

Anselm's Logic of Agency

Sara L. Uckelman
Institute for Logic, Language, and Computation
suckelma@illc.uva.nl

1st October 2007

Abstract

The origins of treating agency as a modal concept go back at least to the 11th century when Anselm, Archbishop of Canterbury, provided a modal explication of the Latin *facere* ‘to do’, which can be formalized within the context of modern modal logic and neighborhood semantics. The agentive logic induced by his conception satisfies the traditional square of opposition, but also has some unique properties which reflect the fact that Anselm’s modal view of agency is grounded strongly in non-logical philosophical and theological considerations. Our conclusion is that Anselm’s logic of agency when presented with the tools of modern formal logic is still a viable theory of agency for current times.

keywords: agency, Anselm of Canterbury, neighborhood semantics, ordinary language, square of opposition

1 Introduction

The idea of treating agency as a modal concept can be traced further back in time than many people would think. In giving a history of modal logics of agency, [Belnap et al. 2001] note that the first person to consider the modal interpretation of agency in any rigorous fashion is Anselm of Canterbury. In fragmentary notes which were compiled and partially organized shortly after his death in 1109, there are discussions of the meaning and functions of the Latin verbs *facere* ‘to do’, *posse* ‘to be able’, and *velle* ‘to will’. In particular, in his discussion of *facere* ‘to do’, Anselm identifies four types of doing and further subdivides each type into six different modes. The relationships between the four types can be placed neatly into a square of opposition. According to [Belnap et al. 2001], it is this square of opposition which “clearly indicates that he [Anselm] had in mind a modal logic of agency” (p. 19).

This modal interpretation of agency found in Anselm shows that the idea of treating agency as a modal concept is far older than many action theorists might have thought. This gives us at least two reasons why a modern logician would be interested in this historical theory. The first reason is the purely formal question of what the modal logic of Anselm’s theory of agency actually is, whether it is identical with any of the standard modern agentive logics or whether Anselm’s constraints resulted in something new. The second reason is

the philosophical question of whether this historical theory has any insights to offer to modern problems and questions of agency.

Our focus in this paper is primarily the formal question. However, in order to answer the formal questions we must consider the philosophical and theological motivations of Anselm which underpin various aspects of this theory. We hence start by giving a brief introduction to the theory, including the non-logical motivations, in §2 and §3. After that, we turn to the details of the theory, the four types and six modes of agency, and the square of opposition in which they can be placed, in §4. The final part of the paper, §§5,6, is devoted to considering how the theory might be formalized using modern techniques. In particular, we look at two proposed syntaxes for Anselm's theory, given in [Walton 1976a] and [Walton 1976b]. We show that because Walton did not have an adequate semantics, his syntaxes introduced features which are not found in Anselm's original theory, and that given an adequate semantics, namely that of neighborhood models, we can give formalizations for a number of different interpretations of Anselm's theory.

2 The theory

The Lambeth manuscript contains both unpublished manuscript notes and letters of Anselm's. It is believed that Anselm composed the parts on *facere*, *posse*, and *velle* while he was archbishop of Canterbury [King –, p. 1]. The Lambeth fragments were first edited and published in [Schmitt 1936], and then again in [Southern & Schmitt 1969, pp. 333–354]. The text is partially translated with detailed commentary in [Henry 1967] and translated, with little commentary, in [Hopkins 1972].

The treatment of *facere* is in the setting of a discussion between a teacher and his student. The opening statement of the teacher contains all of the details of the theory in a nutshell:

Teacher: We commonly use the verb 'to do' in place of all other verbs, regardless of the signification of these other verbs and regardless of whether they are finite or infinite. In fact, 'to do' may even stand for 'not to do'. If you think about it carefully, you will see that when we ask about someone 'What (how) is he doing?' here 'doing' stands for any verb that can be given in answer [Hopkins 1972, p. 218].

An infinite verb is one which is the complement of a finite verb, i.e., one which indicates a finite action. For example, 'run' is a finite verb, and 'not run' is an infinite verb. This terminology is taken from Boethius. An infinite verb should not be confused with the infinitive form of a verb.

The first modern discussions which recognize Anselm's theory as developing a modal interpretation of agency can be found in [Henry 1953] and [Henry 1967]. When discussing the history of modal views of agency, [Belnap et al. 2001] quote a passage from the Lambeth fragments plus Henry's translation of said passage (p. 19):

Quidquid autem 'facere' dicitur, aut facit ut sit aliquid, aut facit ut non sit aliquid. Omne igitur 'facere' dici potest aut 'facere esse' aut 'facere non esse' [Henry 1967, p. 124].

Paraphrase by Henry: For all x , if ‘ x does’ is true, then x does so that something either is so or is not so. Hence the analysis of ‘doing’ will in fact be an analysis of x ’s doing so that p , and of x ’s doing so that not- p , [where ‘ p ’ is a clause describing a state of affairs, and ‘not- p ’ is short for ‘it is not the case that p ’].

The most striking feature of this theory, in this informal introduction, is its breadth. Its breadth comes in two aspects. The first is that, for Anselm, an analysis of ‘to do’ will encompass an analysis of *all* verbs. The second aspect is that in statements of the form ‘ x does’, x can be anything. For example, according to Anselm if I trip over a tree which has fallen across the path, then it is perfectly plausible to ask “What did the tree do?” and respond “It tripped me.” While modern agency theorists focus on explications of ‘ x does’ where x is a causal agent, Anselm says that “investigation will show that general use of the theory need not be confined to such contexts” [Henry 1967, p. 124].

At first, one might find the breadth of this theory worrisome, for it may turn out that the theory is too broad, and that either it cannot be used in particular situations or that too many things end up counting as examples of agency. The worry that we have is this: How can we expect to find a reasonable explication of what *facere* means and how it functions, if we will not be distinguish it from any other verb? Once we understand the philosophical and theological foundations which motivate Anselm’s theory, it will be clear that any more restricted conception of agency would be untenable. Further, while his scope is wider than many modern theories of agency, this breadth does not hurt him and in fact its very generality can be said to be a positive feature which sets his theory above other theories. The very breadth of his account demonstrate Anselm’s capacity to connect logic, ordinary language, and theology into a single, coherent theory. The questions which arise on a first reading of the theory can be answered when we consider the non-logical, that is philosophical and theological, motivations behind Anselm’s theory, which we do in the next section.

This is not to say that a more focused view of agency, pertaining specifically to human agency, might not also be a useful and fruitful exercise. Serene in [Serene 1983] provides an in-depth study of Anselm’s work in the Lambeth fragments which focuses on the connections between doing and willing. This is not a logical article; there is no discussion of axioms, syntax, or semantics. However, she does point out some of the general characteristics of Anselm’s conception of *doing* which must be kept in mind when we look towards developing a logic for his conception of agency:

Because the analysis of *facere* is meant to apply to all instances of agency, whether or not the subject is human, rational, conscious or even an efficient cause of the outcome, it does not constitute a complete or a specific account of human action. His full theory of human agency also includes some explanation of the nature of willing [Serene 1983, p. 144].

As a result, any logic which is developed only to address Anselm’s discussion of *facere* must not turn on any specifically human (or even sentient property). Such facts only have a use when we are trying to give a full account of *human* agency, in which case an evaluation of *velle* (to will) will also be necessary. While it is possible to look only at accounts of human agency in the context

of the broader theory (which is what we will be doing), it seems likely that a full account of specifically human agency, involving a discussion of willing, may cause us to modify the logic we propose for *facere*. Additionally, it is clear that any logical representation of Anselm's theory will differ from many modern theories of human agency.

3 Philosophical and theological motivations

Before looking more closely at the logical aspects of the theory, we first discuss the motivations for Anselm's account. The first motivation is philosophical or methodological. Much of Anselm's discussions of logical matters involve separating questions of logic and logical usage from questions of grammar and everyday usage, separating the use *proprie* from the use *non proprie*. He takes care to point out that everyday usage (*usus loquendi*, *usus non proprie*) of words is often sloppy, and what we say doesn't always accurately represent what we mean. (This topic is a main focus of his *De Grammaticis*, some of the themes of which are echoed in the Lambeth fragments.)

The aim of the grammarian is to explain the *usus loquendi* of terms; his goal is *descriptive*. The logician, on the other hand, has two options. He can either ignore the *usus loquendi* altogether, and make his aim strictly prescriptive, by focusing on the proper, logical uses of the terms involved, even when this explication seems at odds with our everyday uses of the terms. Alternatively, he can allow his logical explication to be broad enough to cover and hence to *explain* to some extent, the *usus loquendi*. Parts of Anselm's logical works take the former route, but in his discussions on agency he always allows for taking into account the latter route. In discussing *facere* he notes that "firstly, the ways in which 'to do' is used in ordinary speech, i.e., in the contexts in which it occurs, are to be codified so that the deviations of these uses from the proper sense become evident" [Henry 1967, p. 123]. We cannot fully understand the proper usage of a term until we understand how ordinary usage differs from proper usage. [Henry 1960] claims that Anselm's discussion of *facere* "is intended as a means of analyzing the senses of verbs as they occur in customary utterance (*usus loquendi*), in non-strict oblique uses as measured against the standard of their precise or strict signification, the latter being shown by exemplifying the simplest overt meaning of the verb in question" (p. 377). The goal, then, is to produce a logical explanation for the result of the grammarian's study of the word.

The other motivations guiding Anselm's account of agency stem from this explanatory motivation, and are mainly theological in nature. Any theory of agency which he proposes should be able to explain why *facere* is used the way that it is in scripture. An explication of agency which does not make sense of scriptural usages of *facere* will not be adequate for Anselm:

Indeed, the Lord Himself in the Gospel uses *facere* and *agere* . . . in place of every other verb when He says, "Whoever does evil hates the light" and "Whoever does the truth comes to the light" (John 3:20–21) . . . Whoever sits or stands where or when he ought not, does evil; and whoever is not present, does not sit, or does not stand where or when he should also does evil . . . Likewise, he does the truth who is present, is sitting, or is standing where and when he

ought, and is not present, is not sitting, or is not standing where and when he ought not. In this way the Lord reduces every verb, whether positive or negative, to a form of ‘to do’ [Hopkins 1972, p. 220].

And just as *usus loquendi* is very broad, so too is scriptural use:

The Lord wishes to convey that ‘to do’ may be used not only in respect of that which is properly asserted to constitute ‘doing’ but also in respect of all verbs ... The ordinary use of language also has this feature, namely, it treats as ‘doing’ both *undergoing* and many other things which are not really cases of *doing* (*De Veritate* in [Henry 1967, pp. 182–183]).

If our logical theory of agency can provide an explanation of the *usus loquendi*, then we will also have an explanation of the theological usage of the word, because the two combine.

The desire to give an adequate account of the scriptural usage of *facere* was not merely an idle exercise in logic and grammar. Because the concept of agency is closely connected to issues in responsibility for actions and hence culpability and sin, an explanation of the proper conditions under which agency can be ascribed will have implications for ethics as well as logic. This issues can be seen in Matthew 25:31–46, where on the day of judgment God will separate the sheep from the goats on the basis of what they did and didn’t do.¹ A similar sentiment is found in *De Casu Diaboli*, which Henry roughly translates as follows:

For when in the latter instance, someone is said to bring it about that the victim is naked, or that the victim is not clothed, the exact import is that although the person was capable of doing so, he did not bring it about that the victim was not naked, or that the victim remained clothed” [Henry 1967, p. 184].

Knowledge of correction ascriptions of agency, both in proper usage and in common usage, is hence important because it gives knowledge concerning eternal culpability.

4 The types and modes of doing

Now that we have seen some of the motivations underlying the informal statement of the theory that we presented in §2, we can turn to the details of the account.

Doing will always result in something being or not being the case. Something can either be or not be the case because it is either caused or not caused. This gives us four types of agency:

A to cause to be

B to cause not to be

¹Matthew 25:40 “The King will reply, ‘I tell you the truth, whatever you did for one of the least of these brothers of mine, you did for me’,” and 25:45 “He will reply, ‘I tell you the truth, whatever you did not do for one of the least of these, you did not do for me’.”

C not to cause not to be

D not to cause to be

Types (A) and (B) are called *affirmative*. Types (C) and (D) are called *negative*. The implication relationships between these four types of agency form a square of opposition. The graphical square itself is not present in Anselm's work, but the verbal descriptions of the relations fix the graphical square uniquely. (See Figure 1.)

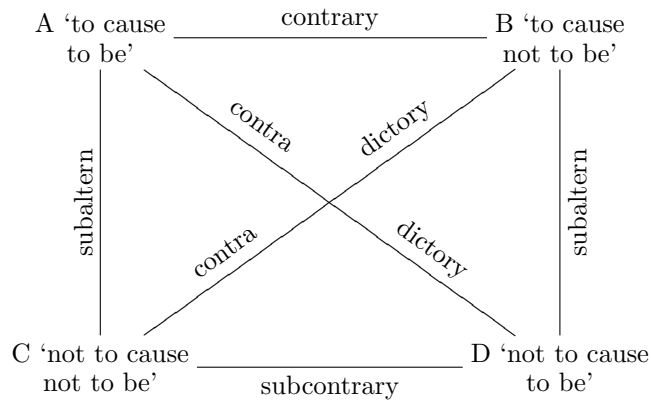


Figure 1: Agentive Square of Opposition

It is interesting to compare this agentive square of opposition with the traditional Aristotelian square of opposition for categorical sentences. (See Figure 2.)

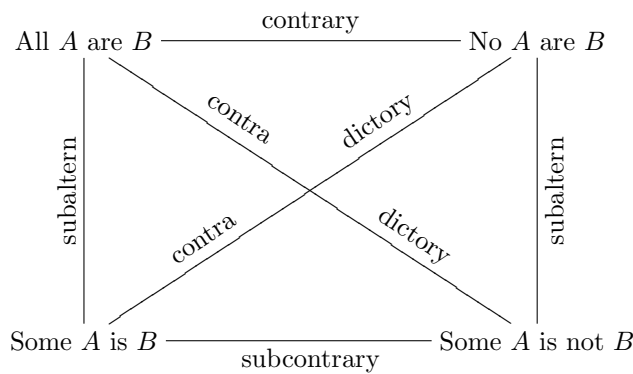


Figure 2: Categorical Square of Opposition

In the traditional Aristotelian square, the affirmative statements are on the left of the square and the negative on the right, whereas in the agentive square, the affirmative statements are on the top and the negative statements on the bottom. This means that in the Aristotelian square, the three oppositions relationships (contrary, subcontrary, and contradictory) always hold between one negative and one positive relationship. In contrast, in the agentive square, the

two positive statements are contraries and the two negative statements are sub-contraries, and a positive statement always implies a negative one.

Each of the four types of action can be further divided into six modes, each of which picks out a different way that the main type of action can be brought about. For example, of type *A* ‘to cause to be’

[w]e say that *x* causes *y* when *x* causes *y* itself to be; or when *x* does not cause *y* itself not to be; or when *x* causes *y* to be by causing *z* to be, by not causing *z* to be, by causing *z* not to be, or by not causing *z* not to be [Hopkins 1972, pp. 221–222].

and

[w]e say that a thing causes something else not to be either because (1) it directly causes this other thing not to be, or (2) it does not directly cause it to be, or (3) it causes an intervening thing to be, or (4) it does not cause an intervening thing to be, or (5) it causes an intervening thing not to be or (6) it does not cause an intervening thing not to be [Hopkins 1972, p. 223].

Let us illustrate these six different modes with an example, ‘to cause to be dead’:

- A1 Killing directly (*facere idipsum esse*)
- A2 Not making not dead (*non facere idipsum non esse*)
- A3 Making the killer have arms (*facere aliud esse*)
- A4 Not arming the victim (*non facere aliud esse*)
- A5 Making the victim not armed (*facere aliud non esse*)
- A6 Not making the killer not armed (*non facere aliud non esse*)

This list distinguishes between positive agency (where the agent does something) and negative agency (where the agent does not do something), as well as direct *per se* agency (where the agent brings about the effect himself) and indirect *per aliud* agency (where he causes some other being to bring the effect about). There is a further distinction that can be made in cases of *per aliud* agency. They divide into cases where the agent indirectly performs an action and where the agent indirectly does not perform an action (we can call this ‘proximal’ and ‘distal’, respectively). Thus, the six types listed above can be classified as follows:

- 1 Positive *per se*
- 2 Negative *per se*
- 3 Positive, proximal, *per aliud*
- 4 Negative, proximal, *per aliud*
- 5 Positive, distal, *per aliud*
- 6 Negative, distal, *per aliud*

At this point we will anticipate §5 by introducing some notation. We use the operator δ for ‘does’, e.g. δp is read ‘does so that p ’. We can represent the six modes of type A ‘to cause to be’ as follows:

A1 δp

A2 $\neg\delta\neg p$

A3 $\delta q \wedge (q \leftrightarrow p)$

A4 $\neg\delta q \wedge (q \leftrightarrow \neg p)$

A5 $\delta\neg q \wedge (q \leftrightarrow \neg p)$

A6 $\neg\delta\neg q \wedge (q \leftrightarrow p)$

At this point we say nothing about how the relationship expressed by $q \leftrightarrow p$ is to be interpreted.

Each of the four types of agency can be expressed in each of the six modes, which means that we have potentially twenty-four types of agency. Within each type, the six modes are all independent; they can neither be defined by each other nor do they imply each other. The relationships between the four types do, in a sense, ‘trickle down’ to the modes within each type. For example, $A1$ is the contrary of $B1$, $D3$ is a contradictory of $A3$, and so on. As a result, it turns out that types $A1$ and $C2$ are identical, and likewise $C1$ and $A2$, and the same for $B1, B2, D1, D2$. The other sixteen combinations of modes and types are all logically independent, hence the result is twenty distinct ways that agentive statements can be expressed.

Anselm’s thesis is that any ascription of doing will be one of these twenty-four forms. In ordinary usage, the twenty logically distinct forms are often used interchangeably, as if they were not distinct but equivalent. Anselm notes that in *usus loquendi*, we often use affirmative claims as a short hand when what we really mean is the negation of the contrary, for example: “‘ x does so that there are evils’ can be used in cases where the more proper expression would be ‘ x does not so that evils are not’” [Henry 1967, p. 125]. In a more detailed explanation, Anselm says:

It must be noted that while the first mode of the negative tables [modes C and D] simply negates, without implying anything else, each of the five subsequent modes in the negative tables contains statements which can be substituted for those statements which appear in that table which is the contrary of their corresponding affirmative table. For example, whoever revives someone may be said “not to cause him to be dead” in the place of “to cause him not to be dead”; and we may also substitute “not to cause him not to be living” for “to cause him to be living”... [Hopkins 1972, p. 227].

That is, forms $C2 - C6$ and $D2 - D6$ are often used equivalently with forms $B2 - B6$ and $A2 - A6$, respectively, even though, strictly speaking, forms $C3 - C6$, $D3 - D6$, $B3 - B6$ and $A3 - A6$ are all nonequivalent. (As noted earlier, $C2$ is equivalent to $A1$, and so one for the first and second modes of each type.)²

²The same phenomenon shows up later in the same philosophical fragments, when Anselm

This is an example of Anselm demonstrating how ordinary usage can be explained in part by their logical definitions and relations. Nevertheless, we need to remember that though we may use the locutions interchangeably, “they are different from each other” [Hopkins 1972, p. 225].

So much for ordinary usage. What about proper usage? According to Anselm, the answer is simple. Only the first mode of each type represents proper usage:

‘ x does so that p ’ has the proper sense, ‘ x does so that p , which was not the case, becomes the case’... [and] is only properly used in the case described by [the positive *per se* mode]’ [Henry 1967, p. 126].

Note that this is partially contrary to Serene’s assertion that “only ascriptions made in mode one are ‘proper’, since this is the only mode in which the agent’s action directly causes the outcome ascribed to him. Ascriptions in mode two are ‘improper’ because the directly relevant factor is the agent’s failure to act rather than his directly doing what is ascribed to him.” [Serene 1980, p. 123]. Ascriptions in the second mode of the negative types ($C2$ and $D2$) must count as proper if the first mode of the two positive types are to count as proper, since they are identical. And likewise, if the first mode of the negative types are to be considered proper ascriptions of agency, then the second mode of the positive types must also be considered proper, for the same reason.

5 Formal approaches

We now turn to the more technical part of this paper. In this section we introduce a semantics for modal logic which we will use in the next section to examine the syntax proposed by Walton in [Walton 1976a] and [Walton 1976b]. It will turn out that the type of modal logic which best expresses the features of Anselm’s account of agency is a non-normal modal logic. Traditional semantics for normal modal logics are not adequate for modeling non-normal modal logics, so we will use instead neighborhood semantics. We begin with a brief review of normal modal logics and their usual semantics. Our base logic is classical propositional logic, to which we add one or more modality operators.

Definition 5.1. A *normal* modal logic is one which contains the following axioms, where \Box represents an arbitrary modality operator:

$$\begin{aligned} (K) \quad & \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q) \\ (M) \quad & \Box(p \wedge q) \rightarrow (\Box p \wedge \Box q) \\ (C) \quad & (\Box p \wedge \Box q) \rightarrow \Box(p \wedge q) \\ (N) \quad & \Box \top \end{aligned}$$

uses his explication of *facere* as a model for his discussion of *esse* ‘to be’, *habere* ‘to have’, and *debere* ‘to be obliged, ought’. He says: “We also say that we are not ‘obliged to sin’ (*non debere peccare* as a substitute for saying that we are ‘obliged not to sin’ (*debere non peccare*)). But properly speaking not everyone who does what he is not obliged to do sins... Now as you remember, we said earlier that ‘not to cause to be’ may be used in place of ‘to cause not to be’. In the same way, we say ‘is not obliged to’ for ‘is obliged not to’, and ‘is not obliged to sin’ for ‘is obliged not to sin’. But our [Latin] usage is such that ‘is not obliged to sin’ we really mean ‘is obliged not to sin’” [Hopkins 1972, pp. 231–232]. In modern linguistics, this phenomenon is called “negation raising” or “neg raising”. For general information on negation raising, see [Horn 1989]. My thanks to Laurence Horn for drawing to my attention this parallel occurrence.

and is closed under *modus ponens*, uniform substitution, and the rule of necessitation *RN* (from $\vdash \phi$ infer $\vdash \Box\phi$).

The minimal modal logic which contains all of these axioms and satisfies these rules of inferences is called *K*. The standard semantics for a normal modal logic are Kripke relational semantics: a structure is a *frame* $\mathfrak{F} = \langle W, R \rangle$ where W is a non-empty set and R is a binary relation on W , and a model is frame plus a valuation function, e.g. $\mathfrak{M} = \langle \mathfrak{F}, V \rangle$, where V is a map from atomic sentence letters to $\mathcal{P}(W)$.

As we will see in more detail in §6, some of the rules of inferences and axioms of normal modal logics are problematic when we try to apply them to agency. We therefore look at axiom systems which are weaker than *K*, namely ones that do not have the necessitation rule and which omit one or more of the axioms listed above. Since *K* is characterized by the class of all Kripke frames, these sub-*K*, non-normal logics cannot have Kripke frames as their semantics. Instead, non-normal modal logics are usually modeled with an alternative method, that of neighborhood models.

Definition 5.2. A *neighborhood model* is a structure $\mathfrak{M} = \langle W, N, V \rangle$ where

- W is a set of points, called worlds.
- N is a function from W to $\mathcal{P}(\mathcal{P}(W))$, such that $N(w)$ is called “the neighborhood of w ”.
- V is a function from atomic sentence letters to W , such that if $w \in V(p)$, we say that p is true at w , or $V(p, w) = 1$.

In these models, each formula is associated with a *truth set*.

Definition 5.3. Let \mathfrak{M} be a neighborhood model and ϕ a formula. The *truth set for ϕ in \mathfrak{M}* is $\|\phi\|^{\mathfrak{M}} = \{w \in W : w \in V(\phi)\}$

The clauses in the truth definition for the propositional connectives are as expected. We add only the following for the modal connective:

Definition 5.4. Let \mathfrak{M} be a neighborhood model, w a world in \mathfrak{M} and ϕ a formula. Then

$$\mathfrak{M}, w \models \Box\phi \text{ iff } \|\phi\|^{\mathfrak{M}} \in N(w)$$

Most modal logics are designed to deal with the dual notions of possibility and necessity, and a second operator \Diamond is introduced (and often, but not necessarily, defined as $\neg\Box\neg$). Because there isn’t any natural dual notion for agency, we omit most references to the \Diamond operator here.

There are various restrictions that we might wish to place on the N function.

Definition 5.5. Let \mathfrak{M} be a neighborhood model. Then, for every world $w \in W$ and all subsets $X, Y \subseteq W$,

m \mathfrak{M} is *monotonic*, or *supplemented*, iff if $X \cap Y \in N(w)$ then $X \in N(w)$ and $Y \in N(w)$.³

c \mathfrak{M} is *closed under intersection* iff if $X \in N(w)$ and $Y \in N(w)$ then $X \cap Y \in N(w)$.

³Equivalently, if $X \subseteq Y$, and $X \in N(w)$ then $Y \in N(w)$.

n \mathfrak{M} contains the unit iff $W \in N(w)$.

t \mathfrak{M} is truth-making iff if $X \in N(w)$ then $w \in X$.

The modal logic characterized by the class of all neighborhood models the logic E (so called in [Chellas 1980]). E has one axiom:

$$(E) \quad \Box\phi \equiv \neg\Diamond\neg\phi$$

and one rule of inference:

Rule 5.6. (RE) From $\vdash \phi \equiv \psi$ infer $\vdash \Box\phi \equiv \Box\psi$

If we restrict the class of models under consideration to ones which have some subset of the properties listed above, we can classify certain other non-normal logics. (Proofs for theorems not otherwise given can be found in [Chellas 1980, pp. 214–217].) Let \mathfrak{M} be a neighborhood model. Then

Theorem 5.7. $\mathfrak{M} \models \Box(p \wedge q) \rightarrow (\Box p \wedge \Box q)$ iff \mathfrak{M} is monotonic.

Theorem 5.8. $\mathfrak{M} \models (\Box p \wedge \Box q) \rightarrow \Box(p \wedge q)$ iff \mathfrak{M} is closed under intersection.

Theorem 5.9. $\mathfrak{M} \models \Box\top$ iff \mathfrak{M} contains the unit.

Theorem 5.10. $\mathfrak{M} \models \Box p \rightarrow p$ iff \mathfrak{M} is truth-making.

Hence, each restriction on the neighborhood function N corresponds to a natural constraint on the models. These constraints on models give rise to different classes of logics.

Definition 5.11. Introduce the following abbreviations:

$$\mathbf{M} \quad \Box(p \wedge q) \rightarrow (\Box p \wedge \Box q)$$

$$\mathbf{N} \quad \Box\top$$

$$\mathbf{C} \quad (\Box p \wedge \Box q) \rightarrow \Box(p \wedge q)$$

$$\mathbf{T} \quad \Box p \rightarrow p$$

We define EM to be the smallest logic containing both E and M and closed under RE , and similarly for EN , EC , ET , EMC , $EMCT$, etc.

The axioms which characterize certain constraints on the neighborhood function N each correspond to a rule of inference.

Lemma 5.12. *Instead of adding to E the axioms M , N , or C , we could have added the following rules of inference:*

m From $\vdash p \rightarrow q$ infer $\vdash \Box p \rightarrow \Box q$

n From $\vdash p$ infer $\vdash \Box p$

The logics between E and K extended by M , N , and C form a boolean lattice where each combination is distinct and $ECMN = K$. Each of these with T as a further axiom are also all distinct logics; we prove a few of the cases:

Theorem 5.13. $EMT \neq C$

Proof. Let $\mathfrak{M} = \langle W, N, V \rangle$ where

$$\begin{aligned} W &= \{0, 1, 2\} \\ N(0) &= \{\{0, 1\}, \{0, 2\}, \{0, 1, 2\}\} \\ N(1) &= \emptyset \\ N(2) &= \{0, 1, 2\} \\ V(p) &= \{0, 1\} \\ V(q) &= \{0, 2\} \end{aligned}$$

Then C is falsified at 0, because $\|p\|^{\mathfrak{M}} = \{0, 1\} \in N(0)$ and $\|q\|^{\mathfrak{M}} = \{0, 2\} \in N(0)$, but $\|p\|^{\mathfrak{M}} \cap \|q\|^{\mathfrak{M}} = \{0\} \notin N(0)$. Further, E , M , and T are true everywhere. (It should be clear that this model also does not satisfy $\Box\top$, because of 1.) \square

Theorem 5.14. $ECT \neq M$

Proof. Let $\mathfrak{M} = \langle W, N, V \rangle$ where

$$\begin{aligned} W &= \{0, 1\} \\ N(0) &= \{0\} \\ N(1) &= \{1\} \\ V(p) &= \{0, 1\} \\ V(q) &= \{0\} \end{aligned}$$

Then, since $\Box p$ is true nowhere, both T and C are satisfied everywhere, but $0 \models \Box(p \wedge q)$, and hence M is falsified. (It should also be clear that this model also does not satisfy $\Box\top$, either.) \square

In the next section we will apply these different classes of models to the syntax developed in [Walton 1976a] and [Walton 1976b] to see whether certain questions which he leaves open can be settled.

6 The syntax of agency

In this section we will look at certain modern syntactical representations of Anselm's theory, and evaluate them with the help of the semantics presented in the previous section.

The first syntax that we look at is both cursory and unfortunate, found in [Danto 1973]. Danto says very little about Anselm; there is a brief mention and then a footnote. He uses Anselm as a justification for introducing the expression mDa , to be read “ m makes happen the event a by doing a ”. He says that

the locution mDa covers the stiltedness of the expression ‘... makes ... happen by ... -ing’ and permits us to treat actions in a generalized manner by treating ‘does’ for the moment as an auxiliary of action verbs, much as ‘knows’ may be an auxiliary of cognitive verbs. In doing so, I follow the illustrious precedent of Anselm of Canterbury

who in discussing the Latin verb *facere* treats it in similar auxiliary fashion (p. 7).⁴

This is an unfortunate case where symbolic notation is introduced as a method of clarifying the underlying structure of the sentences being discussed but where in fact the notation ends up merely hiding the relevant issues without explaining them. We will not investigate this syntax any further.

In §4 we introduced the notation δp ‘does so that p ’. We modify it here slightly. Following Walton, we start with a classical propositional language made up out of an infinite set of propositions P and a (possibly but not necessarily infinite) set of agents A from which we build the following formulas:

$$p \mid \neg\phi \mid \phi \wedge \psi \mid \phi \rightarrow \psi \mid \delta_a\phi \text{ for } a \in A$$

where the operators δ_a , one for each agent in A , are our does modalities.⁵

The addition of a does modality for each agent means that we are actually working within a multi-modal setting, so we need to slightly modify our definition of neighborhood model from the previous section, with a corresponding modification to the truth conditions for agentive formulas.

Definition 6.1. A *multi-modal neighborhood model* is a structure

$$\mathfrak{M} = \langle W, A, N_a \text{ for } a \in A, V \rangle$$

where

- W is a set of points, called worlds.
- Each N_a is a function from W to $\mathcal{P}(\mathcal{P}(W))$, such that $N_a(w)$ is called “the neighborhood of w for a ”.
- V is a function from atomic sentence letters to W , such that if $w \in V(p)$, we say that p is true at w , or $V(p, w) = 1$.

Definition 6.2. Let \mathfrak{M} be a multi-modal neighborhood model, w a world in \mathfrak{M} and ϕ a formula. Then

$$\mathfrak{M}, w \models \delta_a\phi \text{ iff } \|\phi\|^{\mathfrak{M}} \in N_a(w)$$

There is an important respect in which using a language like the one we’ve outlined, and like the one Walton uses in his reconstruction, is best described as *Anselmian*, and *not* Anselm’s actual ideas (beyond the surface difference that Anselm never gave this type of formalism). Anselm explicitly allows as answers to the question “What is he doing?” only atomic actions and negations of atomic actions. Because our language allows *any* type of formula to be substituted in for p in $\delta_a p$, this system cannot be taken as being a reconstruction of Anselm’s

⁴He then quotes Anselm, and footnotes this with a reference to the Lambeth fragment, and notes that a translation of the fragment by Ernst Van Haagen was “scheduled for publication in the *American Philosophical Quarterly*” (p. 199), but I have unfortunately not been able to find any further record of this publication.

⁵We make no assumption about any of the properties of these agents, other than that they are agents, in as weak a sense as possible. This is in line with what we discussed at the end of §2.

actual ideas.⁶ However, because Anselm himself says that the answer to “What is he doing?” can be *any* verb, this extension of our syntax is not unreasonable, because it makes just as much sense to say “He is reading and sitting” and “He makes it the case that if he reads he is sitting” as it does to say “He is reading” or “He is sitting”.

After Walton introduces his syntax, he considers different possible candidate theorems for a logic of Anselmian agency. The first he proposes is both necessary and obvious:

Axiom 6.3 (Success). $\delta_a p \rightarrow p$

This is the agentive parallel to the axiom T introduced in the previous section. It’s intuitive plausibility follows from the fact that after agent a does so that p , then p must be the case, for otherwise, you’re saying that a succeeded in bringing about p , even though p is still false, which makes no sense. Beyond its intuitive plausibility, there is second reason to adopt this axiom. This axiom implies $\neg\delta_a p \vee \neg\delta_a \neg p$, which in turn is equivalent to $\neg(\delta_a p \wedge \delta_a \neg p)$, the truth of which is required for the relations in the square of opposition to hold.⁷

Next Walton considers the following pair of potential axioms:

Proposition 6.4 (Conjunction Elimination). $\delta_a(p \wedge q) \rightarrow (\delta_a p \wedge \delta_a q)$

Proposition 6.5 (Conjunction Introduction). $(\delta_a p \wedge \delta_a q) \rightarrow \delta_a(p \wedge q)$

These are converses of each other. Walton argues that we cannot accept both of these as axioms or theorems. He claims that adding

$$\delta_a(p \wedge q) \equiv (\delta_a p \wedge \delta_a q)$$

is too strong, because this equivalence plus the T axiom is provably equivalent to the standard normal modal logic T [Walton 1976b, p. 303, fn. 17]. He says that this is unacceptable because T , being a normal modal logic and hence an extension of K , both proves versions of the paradoxes of strict implication and also validates the rule of necessitation RN . From an agentive point of view, RN violates intuitions that we have about agency and tautologies. It should not be the case that any agent can *cause* it to be the case that a tautologous state of affairs is obtained. Such states of affairs will obtain vacuously, whether or not we ever do anything, and even in spite of our actions. The problems with this rule also apply to adopting either $\delta_a p \equiv p$ (material equivalence) or $\delta_a p \leftrightarrow p$ (strict or causal equivalence) as theorems.

Walton is wrong in rejecting the acceptance of both Proposition 6.4 and Proposition 6.5 out of hand, for two reasons. The first is that

$$\delta_a(p \wedge q) \equiv (\delta_a p \wedge \delta_a q) + \delta_a p \rightarrow p$$

is equivalent to

$$\delta_a(p \rightarrow q) \rightarrow (\delta_a p \rightarrow \delta_a q) + \delta_a p \rightarrow p$$

only in the presence of the further axiom $\delta_a \top$. Without $\delta_a \top$, RN is not sound. If we wanted to take both Proposition 6.5 and Proposition 6.4 as axioms, we

⁶Walton is aware of this: “St. Anselm did not, to my knowledge, take the next step that would be of interest to a student of modern sentence logic, namely extension to conjunctive, disjunctive, and materially conditional states of affairs” [Walton 1976b, p. 301].

⁷As noted in [Seegerberg 1992, p. 349].

can do so without sacrificing our intuitions about doing. The resulting logic is *EMCT*.

The second reason is that his objection to *RN* relies on a certain narrow conception of agency. Under such a narrow conception, agency is always active and causal. But insisting that we interpreted Latin *facere* as ‘to cause’ is too restrictive. If we remember that the analysis of *facere* is an analysis of doing, not of causation, then it wouldn’t seem that unreasonable if someone said ‘agent *a* does such that $p \vee \neg p$ ’. In fact, I myself am doing such that an infinite number of tautologies are true. Here is case where the ordinary usage (*usus loquendi*) of terms contradicts some intuitions about their potentially more narrow logical functions. For insofar as tautologies are necessary, $\neg\delta_a\neg p$ (‘it is not the case that *a* brings it about that not *p*’) will always be true when *p* is a tautology; and then, as mentioned earlier, it *does* follow that $\delta_a p$ holds whenever *p* is a theorem. If we are interested in the logical properties of *facere* at the possible expense of ordinary usage, then the necessitation rule is unacceptable and we must look elsewhere for axioms and rules. If, however, we are interested in explaining in logical terms our ordinary usage of *facere*, as Anselm appears to be doing, then *T* presents itself as a most plausible choice.

That being said, we will continue to focus on the more strictly logical, rather than common usage, analysis of doing. Walton concludes, incorrectly, that one of Proposition 6.4 and Proposition 6.5 must be given up. He gives up the latter, because this is the route taken in [Fitch 1963]⁸, but his argument for accepting Proposition 6.4 is simply to state what it says, and note that adopting it plus axiom *T* “would give us the rudiments of a seemingly not very contentious, if rather minimal, system of agency” [Walton 1976b, p. 302]. But the same could be said if we took Proposition 6.5 instead of Proposition 6.4.

After accepting Axiom 6.3 and Proposition 6.4, Walton next proposes, and quickly rejects, the following:

Proposition 6.6. $(\delta_a p \wedge (p \rightarrow q)) \rightarrow \delta_a q$

His reason for rejecting this is that this axiom is even stronger than Propositions 6.5 and 6.4 combined. In this he is correct, both in his rejection of the principle and his reason for doing so. Proposition 6.6 is stronger than the axiom *K*, as it implies $(\delta_a p \wedge q) \rightarrow \delta_a q$. This is clearly too strong, so Proposition 6.6 should be rejected.

Instead, Walton offers a version of the *K* axiom as an alternative to Proposition 6.4:

Axiom 6.7. $(\delta_a p \wedge \delta_a (p \rightarrow q)) \rightarrow \delta_a q$

He says that the system combining Axiom 6.3 with Proposition 6.7 is stronger than that containing just Axiom 6.3 and Proposition 6.4, because Proposition 6.7 implies Proposition 6.4 but that “the converse implication does not seem to hold. [The claim] is inconclusive, in the absence of a δ_a -semantics” [Walton 1976b, p. 304]. As we noted earlier, he is wrong in saying that Proposition 6.7 implies Proposition 6.4; it does so only in the presence of the further

⁸Fitch gives no argument for why we should take this over Proposition 6.5. He claims outright that he’s assuming it’s true: “We assume that the following concepts, viewed as classes of propositions, are closed with respect to conjunction elimination: striving (for), doing, believing, knowing, proving” [Fitch 1963, p. 137]. He makes no argument for the truth of this assumption.

axiom $\delta\top$, which we have reason to reject when modeling the proper, logical usage of *facere*. However, now that we have provided a type of δ_a -semantics, we can confirm that his second claim is correct; Proposition 6.4 does not imply axiom 6.7.

Finally, Walton puts forward one further possible axiom or theorem:

Proposition 6.8 (Causal implication). $(\delta_a p \wedge (p \hookrightarrow q)) \rightarrow \delta_a q$

(Here, we use \hookrightarrow to represent causal implication; this is not to be confused with the usage of \leftrightarrow earlier, where we specifically refrained from giving any meaning to the connective.) The reason that this proposition is formulated as causal implication instead of just standard implication is because Walton wishes to block $(\delta_a p \wedge (p \rightarrow q)) \rightarrow \delta_a q$ as a theorem, as this implies $(\delta_a p \wedge q) \rightarrow \delta_a q$, which has as an unfortunate instance the following: “If Socrates scratches his head and Plato dies, then Socrates brings it about that Plato dies” [Walton 1976b, p. 304]. Walton discusses this theorem in the context of agency *per aliud*. Agency of this type only becomes relevant when we are trying to give an analysis of the *usus loquendi* of the term *facere*; it plays no role in the analysis of the strict logical usage of the term. A full analysis of the improper usage of the term is much more difficult, and as it is one best left to the grammarian and linguist, we do not pursue it further here.

7 Conclusions

There is one outstanding issue that must be mentioned before we draw our conclusions. There is a relevant sense in which Walton’s approach, in developing the syntax and leaving any questions of semantics behind, more adequately captures what is found in the Anselmian texts, and in which my semantical proposal is inherently, irredeemably, and perhaps pathologically anachronistic. As Serene notes, Anselm in the texts discussed above “presents the modes as a disjunctive *necessary condition* for ascriptions of agency, but he does not to my knowledge assert that any relationship, no matter how remote, between a subject and a state of affairs provides a sufficient condition for agency” [Serene 1983, p. 146] (emphasis added). This is a crucial feature of his theory. If there were such a sufficient condition for ascriptions of agency, then given how encompassing his theory of action is, it would be possible to make practically every person (or indeed, every object) an agent for every action, because failure to act counts, in his theory, as action. With the ensuing consequence such a move would have for personal culpability and sin, this is clearly a move we do not want to take.

Walton’s syntax can be seen as an extension of Anselm’s necessary conditions, in which any ascription of agency will have one of twenty distinct possible syntactic constructions. When we add semantics, we are essentially adding sufficient conditions; we can say that when such-and-such conditions hold, we can then make a true statement about agency.

In no way, then, should the discussions in the preceding two sections be taken as a formalization of *what Anselm said*, as it is not. The formalizations should be viewed as inspired by, and hopefully capturing, the brilliant insights of his theory of agency. We have seen that the breadth of Anselm’s conception of agency is a point in its favor, and not a reason for discard. Further, Walton’s extension of Anselm’s discussion of agency to non-atomic actions seems thoroughly plausible,

and we can provide both syntax and semantics to accommodate this extension. If we are content to divorce the logical theory from any ethical theory, there is no problem with the addition of semantics from the formal point of view.

With that caveat expressed, we can end with the following. Depending on specific ideas about agency, there are a number of different choices for logics:

T The normal multi-modal logic T , which has as axioms both $\delta_a(p \wedge q) \equiv (\delta_a p \wedge \delta_a q)$ and $\delta_a \top$, corresponds to at least some aspects of our ordinary usage of the word *facere*.

EMT, ECT These both block the unwanted inference of $\delta_a \top$, which is desirable from the standpoint of the logical usage of *facere*, as well as barring the equivalence found in T , thus satisfying the syntax provided by Walton.

EMCT This blocks the unwanted inference of $\delta_a \top$ but allows for the equivalence noted above, for which the only argument against was the incorrect claim that it caused the logic to collapse into T , and for which arguments for can be provided.

We have shown that each of these logics is characterized by a class of models, and hence that each system is sound, and that they are all distinct. Which one should be preferred, in our opinion, depends on the context of usage.

8 Acknowledgments

The author would like to acknowledge the participants of the ILLC project on ancient and medieval logic, January 2007, and those of the Square of Opposition International Congress, June 2007, for stimulating and useful discussion.

References

- [Belnap et al. 2001] Belnap, N., Perloff, M., & M. Xu. (2001). *Facing the Future*. (Oxford: Oxford University Press).
- [Chellas 1980] Chellas, B.F. (1980). *Modal Logic: An Introduction*. (Cambridge: Cambridge University Press).
- [Danto 1973] Danto, A.C. (1973). *Analytical Philosophy of Action*. (Cambridge: Cambridge University Press).
- [Fitch 1963] Fitch, F.B. (1963). A Logical Analysis of Some Value Concepts. *Journal of Symbolic Logic* 28, no. 2, 135–142.
- [Henry 1953] Henry, D.P. (1953). St. Anselm and the Varieties of Doing. *Theoria* 19, 178–185.
- [Henry 1960] Henry, D.P. (1960). The Scope of the Logic of St. Anselm. (In L. de Raeymaeker (Ed.), *L'homme et son destin d'après les penseurs du Moyen Âge: actes du premier Congrès international de*

- philosophie médiévale, Louvain-Bruzelles, 28 août-4 septembre 1958* (pp. 377–383). Louvain: Nauwelaerts, Paris: Béatrice-Nauwelaerts, 1960.)
- [Henry 1967] Henry, D.P. (1967). *Logic of Saint Anselm*. (Oxford: Oxford University Press).
- [Hopkins 1972] Hopkins, J. (1972). *A Companion to the Study of St. Anselm*. (Minneapolis: University of Minnesota Press).
- [Horn 1989] Horn, L. (1989). *A Natural History of Negation*. (Chicago: University of Chicago Press).
- [King –] King, P. (St.) Anselm of Canterbury [1033–1109], forthcoming in *Encyclopedia of Philosophy* 2nd Ed., preprint available at <http://individual.utoronto.ca/pking/articles/Anselm.EP.pdf>
- [Schmitt 1936] Schmitt, F.S. (1936). *Ein neues unvollendetes werk des hl. Anselm von Canterbury*. (Aschendorff: Münster i W.)
- [Serene 1980] Serene, E. (1980). Anselm’s Modal Conceptions. (In S. Knuuttila (Ed.), *Reforging the Great Chain of Being* (pp. 117–162) Dordrecht: D. Reidel Publishing Company.)
- [Serene 1983] Serene, E. (1983). Anselmian Agency in the *Lambeth Fragments*: A Medieval Perspective on the Theory of Action. *Anselm Studies, An Occasional Journal* 1, 143–156.
- [Seegerberg 1992] Seegerberg, K. (1992). Getting Started: Beginnings in the Logic of Action. *Studia Logica* 51, nos. 3/4, 347–478.
- [Southern & Schmitt 1969] Southern, R.W. & Schmitt, F.S. (Eds.) (1969). *Memorials of St. Anselm Auctores Britannici Medii Aevi I*. (London: Oxford University Press for the British Academy), specifically p. 334ff.
- [Walton 1976a] Walton, D. (1976a). Logical Form and Agency. *Philosophical Studies* 29, 75–89.
- [Walton 1976b] Walton, D. (1976b). St. Anselm and the Logical Syntax of Agency. *Franciscan Studies* 36, 298–312.