be changed in order to acquire a belief is a psychological state of some sort, the anti-voluntarist must show that there exists no one-one correspondence between such states of a subject and the subject's beliefs
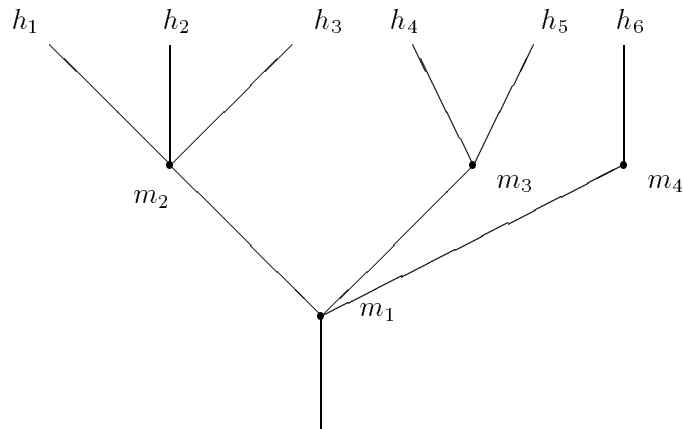
Figure 1: A branching tree of moments of time.

as dispositions, and that the states in question cannot be changed at will. The argument envisaged by Bennett really is insufficient.

3   Deciding to believe

The models for belief ascriptions to be presented have been put forward by von Kutschera (1986) and Horty (1989); they are simplifications of the models for the achievement-stit operator investigated by Belnap, Perloff and Xu (see, for example, (Belnap 1991), (Belnap and Perloff 1988), (Belnap, Perloff, and Xu 1996)). These models are based on the idea of trees of moments of time branching towards the future. Whereas forward branching reflects that the future is open, the requirement that there is no backward branching reflects the determinateness of the past. Linearly ordered sets of moments in such a tree are called histories if they are maximal, i.e. if they are not contained in any larger linearly ordered set of moments in the tree. Intuitively, a history can be thought of as a complete possible development of the world. Since it is assumed that the future is open, moments are typically such that more than just one history passes through them. Figure 1 depicts part of a tree with moments m_1 to m_4 and histories h_1 to h_6.

More formally, a *branching temporal frame* is a structure $\langle T, \leq \rangle$, where T is a non-empty set of moments, and $\leq$ is a partial order on T satisfying *historical connectedness* ($\forall m_1 \forall m_2 \exists m (m \leq m_1 \wedge m \leq m_2)$) and *no backward*
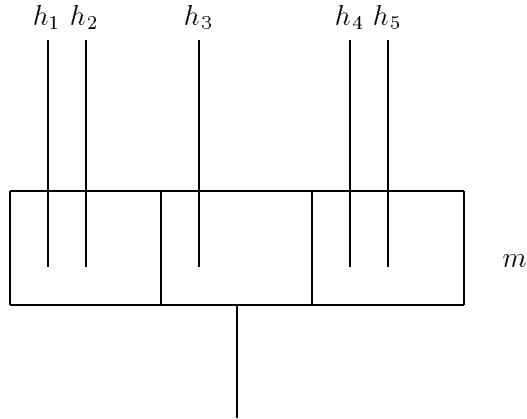
8

Figure 2: Choice cells of α at moment m.

branching $(\forall m \forall m_1 \forall m_2 ((m_1 \leq m \wedge m_2 \leq m) \supset (m_1 \leq m_2 \vee m_2 \leq m_1)))$. A *history in T* is a maximal set of moments (in T) linearly ordered by $<$, where $m < m'$ iff $m \leq m'$ and $m \neq m'$. H_m (the set of histories passing through moment m) is defined as $\{h \mid h$ is history and $m \in h\}$.

There are reasons for evaluating formulas not just in moments of a tree, but in moment/history pairs (m, h). Suppose that in the structure depicted in Figure 1, formula A is true at moment m_4 and untrue at every other moment. Then at moment m_1 nothing decisive can be said about whether A is true sometimes in the future. At m_1 A is true sometimes in the future with respect to h_6 but untrue with respect to histories h_1 to h_5. Therefore, in order to be able to evaluate tense logical formulas $\langle F \rangle A$ ('sometimes in the future A'), *every* formula is evaluated at moment/history pairs. We now assume that our doxastic subjects are also agents who can influence the future course of events by their actions. For this purpose it is assumed that for every agent α the histories passing through a moment m are partitioned into sets histories choice-equivalent for α at m. The idea is that at m, α cannot distinguish by her or his actions between histories that are choice-equivalent for α at m. These 'choice cells' represent the options available to α at m. A natural requirement then is that for every agent α, histories dividing at a moment later than m are choice-equivalent for α at m. The partition of a moment into sets of histories choice-equivalent for an agent α can also be nicely graphically represented, see Figure 2.

If $\langle \mathrm{T}, \leq \rangle$ is a branching temporal frame, then $\langle \mathrm{T}, \leq, Agent, Choice \rangle$ is called a *dstit frame*, if *Agent* is a nonempty set (of agents) and *Choice* is a function mapping every agent/moment pair (α, m) to a partition of H_m (the histories *choice-equivalent for α at m*) satisfying *no choice between undivided histories* $(\forall H \in Choice(\alpha, m))\forall h \forall h'[(h \in H \wedge \exists m'(m < m' \wedge m' \in h \cap h')) \supset h' \in H]$. If $h \in H_m$, then $Choice_\alpha^m(h)$ is the particular choice in $Choice(\alpha, m)$ containing h. A *dstit model* is a structure $\langle T, \leq, Agent, Choice, v \rangle$, where $\langle T, \leq, Agent, Choice \rangle$ is a dstit frame, and v is a valuation function that sends atomic formulas to sets of moment/history pairs. The truth definition for formulas $[\alpha \; dstit: A]$ ("α deliberatively sees to it that A") is as follows:

Definition 1. $[\alpha \; dstit: A]$ is true in a dstit model $\langle T, \leq, Agent, Choice, v \rangle$ at a moment/history pair (m, h) $((m, h) \models [\alpha \; dstit: A])$ iff (i) $\forall h' \in Choice_\alpha^m(h)$, A is true at (m, h'), and (ii) $\exists h'$ such that $m \in h'$ and A is untrue at (m, h').

In this definition, (i) is called the positive condition and (ii) the negative condition. If the negative condition fails, Q is said to be *settled true* at m. Note that it cannot be true that one (deliberatively) sees to it that something is settled true, i.e. "α (deliberatively) sees to it that it is settled true that α does Q" is false at every moment/history pair. "It is settled true that Q" is non-agentive.

In stit-theory (Belnap, Perloff and Xu 1996), a sentence A is considered to be agentive in agent α if A is equivalent to 'α sees to it that A'. If 'α voluntarily acquires the belief that A' is an agentive sentence, it thus ought to be equivalent to 'α deliberatively sees to it that α voluntarily acquires the belief that A'. But what can be said about the semantics of ascriptions of voluntary *belief acquisition* in particular? Very little. What *is* clear is that we can acquire untrue beliefs. If we acquire untrue beliefs, we are mistaken or, in other words, subject to error.[4] The idea now is to introduce for every agent $\alpha \in Agent$ a sentence letter E_α. Intuitively, E_α is true at a moment/history pair (m, h) iff α is mistaken at (m, h). My suggestion (see also (Wansing 1999)) is to understand 'α voluntarily acquires the belief that A' as

$$[\alpha \; dstit: ((\neg A \supset E_\alpha) \wedge (A \supset \neg E_\alpha))],$$

in other words, given a classical understanding of negation and implication, α deliberatively sees to it that either A or α is mistaken.[5] Note that I do not claim that $((\neg A \supset E_\alpha) \wedge (A \supset \neg E_\alpha))$ is an adequate representation of 'α believes that A'. In order to satisfy the above constraint (*), the truth

conditions of 'α voluntarily acquires the belief that A' ought to be part of the truth conditions of 'α believes that A'. The idea is to define belief in terms of belief acquisition and abandonment. If one assumes that every doxastic subject always has some untrue beliefs, introducing constants E_α is not enough, since then, intuitively, E_α is true at every moment/history pair. A possible refinement consists in introducing for every agent α and every formula A a propositional constant $E_{\alpha,A}$ to be understood as 'α is mistaken with respect to A'. If 'α voluntarily acquires the belief that A' is abbreviated as $[\alpha\ vab:\ A]$, then we obtain the following definition:

Definition 2. $[\alpha\ vab:\ A]$ is true in a dstit model $\langle T, \leq, Agent, Choice, v\rangle$ at a moment/history pair (m, h) iff (i) (positive condition) $\forall h' \in Choice_\alpha^m(h)$, $(m, h') \models ((\neg A \supset E_{\alpha,A}) \wedge (A \supset \neg E_{\alpha,A}))$ and (ii) (negative condition) $\exists h'$ such that $m \in h'$ and not $(m, h') \models ((\neg A \supset E_{\alpha,A}) \wedge (A \supset \neg E_{\alpha,A}))$.

If ascriptions of voluntary belief acquisition are interpreted according to the above definition, dstit models can be viewed as models of universal weak factual voluntarism. Given this interpretation of belief acquisition, a truth definition for belief ascriptions can be formulated satisfying constraint (*). The suggestion I would like to make is that with respect to a certain history, α believes that A at the moment m iff there exists a moment m' earlier than m such that at m', α acquires the belief that A and since then α has not given up the belief that A. In addition to a precise notion of voluntary belief acquisition we therefore also need a concept of voluntary belief abandonment.

Definition 3. $[\alpha\ gub:\ A]$ ("α voluntarily gives up the belief that A") is true in a dstit model $\langle T, \leq, Agent, Choice, v\rangle$ at a moment/history pair (m, h) iff $(m, h) \models [\alpha\ dstit:\ \neg((\neg A \supset E_{\alpha,A}) \wedge (A \supset \neg E_{\alpha,A}))]$.

The truth definition for belief ascriptions then takes the following form:

Definition 4. $[\alpha\ bel:\ A]$ ("α believes that A") is true in a dstit model $\langle T, \leq, Agent, Choice, v\rangle$ at a moment/history pair (m, h) iff $\exists m' \in T$ such that $m' \leq m$ and $(m', h) \models [\alpha\ vab:\ A]$ and, moreover, $\neg \exists m'' \in T$ such that $m' \leq m'' \leq m$ and $(m'', h) \models [\alpha\ gub:\ A]$.

If negation and implication (and conjunction) are interpreted *classically*, then giving up the belief that A amounts to acquiring the belief that $\neg A$, since the following formula schemata are pairwise logically equivalent:

$$\neg((\neg A \supset B) \wedge (A \supset \neg B)), \ \neg(A \overset{\circ}{\vee} B), \ (\neg A \overset{\circ}{\vee} B), \ (A \equiv B),$$

where '$\overset{\circ}{\vee}$' denotes exclusive disjunction. This results in a semantics for consistent belief not closed under logical consequence.

Definition 5. $[\alpha\ cbel{:}\ A]$ ("α consistently believes that A") is true in a dstit model $\langle T, \leq, Agent, Choice, v\rangle$ at a moment/history pair (m, h) iff $\exists m' \in T$ such that $m' \leq m$ and $(m', h) \models [\alpha\ vab{:}\ A]$ and, moreover, $\neg\exists m'' \in T$ such that $m' \leq m'' \leq m$ and $(m'', h) \models [\alpha\ vab{:}\ \neg A]$.

Closure under logical consequence fails due to the negative condition in the definition of '$[\alpha\ vab{:}\ A]$'. But this condition is not an *ad hoc* constraint for the sake of avoiding closure under logical consequence. It is just a special case of the negative condition in the truth definition for the dstit operator and therefore a completely natural consequence of adopting the point of view of doxastic voluntarism. If $\perp$ denotes the constantly false proposition and $\top$ the constantly true proposition, we obtain:

$$[\alpha\ vab{:}\ \perp] \quad \text{iff} \quad [\alpha\ dstit{:}\ E_{\alpha, \perp}]$$
$$[\alpha\ vab{:}\ \top] \quad \text{iff} \quad [\alpha\ dstit{:}\ \neg E_{\alpha, \top}]$$

To illustrate Definition 5, Figure 3 exhibits a moment m and a history h passing through m such that at (m, h) the agent α under consideration believes that A, since at (m_2, h), α acquires the belief that A, and neither at (m, h) nor at (m_1, h) or (m_2, h), does α acquire the belief that $\neg A$. The depicted formulas are supposed to be true at the associated moment/history pairs.

Note also that defining belief in terms of belief acquisition and belief abandonment as suggested above allows one to draw various quite subtle distinctions, for instance between giving up the belief that A and refraining from believing that A:

$[\alpha\ gub{:}\ A]$	α gives up the belief that A
$[\alpha\ dstit{:}\ \neg[\alpha\ bel{:}\ A]]$	α refrains from believing that A
$[\alpha\ bel{:}\ A]$	α believes that A
$[\alpha\ dstit{:}\ \neg[\alpha\ gub{:}\ A]]$	α refrains from giving up the belief that A
$\neg[\alpha\ bel{:}\ A]$	α does not believe that A
$[\alpha\ dstit{:}\ \neg[\alpha\ vab{:}\ A]]$	α refrains from voluntarily acquiring the belief that A

Another distinction is that between α seeing to it that α believes that A and α acquiring the belief that A. The former, not the latter, implies that there is an earlier moment at which α acquired the belief that A.
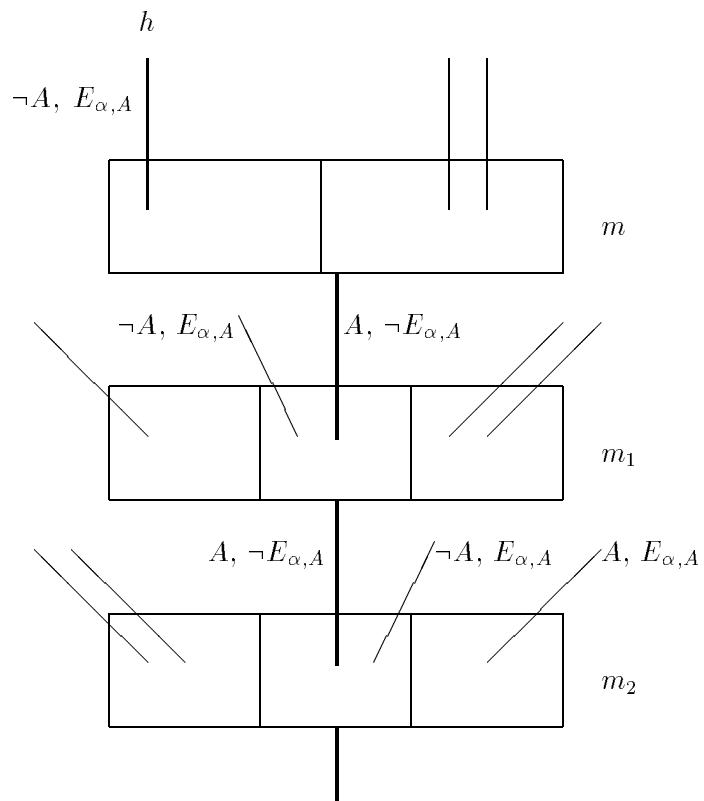
Figure 3: A model in which α believes that A at moment/history pair (m, h).

4 Group beliefs

The richness of the present approach does not only reside in the possible variation resulting from a replacement of classical logic as the non-modal base logic by some more suitable non-classical logic, so that 'α gives up the belief that A' fails to be logically equivalent to 'α acquires the belief that $\neg A$'. Given the reduction of doxastic logic to action logic, the semantics of ascriptions of group agency determines the semantics of ascriptions of group beliefs. The emerging semantics of ascriptions of group beliefs is such that if a group Γ believes that A at a moment history/pair (m,h), it does not follow that every group member α also believes that A at (m,h). This is a property of any sufficiently complex notion of group belief. Normally, the beliefs of groups like political parties, trade unions, scientific associations etc. fail to be closed under membership. If $\Gamma \subseteq Agent$ and h is a history passing through $m \in T$, the set $Choice_\Gamma^m(h)$ of histories choice-equivalent with h for Γ at moment m is defined as $\{h' \mid (\forall \alpha \in \Gamma)\, h' \in Choice_\alpha^m(h)\}$.

Definition 6. [Γ *dstit*: A] is true in a dstit model $\langle T, \leq, Agent, Choice, v \rangle$ at a moment/history pair (m,h) $((m,h) \models$ [Γ *dstit*: A]) iff (i) $\forall h' \in Choice_\Gamma^m(h)$, A is true at (m,h'), and (ii) $\exists h'$ such that $m \in h'$ and A is untrue at (m,h').

Let s_m be any mapping from $Agent$ into the powerset of the powerset of H_m such that $s_m(\alpha) \subseteq Choice(\alpha,m)$. It has been suggested in stit-theory to capture the independence of agents by requiring that for every function s_m,

$$\bigcap\nolimits_{\alpha \in Agent} (s_m(\alpha) \neq \emptyset).$$

In Figure 4, the horizontal partition indicates α's choice cells at m, whereas the vertical partition indicates the sets of histories choice-equivalent for β at m. The choice cells of $\{\alpha, \beta\}$ are denoted as K_1, ..., K_6. In this situation, the agents α and β are not independent of each other because K_2 is empty. In order to define truth conditions for 'Γ voluntarily acquires the belief that A' ([Γ *vab*: A]) and 'Γ voluntarily gives up the belief that A' ([Γ *gub*: A]) in analogy with Definitions 2 and 3, we need a sentence letter $E_{\Gamma,A}$ ("Γ is mistaken with respect to A") for every $\Gamma \subseteq Agent$ and every formula A. [Γ *vab*: A] is defined as [Γ *dstit*: $((\neg A \supset E_{\Gamma,A}) \wedge (A \supset \neg E_{\Gamma,A}))$]; [$\Gamma$ *gub*: A] is defined as [Γ *dstit*: $(\neg(\neg A \supset E_{\Gamma,A}) \wedge (A \supset \neg E_{\Gamma,A}))$]. We then obtain the following notion of group belief:

Definition 7. [Γ *bel*: A] ("Γ believes that A") is true in a dstit model $\langle T, \leq, Agent, Choice, v \rangle$ at a moment/history pair (m,h) iff $\exists m' \in T$ such
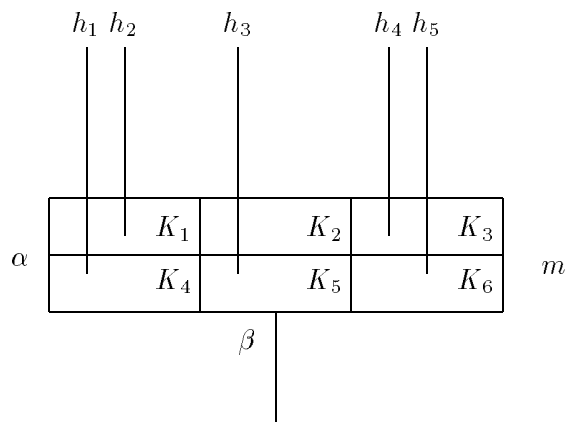
14

Figure 4: Choice cells $K_1, \ldots, K_6$ of $\{\alpha, \beta\}$ at moment m.

that $m' \leq m$ and $(m', h) \models [\Gamma \ vab: A]$ and, moreover, $\neg \exists m'' \in T$ such that $m' \leq m'' \leq m$ and $(m'', h) \models [\Gamma \ gub: A]$.

Figure 5 exemplifies the failure of closure of group belief under membership. The group $\{\alpha, \beta\}$ believes that A at m. Closure under membership fails because at m, β gives up the belief that A.

5 Concluding remarks

In this paper I have suggested a reduction of doxastic logic, the logic of belief, to the logic of concrete actions. The development of this reduction is mainly conceptual and based on a certain epistemological position, namely a variant of doxastic voluntarism. The proposed reduction makes use of the formal apparatus of stit-theory; in particular it uses the semantics of the dstit operator and the idea of introducing sentence letters in order to define belief in terms of belief acquisition and belief abandonment. The striking features of the resulting doxastic logic are: failure of closure of belief under logical consequence and failure of closure of group belief under membership. One might suspect that the constants $E_{\alpha,A}$ and $E_{\Gamma,A}$ introduce a syntactic component into the semantic representation in a way quite similar to the introduction of awareness functions into possible worlds models. In the former approach the idea is that α may or not be mistaken with respect to A, in the latter the idea is that α may or not be aware of A. However, aware-
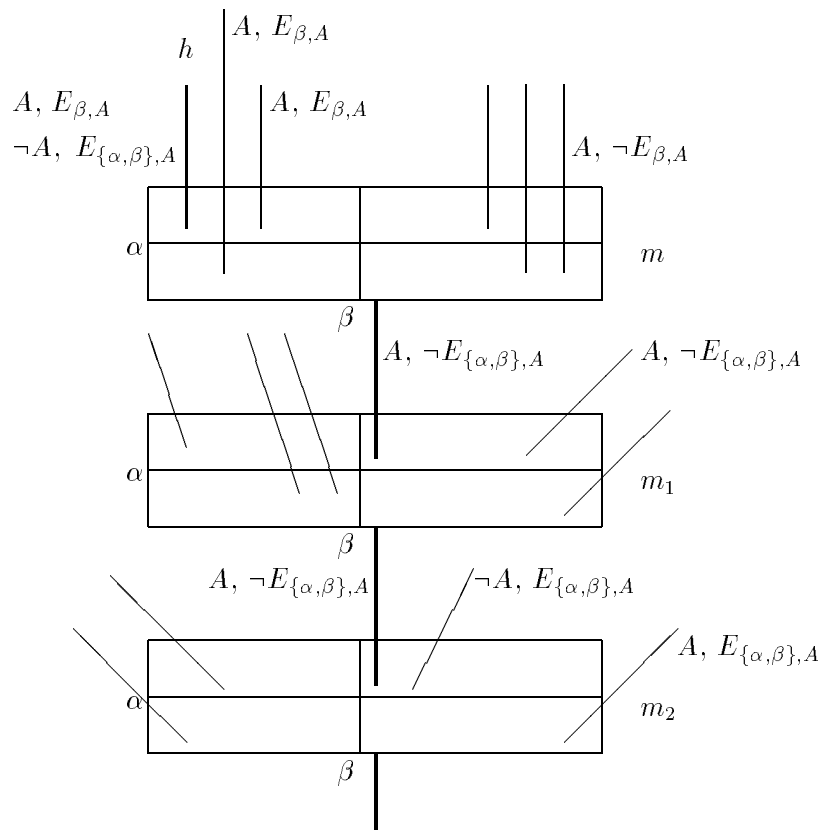
Figure 5: A model in which $(m,h) \models [\{\alpha, \beta\}\ bel\colon A]$, although not $(m,h) \models [\beta\ bel\colon A]$.

ness functions operate as syntactic filters on 'implicit belief' represented as a normal modal operator closed under logical consequence, thereby giving rise to a notion of 'explicit belief' not closed under logical consequence. The role of $E_{\alpha,A}$ and $E_{\Gamma,A}$, however, is completely conceptual. In the present approach, failure of closure under logical consequence and failure of closure of group belief under membership result from the agentive conception of belief acquisition and belief abandonment used in the semantic representation of belief ascriptions together with the definition of the belief operators. Conjunction distribution $[\alpha\ bel{:}\ (A \wedge B)] \supset ([\alpha\ bel{:}\ A] \wedge [\alpha\ bel{:}\ B])$, for example, fails to be valid because the negative condition may be fulfilled for $(A \wedge B)$ without being fulfilled for both A and B. The converse of conjunction distribution fails to be valid because there may be previous moments m' and m'' such that A is true at (m',h) and B is true at (m'',h), although there is no previous moment m such that $(A \wedge B)$ is true at (m,h).

Notes

1 Of course, further variants are possible. Instead of quantifying over all or some beliefs, one might for instance want to use some complex generalized quantifier.

2 It is assumed here that the modality has wide scope. Moreover, acquisition of arbitrary beliefs is to be understood either as acquisition of beliefs with an arbitrary content or as acquisition of dispositions towards one's own actions based on one's own intentions and other beliefs of oneself.

3 Another reconstruction of the argument can be found in (Winters 1979). In this reconstruction, the valid argument rests on a false assumption. The main observation is that if one assumes that a belief has been voluntarily acquired in full consciousness, it is nevertheless possible to maintain the belief "for reasons other than those involved in the original acquisition" (1979, p. 253). Winters then defends the impossibility of voluntary belief *sustainment* in full consciousness.

4 Although Evans (1963) draws a distinction between error and mistaken belief, I see no reason to do so.

5 This procedure resembles A.R. Anderson's introduction of a sanction constant into deontic logic, see also (Wansing 1998). Intuitively, the sanction constant is true at a moment/history pair (m,h) iff the agent under consideration does wrong at (m,h). It seems natural to explicate obligation in terms of wrongdoing, and it also seems natural to explicate belief in terms of error.

References

[1] N. Belnap. Backwards and Forwards in the Modal Logic of Agency. *Philosophy and Phenomenological Research* 51 (1991), 8–37.

[2] N. Belnap and M. Perloff. Seeing to it that: a Canonical Form for Agentives. *Theoria* 54 (1988),175–199.

[3] N. Belnap, M. Perloff, and M. Xu. *Facing the Future.* Manuscript, Philosophy Department, University of Pittsburgh, 1996.

[4] J. Bennett. Why Belief is Involuntary. *Analysis* 50 (1990), 87–107.

[5] J. Evans. Error and the Will. *Philosophy* 38 (1963), 136–148.

[6] R. Fagin and J. Halpern. Belief, Awareness and Limited Reasoning. *Artificial Intelligence* 34 (1988), 39–76.

[7] J. Horty. An Alternative stit-operator. Manuscript, Philosophy Department, University of Maryland, 1989.

[8] J. Horty and N. Belnap. The Deliberative stit: a Study of Action, Omission, Ability, and Obligation. *Journal of Philosophical Logic* 24 (1995), 583–644.

[9] R. Hoyler. Belief and Will Revisited. *Dialogue* 22 (1983), 273–290.

[10] W. James. The Will to Believe. 1896. In: *Essays in Pragmatism*, New York, Hafner Press, 1969, 88-105.

[11] F. von Kutschera. Bewirken. *Erkenntnis* 24 (1986), 253–281.

[12] D. Pearce and H. Wansing. On the Methodology of Possible Worlds Semantics, II: Non-normal Worlds and Propositional Attitudes. Freie Universität Berlin, Gruppe für Logik Wissenstheorie und Information, Bericht 2/89, 1989.

[13] L. Pojman. Believing and Willing. *Canadian Journal of Philosophy* 15 (1985), 37–55.

[14] V. Rantala. Impossible Worlds Semantics and Logical Omniscience. *Acta Philosophica Fennica* 35 (1982a), 106–115.

[15] V. Rantala. Quantified Modal Logic: Non-normal Worlds and Propositional Attitudes. *Studia Logica* 41 (1982b), 41–65.

[16] E. Thijsse. *Partial Logic and Knowledge Representation.* Eburon Publishers, Delft, 1992.

[17] E. Thijsse and H. Wansing. A Fugue on the Themes of Awareness Logic and Correspondence. *Journal of Applied Non-Classical Logic* 6 (1996), 127–136.

[18] H. Wansing. Bemerkungen zur Semantik nicht-normaler möglicher Welten. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 35 (1989), 551-557.

[19] H. Wansing. A General Possible Worlds Framework for Reasoning about Knowledge and Belief. *Studia Logica* 49 (1990), 523–539, 50 (1990) 359.

[20] H. Wansing. Nested Deontic Modalities: Another View of Parking on Highways. *Erkenntnis* 49 (1998), 185–199.

[21] H. Wansing. Handlungstheoretische Aspekte der Theoriewahl. Manuscript, Department of Philosophy, Dresden University of Technology, 1999, submitted for publication.

[22] B. Williams. Deciding to Believe. In: *Problems of the Self*, New York, Cambridge University Press, 1973, 136–151.

[23] B. Winters. Believing at Will. *Journal of Philosophy* 76 (1979), 243–256.

[24] M. Xu. Axioms for Deliberative *Stit. Journal of Philosophical Logic* 27 (1998), 505–552.