

MODALITY AND AXIOMATIC THEORIES OF TRUTH I: FRIEDMAN-SHEARD

JOHANNES STERN

Munich Center for Mathematical Philosophy, LMU Munich

Abstract. In this investigation we explore a general strategy for constructing modal theories where the modal notion is conceived as a predicate. The idea of this strategy is to develop modal theories over axiomatic theories of truth. In this first paper of our two part investigation we develop the general strategy and then apply it to the axiomatic theory of truth Friedman-Sheard. We thereby obtain the theory Modal Friedman-Sheard. The theory Modal Friedman-Sheard is then discussed from three different perspectives. First, we show that Modal Friedman-Sheard preserves theoremhood modulo translation with respect to modal operator logic. Second, we turn to semantic aspects and develop a modal semantics for the newly developed theory. Third, we investigate whether the modal predicate of Modal Friedman-Sheard can be understood along the lines of a proposal of Kripke, namely as a truth predicate modified by a modal operator.

§1. Introduction. In philosophical logic, there are two ways of formalizing modal notions. The first way, the orthodox approach, conceives of modalities as sentential operators which applied to a sentence yield a new sentence. The second, alternative way conceives of modalities as sentential predicates which applied to the name of a sentence yield a new sentence. This second approach has the advantage that it respects the fundamental intuition that truth and the modal notions are expressions of the same grammatical category and should therefore be treated uniformly. The predicate approach also ties in well with the so-called relational analysis of propositional attitudes and, finally, has a greater expressive strength than the operator approach, at least as this latter approach is standardly conceived.

Despite the initial appeal of the predicate approach to modality only few attempts have been made so far in developing a workable and philosophically attractive approach of this kind. The problem we face when we take up this task is that the seemingly impeccable and intuitive modal principles we know from modal operator logic lead straight to paradox if we adopt them in the predicate setting. In this two-part investigation we propose a strategy that overcomes these difficulties and apply this strategy in developing two particular modal theories, the theory Modal Friedman-Sheard and the theory Modal Kripke-Feferman. In a nutshell our strategy is to characterize the modal notion viz. the modal predicate by its interaction with the notion of truth viz. the truth predicate. In doing so we can analyze the modal paradoxes as manifestations of the well-known paradoxicality of the concept of truth. Consequently, or so we shall argue, if we adopt a consistent system of truth, we arrive at a consistent system of truth and modality.

In this first paper of our investigation we lay out the general strategy we apply in developing the modal theories and we construct the theory Modal Friedman-Sheard. This

Received: August 8, 2013.

newly developed theory Modal Friedman-Sheard will then be shown to satisfy a minimal adequacy condition we think an interesting modal theory ought to fulfill. We show that the theory preserves theoremhood modulo translation with respect to modal operator logic. A modal theory should satisfy this criterion for otherwise one would take a revisionist stance with respect to important work in philosophical and mathematical logic. If one is aiming toward a general modal theory that is neutral with respect to the particular philosophical theory of modality adopted, we think it is advisable to avoid revisionist positions in connection with established work.

We then turn to more semantic aspects and provide, building upon work by Gupta (1982), Herzberger (1982), Belnap & Gupta (1993), and Halbach (1994), a novel modal semantics, namely modal revision semantics. As the name indicates the semantics combines ideas from revision semantics with possible world semantics for modal operator logic. We show that the theory Modal Friedman-Sheard axiomatizes in some relevant sense this very semantics. That is, Modal Friedman-Sheard stands in the same relation to modal revision semantics as the theory of truth Friedman-Sheard stands to ordinary revision semantics: Modal Friedman-Sheard axiomatizes modal revision semantics up to the first transfinite ordinal ω .

Finally, we take up a third theme and investigate whether the modal predicate we have axiomatized can be understood along the lines of an idea put forward by Kripke in his seminal “Outline of a Theory of Truth” Kripke (1975). Kripke suggested defining (or understanding) a modal predicate by a truth predicate which has been modified by a modal operator, that is, the modal operator acts as a predicate modifier on the truth predicate to yield a modal predicate. In the case of an alethic modality the proposal would amount to understanding the predicate ‘is necessary’ (N) to be an abbreviation for the complex predicate ‘is necessarily true’ ($\Box T$).

Kripke’s proposal is sometimes understood as a strategy for recovering the expressive strength of the predicate approach within the operator setting and, consequently, as an argument in favor of operator approaches to modality.¹ We do not think that this is a good understanding of Kripke’s proposal because at best the proposal serves as an argument in favor of treating modal notions as predicate modifiers rather than as sentential operators. Furthermore, understanding modal notions as operators violates one of the main grammatical constraints on formal treatments of modality which is to treat truth and the modalities in a grammatical uniform way.

In fact, Kripke himself advocated a very different understanding of his proposal. He viewed his proposal as a vindication of predicate approaches to modality against the back-drop of the liar-like paradoxes that threaten these approaches:

Ironically, the application of the present approach [his theory of truth] to languages with modal operators may be of some interest to those who dislike intensional operators and possible worlds and prefer to take predicates true of sentences (or sentence tokens). (...) Now if a necessity operator and a truth predicate are allowed, we could define a necessity predicate $Nec(x)$ applied to sentences, either by $\Box T(x)$ or by $T(\Box x)$ [slight change of notation] according to taste (...).²

¹ Cf. Halbach & Welch (2009) for an understanding of Kripke’s proposal along these lines.

² See Kripke (1975, p. 713/14).

Following Kripke we investigate whether the primitive modal predicate we have introduced can be understood in this way. As our study is carried out from a more proof-theoretic perspective we also try to answer this question in a proof-theoretic way. To this end we formulate the theory of truth Friedman-Sheard in a modal operator language with a truth predicate and show that the theory Modal Friedman-Sheard is reducible to this very theory. That is, if we translate the modal predicate by the truth predicate modified by the modal operator and keep the remaining vocabulary fixed, then theoremhood is preserved modulo translation from the theory Modal Friedman-Sheard to the theory of truth Friedman-Sheard as formulated in the modal operator language.³ The success of this reduction which we label the “Kripke reduction” provides a conceptual insight into why the strategy of characterizing the modal predicate by appeal to its interaction with the truth predicate works so well—at least in the present case. In applying this strategy we have axiomatized the modal predicate in a way that allows for an understanding of the modal predicate as defined by the truth predicate and a modal operator. Since the modal operator we consider behaves in a standard way the reduction shows that if the modal predicate is axiomatized in the way we have proposed there remains only one potentially paradoxical concept, that is the concept of truth. Consequently, if we guarantee the consistency of the theory of truth we adopt we get consistency of our modal theory for free.

1.1. Structure of the paper. The paper structures as follows. In the first section after the introduction we lay out the general strategy and methodology we employ in constructing our modal theory. The modal theory, that is Modal Friedman-Sheard, is presented in the next section in which we also show that the theory is adequate from the perspective of modal operator logic. In a further section we turn to semantic considerations and present modal revision semantics. We end the discussion of semantic aspects of our approach by establishing the semantic adequacy result we have alluded to before. That is, we show that Modal Friedman-Sheard axiomatizes modal revision semantics up to the first transfinite ordinal ω . In the last technical section we present the details of the “Kripke reduction” we have briefly outlined above. The paper ends with a very short summary of the results we have obtained and an outlook to future work.

1.2. Notation. In what is to come we let \mathcal{L}_{PA} be the language of arithmetic containing the logical constants \neg , \wedge , and \forall , the identity symbol $=$ and the usual arithmetical vocabulary 0 , S , $+$, \times . Moreover, we assume the language to contain a finite number of function symbols for certain primitive recursive functions such as the substitution function. Additional logical constants that we use in our study are taken to be defined in the usual way. We take \mathcal{L}_{PAT} (\mathcal{L}_{PAN}) to be the extension of \mathcal{L}_{PA} by a one-place predicate T (N). Correspondingly, \mathcal{L}_{PATN} is \mathcal{L}_{PA} extended by two one-place predicates T and N . PA is the theory Peano arithmetic supplemented by the defining equations for the function symbols of the language. PAT (PAN , $PATN$) is PA in the language \mathcal{L}_{PAT} (\mathcal{L}_{PAN} , \mathcal{L}_{PATN}) where we allow for formulas of the extended language in the induction scheme. In general, if we speak of a theory in an extended language we also allow formulas of the extended language to appear in the induction scheme. Similarly, if we speak of a theory in an extended language, we also tacitly assume the syntactical predicates and function symbols

³ A similar reduction has recently been carried out by Halbach & Welch (2009). Yet, Halbach and Welch’s investigation is purely semantic in character as they show the translation to preserve truth between two intended semantics.

occurring in the axioms of the theory to be replaced by the corresponding predicates and function symbols of the extended language.

We assume some appropriate coding scheme for expressions of \mathcal{L}_{PA} and its extensions and denote the code, that is, the Gödel number of an expression ζ by $\#\zeta$. The numeral of $\#\zeta$ will be conveyed by $\lceil \zeta \rceil$. We let the sets $Sent_{\mathcal{L}}$ (“ \mathcal{L} -sentences”), $Term_{\mathcal{L}}$ (“ \mathcal{L} -terms”), and $Cterm_{\mathcal{L}}$ (“ \mathcal{L} -closed terms”) represent themselves and, moreover, drop the subscript when no confusion may arise. In most cases, if \triangleright is a syntactic operation we represent it by \triangleright . However, there are a few exceptions to this rule: $num^\bullet(\cdot)$ represents the function that yields the Gödel number of the numeral of the argument as an output; we represent the ternary substitution function by $x(y/v)$ and, moreover, let $Val(\cdot)$ represent the function that takes codes of closed terms as arguments and provides their denotation as output. Importantly, $Val(\cdot)$ is not a function symbol of the language. We follow Halbach (2011) in writing $\forall t(\phi(t))$ as abbreviation of the lengthy $\forall x(Cterm_{\mathcal{L}}(x) \rightarrow \phi(x))$ and, similarly, abbreviate $\lceil \phi(x) \rceil(t/\lceil x \rceil)$ by $\lceil \phi(t) \rceil$ where, again, t is considered to be a name of a closed term of the language.

§2. The strategy. In this section we outline our general strategy which we will then implement in developing the theory Modal Friedman-Sheard. If we treat modalities as predicates and allow for self-referential sentences in our language we run into paradoxes similar to those of the notion of truth. And, indeed, as has been pointed out by Montague (1963) we cannot adopt the most basic principles of modal operator logic within the predicate setting for sake of paradox:

THEOREM 2.1 (Montague). *Let Σ be a theory extending Robinson arithmetic in \mathcal{L}_{PAN} and ‘ N ’ a one-place predicate. If for every sentence $\phi \in \mathcal{L}_{PAN}$*

- (i) $\Sigma \vdash N[\phi] \rightarrow \phi$
- (ii) $\Sigma \vdash \phi \Rightarrow \Sigma \vdash N[\phi]$

then Σ is inconsistent.

Now the principles of modal operator logic corresponding to (i) and (ii) above

- (T) $\Box\phi \rightarrow \phi$
- (Nec) $\frac{\phi}{\Box\phi}$

are of course used in axiomatizing many philosophically salient modal notions. Therefore, it seems we fall short of providing a philosophically satisfactory treatment of modality within the predicate setting. As a matter of fact, this has led many to conclude that we cannot have the most intuitive modal principles once we treat modalities as predicates.⁴ Yet, once we introduce a truth predicate into the picture, that is, once we work in a setting where we have an interacting truth and modal predicate at our disposal there may be a possible way out of this dilemma. For what seems to motivate the modal principle (T) and thus condition (i) in Theorem 2.1 is the plausibility of

“Everything that is necessary is true.”

⁴ See Stern (forthcoming) for a careful discussion of Montague’s theorem and its philosophical consequences.

However, having the truth predicate at our disposal this natural language sentence appears to be best formalized within the predicate setting by

$$(T') \quad \forall x (Sent(x) \rightarrow (Nx \rightarrow Tx))$$

And whether paradox arises once (T') is assumed instead of condition (i) remains open and depends on the properties one ascribes to the truth predicate. The main difference between (i) and (T') is of course that where in (i) a sentence is disquoted, that is, we eliminate an occurrence of a sentential predicate, no disquotation takes place within (T') . Rather, disquotation is licensed only if the properties of the truth predicate allow us to apply disquotation to (an instance of) the consequent of (T') .

This reaction to Montague's paradox is of course congenial to the prominent philosophical idea that all the liar-like paradoxes have a common origin, namely the paradoxicality of the concept of truth. For example Horsten (2002) writes:

“There is a view which holds that the relation between these paradoxes [the modal paradoxes, J.S.] goes deeper. The idea is that there is an underlying *conceptual connection* between the notions [the modal notions; J.S.] involved which explains *why* they are all paradoxical. This view is usually combined with the belief that at bottom there is only *one* paradoxical concept: truth. All the Liar-like paradoxes are just manifestations of the paradoxicality of the concept truth.” (See Horsten (2002, p. 215); the italics are due to Horsten.)

While we do not wish to embrace the view outlined by Horsten, it seems worthwhile to take it seriously and to see where it leads us. In a way the view formally amounts to avoiding introduction and elimination of the modal predicate independently of the truth predicate. The idea is that if we have, say, a sentence $N[\phi]$ we should not stipulate a principle or rule that allows us to infer ϕ . And the same holds for the converse direction. Instead, the transition between a sentence in its quoted form within a sentence $N[\phi]$ and the assertion of the sentence itself should always go via the detour of the truth predicate, that is, the sentence $T[\phi]$. Indeed spelled out in this way the view we have outlined seems to be well motivated if one adopts the deflationist idea that quotation and disquotation are the function of the truth predicate. Consequently, quotation and disquotation of sentences is not the task of the modal predicate and in formulating modal principles we should therefore avoid the introduction or elimination of the modal predicates without the detour via the truth predicate.⁵ In the remainder of the paper we adopt the idea that the transition between a sentence in its quoted form and the assertion of the sentence itself should always go via the detour of the truth predicate as our general guideline or strategy in formulating the modal theories.

Whereas (T') fits well with the view we have just outlined, condition (i) of Theorem 2.1 eliminates an occurrence of the modal predicate without taking the detour via the truth predicate. Thereby the sentence to which the modal predicate was applied gets disquoted. This suggests adopting (T') as the predicate version of the principle (T) of modal operator logic. Similarly, the second assumption used in establishing Montague's paradox, that is the predicate version of the rule of necessitation, does not fit well with the view outlined as this time the modal predicate is introduced without taking the detour via the truth predicate.

⁵ The idea of connecting the view we have outlined to deflationism was suggested to us by Martin Fischer.

Rather in alignment with the above view we should modify the rule of necessitation to the effect that we are allowed to infer $N[\phi]$ whenever we have derived $T[\phi]$. This would leave us with the rule

$$(T-Nec) \quad \frac{T[\phi]}{N[\phi]}$$

If we consider the characteristic principles (4) and (E) of the philosophically most interesting modal operator logics *S4* and *S5*,⁶ that is

$$\begin{aligned} (4) \quad & \Box\phi \rightarrow \Box\Box\phi \\ (E) \quad & \neg\Box\phi \rightarrow \Box\neg\Box\phi \end{aligned}$$

in their straightforward formulation once we take the modal notion to be a predicate we arrive at the following principles⁷

$$\begin{aligned} (P4) \quad & \forall t (Sent(Val(t)) \rightarrow (NVal(t) \rightarrow N \dot{N} t)) \\ (PE) \quad & \forall t (Sent(Val(t)) \rightarrow (\neg NVal(t) \rightarrow N \dot{\neg} N t)) \end{aligned}$$

But again the principles (P4) and (PE) violate the condition we have proposed because both principles introduce an additional modal predicate without taking the detour via the truth predicate. In doing so the modal sentence asserted in the antecedens of the principles gets quoted in the consequence of the principles. Yet, we can bring (P4) and (PE) in compliance with our general strategy by pushing the modal sentence in the antecedens of (P4) and (PE) in the scope of the truth predicate. Thus (P4) and (PE) transform into the modal principles

$$\begin{aligned} (4') \quad & \forall t (T \dot{N} t \rightarrow N \dot{N} t). \\ (E') \quad & \forall t (T \dot{\neg} N t \rightarrow N \dot{\neg} N t) \end{aligned}$$

Moreover, in general, it seems that the modal properties which are meant to be expressed by the principles of modal operator logic (T), (4), and (E), and the rule (Nec) are also conveyed by the principles (T'), (4'), and (E'), and the rule (T-Nec) as the truth predicate should not have any impact in this respect. We also think that the principles (T'), (4'), and (E') and the rule (T-Nec) are supported by our natural language intuition to the same degree the principles of modal operator logic are. Of course these kind of natural language intuitions should be met with some care in an area where paradox is lurking behind every corner. However, there seems to be nothing problematic about these principles and there seems to be no reason why we should avoid them or no longer pursue the formal consequences of the view outlined above.

Now, if we pursue this strategy, we intimately bind our theory of modality to the properties of the truth predicate. Depending on which properties one ascribes to the truth

⁶ The modal operator logics *S4* and *S5* are normal modal logics, that is both extend classical propositional logic by the modal principle (K) (cf. below) and are closed under necessitation (Nec). *S4* extends the basic normal logic *K* (which precisely consists of the propositional tautologies, the modal principle (K) and the rule (Nec)) by the modal axiom (T) and (4) while *S5* extends *S4* by the modal axiom (E). For more on modal operator logic, see Blackburn et al. (2001) and, Fitting & Mendelsohn (1998).

⁷ As one of the advantages of the predicate approach is to quantify over the argument position of the modal notion at stake by the means of simple first-order quantification, we replace schematic formulations by their universally quantified counterparts.

predicate, that is, which axiomatic theory of truth one adopts, a different theory of modality will arise. In the remainder of the paper we shall then test our strategy with respect to the axiomatic theory of truth Friedman-Sheard. We will therefore investigate what modal theory Friedman-Sheard gives rise to, if supplemented by the modal axioms (T') , $(4')$, and (E') , and the rule $(T-Nec)$. We apply the strategy to the theory of truth Kripke-Feferman in a sequel to this paper. This should allow us to see how widely applicable our strategy is and to compare the modal theories arising from the two different theories of truth.

Besides these three modal principles and the rule $(T-Nec)$ we shall also adopt the predicate version of the basic axiom of normal modal operator logic

$$(K) \quad \Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi).$$

That is, we adopt the principle

$$(K_N) \quad \forall x, y (Sent(x \rightarrow y) \rightarrow (N(x \rightarrow y) \rightarrow (Nx \rightarrow Ny)))$$

For rather instrumental reasons we also take the modal principles

$$(Reg_N) \quad \forall x, v, s, t (Sent(\forall vx \rightarrow (Val(s) = Val(t) \rightarrow (Nx(s/v) \leftrightarrow Nx(t/v))))$$

$$(ND) \quad \forall s, t (Val(s) \neq Val(t) \rightarrow Ns \neq t)$$

$$(BF) \quad \forall x, v (Sent(\forall vx \rightarrow (\forall t Nx(t/v) \rightarrow N \forall vx)))$$

to be axioms of the modal theory. (Reg_N) is a formalized version of the substitution principle and forces the modal context to be referentially transparent. (ND) , “necessity of distinctness”, asserts that if the value of two terms is distinct, then the two terms will necessarily have distinct values. (BF) is a formalized ω -rule which allows us to infer from the necessity of every instantiation of an universally quantified formula to the necessity of the formula itself. A more philosophical gloss on (BF) is that it tells us that we have names (or terms) in the language for everything that could possibly exist. While this is surely not a plausible assumption on the general scheme, it seems acceptable once we work within an arithmetical setting and, indeed, the three mentioned axioms are tailor made for a framework in which we appeal to an arithmetical theory as our base and syntax theory: it is often thought that arithmetical truths are necessary truths, that the terms of the arithmetical language refer rigidly and that what could possibly arithmetically exist, does exist.

In fact, we wish the arithmetical truths to come out as necessary. Incidentally, this was the main reason for adopting (Reg_N) , (ND) , and (BF) . We want arithmetical truth to turn out necessary because we want our syntax theory to be rigid: we need to guarantee that the names of the syntactical objects and operations do in fact always designate these objects and operations. Otherwise we would not be sure whether the modal axioms convey the meaning we intend them to convey. If, for example, a numeral codes a closed term in certain circumstances but a parenthesis, say, in others then clearly the axioms $(4')$ and (E') would no longer express the modal properties we intend them to convey, but would instead be nonsensical.⁸

Whereas the previous principles all have been theory-unspecific, that is they should be axioms of any modal theory no matter which axiomatic theory of truth one assumes, there

⁸ The reader might wonder why we have added (ND) and (BF) to our list of axioms but have omitted the principles

$$(NI) \quad \forall s, t (Val(s) = Val(t) \rightarrow Ns \equiv t)$$

$$(CBF) \quad \forall x, v (Sent(\forall vx \rightarrow (N \forall vx \rightarrow \forall t Nx(t/v)))$$

will also be theory-specific principles. For instance in the present case we consider a modal principle which fits the theory Modal Friedman-Sheard we are about to develop but which would seem implausible if an alternative theory of truth were assumed. We will turn to this very principle once we present the details of the theory Modal Friedman-Sheard.

§3. The theory Modal Friedman-Sheard. We shall now begin the construction of the modal theory we obtain, if the theory “Friedman-Sheard” is assumed as our underlying theory of truth. We first recall the essentials of the theory Friedman-Sheard.

3.1. Friedman-Sheard. The theory “Friedman-Sheard” first appeared as system *D* in Friedman and Sheard’s “An Axiomactical Approach to Self-Referential Truth” Friedman and Sheard (1987) and was further discussed by Sheard (1994, 2001) and Halbach (1994) who coined the name *FS*. The following axiomatization of *FS* is due to Halbach (2011):

DEFINITION 3.1 (FS). *The theory FS consists of all axioms of PA in the language \mathcal{L}_{PAT} and the following axioms*

- (FS1) $\forall s \forall t (T(s \doteq t) \leftrightarrow Val(s) = Val(t))$
 (FS2) $\forall x (Sent(x) \rightarrow (T \neg x \leftrightarrow \neg Tx))$
 (FS3) $\forall x, y (Sent(x \wedge y) \rightarrow (T(x \wedge y) \leftrightarrow Tx \wedge Ty))$
 (FS4) $\forall v \forall x (Sent(\forall v x) \rightarrow (T(\forall v x) \leftrightarrow \forall t (Tx(t/v))))).$

FS is closed under the rules

$$(T\text{-Elim}) \quad \frac{T[\phi]}{\phi} \qquad (T\text{-Intro}) \quad \frac{\phi}{T[\phi]}$$

for all $\phi \in Sent_{\mathcal{L}_{PAT}}$.

The distinguishing feature of *FS* is the axiom (FS2) together with the fact that it is closed under *T-Intro* and *T-Elim*. This leads to the truth predicate being consistent and complete and implies that the inner and the outer logic of *FS* are identical. Thus *FS* is symmetric and the truth predicate of *FS* is normal in the sense of modal operator logic, that is, K_T

$$(K_T) \quad \forall x, y (Sent(x \rightarrow y) \rightarrow (T(x \rightarrow y) \rightarrow (Tx \rightarrow Ty)))$$

is a theorem of *FS*.

THEOREM 3.2. *The theory FS is consistent.*

Proof. See Friedman & Sheard (1987). Later in this paper we sketch an alternative proof due to Halbach (1994, 2011). \square

A final noteworthy fact concerning *FS* is that it proves the global reflection principle for *PAT*, that is

$$(Ref_T) \quad \forall x (Sent(x) \rightarrow (Bew_{PAT}(x) \rightarrow Tx))$$

is a theorem of *FS*.

that is “Necessity of Identity” and “The Converse Barcan Formula”. These two principles, however, will be provable—or at least pertinent versions of these principles—in the theories we are about to construct.

3.2. Modal Friedman-Sheard. The modal theory arising from the theory of truth FS in the language \mathcal{L}_{PATN} by addition of the modal principles previously discussed will be called “Modal Friedman-Sheard” (MFS). Besides these modal principles we introduce one further theory specific interaction axiom, namely

$$(IA) \quad \forall t (N \dot{T} t \leftrightarrow T \dot{N} t)$$

The gloss on (IA) in natural language is “every sentence for which it is necessary that it is true it is true that it is necessary and vice versa” which has some initial plausibility and seems philosophically acceptable.⁹ More importantly, (IA) does not introduce or eliminate the truth or the modal predicate and therefore seems to fit well with the general theoretical design of FS which precisely avoids axioms of this type.

As we shall see (IA) plays an important theoretic role in our investigation and can be considered as the central axiom governing the interaction of the modal and the truth predicate in the context of FS independently of the particular modal properties assumed. We now define the theory MFS :

DEFINITION 3.3 (MFS). MFS consists of all axioms of FS in the language \mathcal{L}_{PATN} ¹⁰ and the axioms

$$(Reg_N) \quad \forall x, v, s, t (Sent(\forall v x) \rightarrow (Val(s) = Val(t) \rightarrow (Nx(s/v) \leftrightarrow Nx(t/v))))$$

$$(ND) \quad \forall s \forall t (Val(s) \neq Val(t) \rightarrow N(s \neq t))$$

$$(BF) \quad \forall v, x (Sent(\forall v x) \rightarrow (\forall t Nx(t/v) \rightarrow N \dot{\forall} vx))$$

$$(K_N) \quad \forall x, y (Sent(x \rightarrow y) \rightarrow (N(x \rightarrow y) \rightarrow (Nx \rightarrow Ny)))$$

$$(IA) \quad \forall t (N \dot{T} t \leftrightarrow T \dot{N} t)$$

$$(T') \quad \forall x (Sent(x) \rightarrow (Nx \rightarrow Tx))$$

$$(4') \quad \forall t (T \dot{N} t \rightarrow N \dot{N} t)$$

$$(E') \quad \forall t (T \dot{\neg} \dot{N} t \rightarrow N \dot{\neg} \dot{N} t).$$

In addition to the rules of FS , the following rule is adopted.

$$(T-Nec) \quad \frac{T[\phi]}{N[\phi]}$$

for all $\phi \in Sent_{\mathcal{L}_{PATN}}$.

The theory MFS is consistent, indeed it is \mathcal{L}_{PA} -conservative over FS .

⁹ Of course, (IA) is plausible only if we “rigidify” our syntax theory so that the name of a sentence ϕ rigidly denotes the sentence ϕ . For otherwise, for example, the sentence ‘ $2 + 2 = 4$ ’ could mean that grass is green and thus while it would be true that ‘ $2 + 2 = 4$ ’ is necessary it would not be necessary that ‘ $2 + 2 = 4$ ’ is true. Of course, similar problems would affect the modal axioms $(4')$ and (E') . As we have pointed out, these problems were the main reason why we insisted on the fact that arithmetics, that is syntax theory, should come out as necessary true. From a more philosophical perspective rigidifying the syntax theory amounts to reading a language parameter into the modal predicate (“is necessary in English”, “is necessary in \mathcal{L}_{PATN} ” instead of “is necessary” simpliciter). See, for example, Peacocke (1976) for a discussion of these problems and a solution along our lines.

¹⁰ As pointed out in section 1.2 this means that the syntactical predicates and function symbols occurring in the axioms of FS are now the corresponding expressions of \mathcal{L}_{PATN} . Moreover, the induction scheme has been extended to \mathcal{L}_{PATN} .

THEOREM 3.4. *MFS is a \mathcal{L}_{PA} -conservative extension of FS.*

Proof. Using the recursion theorem¹¹ we may define a translation function $g : \mathcal{L}_{PATN} \rightarrow \mathcal{L}_{PAT}$ which translates the modal predicate ‘ N ’ by the truth predicate and which keeps the remaining vocabulary fixed. It is easy to verify that the translations of all axioms of *MFS* are theorems of *FS* and also straightforward to see that proofs in *MFS* translate to proofs in *FS*. That is, we have for all $\phi \in \mathcal{L}_{PATN}$:

$$MFS \vdash \phi \Rightarrow FS \vdash g(\phi) \quad \square$$

REMARK 3.5. *The interpretation g is not faithful! For consider $\forall x, y (Sent(x \vee y) \rightarrow (N(x \vee y) \rightarrow Nx \vee Ny))$ which translates to $\forall x, y (Sent_{\mathcal{L}_{PATN}}(x \vee y) \rightarrow (T(x \vee y) \rightarrow Tx \vee Ty))$. The latter is a theorem of *FS* though the former is not a theorem of *MFS*.*

The reader might find Theorem 3.4 disappointing, as it does not only show the consistency but also the conservativeness of *MFS* over *FS* whereas initially one could have hoped that the interaction of the two predicates increases the deductive power, that is, the proof-theoretic strength, of the theory as was the case in the examples studied by Halbach (2001, 2002). On the other hand, one may argue that any reasonable conception of an alethic modality allows for the trivial reading, that is it allows for a reading of the modal notion as truth. From this perspective Theorem 3.4 is not only unproblematic but desirable. Remark 3.5 is in line with this latter view since obviously the trivial reading is not a faithful reading of an alethic modality.

Moreover, it is unlikely that we will be able to achieve a (nontrivial) proof-theoretic strengthening of *MFS* from the interaction of the truth and the modal predicate. It turns out that virtually every introduction or elimination principle produces inconsistency once it is added to the axioms of *MFS*. We provide some sample cases in the theorem below but the reader is invited to check further introduction and elimination principles in this respect.

THEOREM 3.6. *Let Σ be a modal theory which proves all axioms of *MFS*, is closed under (T-Elim), (T-Intro), and (T-Nec) and which proves one of the following*

- (i) $\forall t (Sent(Val(t)) \rightarrow (T \dot{N} t \rightarrow \neg N \dot{\neg} Val(t)))$
- (ii) $\forall x (Sent(x) \rightarrow (Nx \rightarrow N \dot{\neg} \dot{N} num^\bullet(\dot{\neg} x)))$
- (iii) $\forall t (Sent(Val(t)) \rightarrow (NVal(t) \rightarrow T \dot{T} t))$
- (iv) $\forall t (Sent(Val(t)) \rightarrow (N \dot{N} t \rightarrow TVal(t)))$.

Then Σ is inconsistent.

We show the latter theorem by appeal to the following lemma.

LEMMA 3.7. *Let (i)–(iv) be as in Theorem 3.6. Then*

- (a) $MFS + (ii) \vdash (i)$
- (b) $MFS + (iii) \vdash (ii)$
- (c) $MFS + (iv) \vdash (ii)$.

Proof. We deal with the three items in turn. Throughout we assume ‘ t ’ to be the code of a closed term. Accordingly, if we write $\phi(t)$ we mean $Cterm(x) \rightarrow \phi(x)$. Notice that

¹¹ For an exposition of the recursion theorem see, for example, Monk (1976).

when we apply necessitation we may derive $Cterm(x) \rightarrow N[\phi(x)]$ using (K) , (Nec_N) and Lemma 3.9.

Case (a)

1. $Sent(Val(s)) \rightarrow (NVal(s) \rightarrow N \neg N num^\bullet(\neg Val(s)))$ (ii)
2. $Sent(Val(s)) \rightarrow (NVal(s) \rightarrow T \neg N num^\bullet(\neg Val(s)))$ 1, (T')
3. $Sent(Val(s)) \rightarrow (NVal(s) \rightarrow \neg T N num^\bullet(\neg Val(s)))$ 2, $(FS2)$
4. $Sent(Val(s)) \rightarrow (T N num^\bullet(\neg Val(s)) \rightarrow \neg NVal(s))$ 3
5. $Sent(\neg Val(t)) \rightarrow (T N num^\bullet(\neg \neg Val(t)) \rightarrow \neg N \neg Val(t))$
 $(T N num^\bullet(\neg \neg Val(t)) \rightarrow \neg N \neg Val(t))$ 4, $s \doteq num^\bullet(\neg Val(t))$
6. $Sent(Val(t)) \rightarrow (T N t \rightarrow \neg N \neg Val(t))$ $FS, (Nec_N), (K_N), (Reg_N)$
7. $\forall t(Sent(Val(t)) \rightarrow (T N t \rightarrow \neg N \neg Val(t)))$ 6

Case (b)

1. $Sent(Val(t)) \rightarrow (NVal(t) \rightarrow T T t)$ (iii)
2. $Sent(Val(t)) \rightarrow (\neg T T t \rightarrow \neg NVal(t))$ 1
3. $Sent(Val(t)) \rightarrow (T T num^\bullet(\neg Val(t)) \rightarrow \neg NVal(t))$ 2, $(FS2)$
4. $Sent(Val(t)) \rightarrow (N N num^\bullet(\neg Val(t)) \rightarrow \neg NVal(t))$ 3, (T') , (IA)
5. $Sent(Val(t)) \rightarrow (NVal(t) \rightarrow \neg N N num^\bullet(\neg Val(t)))$ 4
6. $Sent(Val(t)) \rightarrow (NVal(t) \rightarrow \neg T N num^\bullet(\neg Val(t)))$ 5, $(4')$
7. $Sent(Val(t)) \rightarrow (NVal(t) \rightarrow T \neg N num^\bullet(\neg Val(t)))$ 6, $(FS2)$
8. $Sent(Val(t)) \rightarrow (NVal(t) \rightarrow N \neg N num^\bullet(\neg Val(t)))$ 7, (E')
9. $\forall x(Sent(x) \rightarrow (Nx \rightarrow N \neg N num^\bullet(\neg x)))$ 8, $x \doteq Val(t)$ ¹²

Case (c)

1. $Sent(Val(s)) \rightarrow (N N s \rightarrow T Val(s))$ (iv)
2. $Sent(Val(s)) \rightarrow (\neg T Val(s) \rightarrow \neg N N s)$ 1
3. $Sent(Val(s)) \rightarrow (T \neg Val(s) \rightarrow \neg T N s)$ 2, $FS, (4')$
4. $Sent(Val(s)) \rightarrow (N \neg Val(s) \rightarrow T \neg N s)$ 3, $FS, (T')$
5. $Sent(Val(s)) \rightarrow (N \neg Val(s) \rightarrow N \neg N s)$ 4, (E')
6. $Sent(Val(\neg t)) \rightarrow (N \neg Val(\neg t) \rightarrow N \neg N \neg t)$ 5, $s \doteq num^\bullet(\neg x)$
7. $Sent(\neg x) \rightarrow (N \neg \neg x \rightarrow N \neg N num^\bullet(\neg x))$ 6
8. $Sent(x) \rightarrow (Nx \rightarrow N \neg N num^\bullet(\neg x))$ 7
9. $\forall x(Sent(x) \rightarrow (Nx \rightarrow N \neg N num^\bullet(\neg x)))$ 8

□

Proof of Theorem 3.6. By Lemma 3.7 it suffices to show (i). We have

1. $\delta \leftrightarrow \neg N[\delta]$ diagonal lemma
2. $\neg \delta \leftrightarrow N[\delta]$ 1
3. $T[\neg \delta \leftrightarrow N[\delta]]$ 2, $(T-Intro)$
4. $N[N[\delta]] \leftrightarrow N[\neg \delta]$ 3, $(T-Nec), (K_N)$
5. $T[N[\delta]] \rightarrow N[\neg \delta]$ 4, $(4')$
6. $T[N[\delta]] \rightarrow \neg N[\neg \delta]$ (i)
7. $\neg T[N[\delta]]$ 5, 6
8. $T[\neg N[\delta]]$ 7, $(FS2)$
9. $\neg N[\delta]$ 8, $(T-Elim)$
10. δ 1, 9
11. $T[\delta]$ 10, $(T-Intro)$
12. $N[\delta]$ 11, $(T-Nec)$

□

Due to this inconsistency result we suggest that the proponent of proof-theoretically strong modal theories should adopt a strategy in formulating these modal theories which is different from the one we have proposed.

¹² Note that we have $PATN \vdash Sent(x) \rightarrow Cterm(num^\bullet(x))$ which allows us to get rid of $Cterm(x)$ which was our implicit antecedens condition in lines 1–8.

3.2.1. Subtheories of MFS . In formulating MFS we have not focused on the question whether the axioms are mutually independent and, indeed, they are not. The right-to-left direction of (IA) is provable by appeal to the remaining axioms, in particular by the axioms (T') and $(4')$. We nonetheless included the biconditional in our axiomatization, as we will be interested in subtheories of MFS in which (T') and $(4')$ are not available, but we always want (IA) to be one of the axioms. Most importantly, we are interested in the subtheory “Basic Modal Friedman-Sheard” ($BMFS$):

DEFINITION 3.8 (BMFS). *$BMFS$ consists of all axioms and rules of MFS except (T') , $(4')$, and (E') . The modal theory $BX_1 \dots X_nMFS$ is the theory resulting from $BMFS$ by addition of principles X_1 to X_n to the theory.*

According to Definition 3.8, MFS and $BT'4'E'MFS$ are one and the same theory.

We shall now discuss some consequences and state some theorems of MFS and $BMFS$ respectively. First, note that since $(T-Intro)$ is a rule of $BMFS$ the standard rule of necessitation is an admissible rule of $BMFS$ and thus MFS . Moreover, due to the rule of $(T-Elim)$ the rule of conecessitation is also admissible if we add (T') to the axioms of $(BMFS)$; that is, conecessitation is admissible in MFS . Thus MFS is closed under the two following rules:

$$(Nec_N) \quad \frac{\phi}{N[\phi]} \qquad (Conec_N) \quad \frac{N[\phi]}{\phi}$$

for all $\phi \in Sent_{\mathcal{L}_{PATN}}$.

Importantly, syntax theory, that is, PA turns out to be “rigid” in $BMFS$ as we have intended it to be.

THEOREM 3.9. *For all $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PA}$ we have:*

$$BMFS \vdash \forall t_1, \dots, t_n (\phi(Val(t_1), \dots, Val(t_n)) \rightarrow N[\phi(t_1, \dots, t_n)])$$

Proof of Theorem 3.9. By induction over the positive built-up of $\phi(x_1, \dots, x_n)$. See Stern (2012). \square

Amongst other things the proof of Theorem 3.9 hinges on the modal axioms (Reg_N) , (ND) , and (BF) and as we have indicated this was our principal reason for adding these principles to the axioms of $BMFS$ and MFS respectively. We take it to be an interesting and important question to develop versions of $BMFS$ and MFS which use alternative syntax theories and which could thus be formulated without appeal to (Reg_N) , (ND) , and (BF) in their full generality, but this is a topic for possible future research.

3.2.2. Contingent modal theories. So far we have constructed modal theories over a theory of truth and a given syntax theory, which we took to be PA . Within this setting no contingent vocabulary is added to the language and it might seem that we have failed to construct a philosophically interesting modal theory, as the interesting characteristics of the modal notions only arise in the presence of contingent vocabulary. However, the main reason for this omission is that nothing but further complication is gained with respect to our particular investigation if contingent vocabulary is added to the language. For this reason we shall stick to the framework we have presented so far in the remainder of the paper, yet we sketch out how a contingent modal theory may be constructed.

To this end we consider the extension of the language \mathcal{L}_{PATN} by a contingent n -place predicate P^n . We call this language \mathcal{L}_C . We try to keep the framework as simple as

possible, so we only consider extensions of the language by one predicate, but an extension by multiple predicates would follow the same outlines.¹³ In the presence of the contingent predicate P^n the underlying theory of truth needs to be supplemented by an additional axiom dealing with this predicate:

$$(FSC) \quad \forall t_1, \dots, t_n (P^n Val(t_1), \dots, Val(t_n) \leftrightarrow T P t_1, \dots, t_n)$$

Let Π be a modal theory in the language \mathcal{L}_{PATN} . We call Π in the language \mathcal{L}_C supplemented by $(FSC) \Pi^C$. The contingency of the newly introduced predicate P^n will depend on the contingency of the theory Σ axiomatizing it. In particular, this means that the rules introducing the modal predicate are only applicable if a formula has been derived within the corresponding theory Π^C without appeal to the axioms or rules of the contingent theory Σ . If $(T-Nec)$ is the only rule introducing the modal predicate, the theory $\Sigma + \Pi^C$ will be the theory consisting of all the rules and axioms of Π^C and Σ but where the rule $(T-Nec)$ is restricted in the afore mentioned way.

After these preliminary observations about MFS , its subtheories and how the theories relate to modal theories with contingent vocabulary we shall now show that MFS is an adequate modal theory, in the sense that it preserves theoremhood with respect to the modal operator logic $S5$.

3.3. Modal operator logic and MFS . We show that the translations of all theorems of $S5$ are theorems of MFS . Implicitly we even provide a more general result. We show—at least for the modal principles we have discussed—that the modal theory $BMFS + X'$ preserves theoremhood modulo translation with respect to a given normal modal logic KX . As modal operator logic is well entrenched in mathematical and philosophical practice this criterion seems to be a minimal requirement that an interesting modal theory should fulfill.¹⁴

However, one disclaimer seems appropriate at this point. Throughout this paper we shall only be concerned with *de dicto* modality and ignore problems arising with respect to *de re* modality.¹⁵ This will show in the formation rules of the modal operator language we are about to construct over \mathcal{L}_{PA} since we only allow the modal operator to apply to sentences of the language and not to arbitrary (open) formulas. As a matter of fact the modal operator language will be a multimodal language since we introduce a further modal operator to account for truth, that is, we consider ‘ \Box ’ as the modal operator and ‘ \blacksquare ’ as the truth operator.

DEFINITION 3.10. *The multimodal operator language $\mathcal{L}_{PA}^{\Box\blacksquare}$ consists of the vocabulary of \mathcal{L}_{PA} together with the modal operators ‘ \Box ’ and ‘ \blacksquare ’. The formation rules of $\mathcal{L}_{PA}^{\Box\blacksquare}$ are given by:*

¹³ We do not introduce new names to the language because these could of course be considered as one-place predicates.

¹⁴ Amongst others, des Rivières & Levesque (1986) already showed that the translation of the deductive closure of a consistent modal operator logic into \mathcal{L}_{PAN} forms a consistent set. Yet, to our knowledge no axiomatization of these sets has been provided by means of principled and unrestricted axioms within a language containing a modal predicate.

¹⁵ As we have set up things there is no such thing as *de re* modality in the predicate approach. Quine (1977) and Belnap & Gupta (1993) (cf. 237/38) suggest treating the modal predicate as a two-place predicate applying to the code of a formula and the code of a sequence of objects (thus similar to a satisfaction predicate) in order to recover *de re* reading in this context. See also Halbach & Welch (2009) for remarks to this effect.

$$\phi ::= \psi \mid \neg\phi \mid \Box\phi \mid \blacksquare\phi \mid \phi \wedge \phi$$

with $\psi \in \text{Sent}_{\mathcal{L}_{PA}}$.

Next we define a translation from $\mathcal{L}_{PA}^{\Box\blacksquare}$ to \mathcal{L}_{PATN} .

DEFINITION 3.11. Let τ be a function from $\mathcal{L}_{PA}^{\Box\blacksquare}$ to \mathcal{L}_{PATN} . τ is called a translation function iff τ is regular, that is, it commutes with the logical connectives, is the identity function on \mathcal{L}_{PA} and respects the conditions:

$$\tau(\phi) := \begin{cases} T[\tau(\psi)] & , \text{ if } \phi \doteq \blacksquare\psi \\ N[\tau(\psi)] & , \text{ if } \phi \doteq \Box\psi \end{cases}$$

Before we can give the main theorem of this section we need to say a bit more about the modal operator logic we shall work with. Above we stated our aim as showing that *MFS* preserves theoremhood with respect to the modal operator logic *S5*, but this obviously only accounts for the modal principles of the modality under consideration but not for truth. In the modal operator setting we take truth, that is the operator ' \blacksquare ', to be aptly characterized by the Tarski biconditionals, that is the modal principle

$$(Tr) \quad \blacksquare\phi \leftrightarrow \phi$$

THEOREM 3.12. Let *PAS5* be the theory resulting from the addition of the modal logic *S5* axiomatizing the \Box -operator and the modal principle *(Tr)* axiomatizing \blacksquare to the axioms of *PA*. Furthermore, let τ be a translation function in the sense of Definition 3.11. Then for all $\phi \in \mathcal{L}_{PA}^{\Box\blacksquare}$

$$PAS5 \vdash \phi \Rightarrow MFS \vdash \tau(\phi).$$

For the proof of Theorem 3.12 we rely heavily on the following lemma which states that for all sentences of \mathcal{L}_{PATN} in the range of τ we can derive the Tarski biconditionals. We omit a detailed proof of this lemma.

LEMMA 3.13. If τ is a translation in the sense of Definition 3.11, then for all $\phi \in \mathcal{L}_{PA}^{\Box\blacksquare}$

$$BMFS \vdash T[\tau(\phi)] \leftrightarrow \tau(\phi)$$

Proof. By induction on the complexity of ϕ . □

Proof of Theorem 3.12. The theorem is established by an induction on the length of a proof in *PAS5*. For sake of illustration we give the proof in full.

- (i) We need to show that the translation of all the axioms of *S5* are provable in *MFS*. This is obvious for the arithmetical axioms. For the remaining axioms we observe:

1. By Lemma 3.13 we have $MFS \vdash \tau(\blacksquare\phi \leftrightarrow \phi)$ which settles the case of *(Tr)*.
2. *(T)*—By *(T')* we have for all $\phi \in \mathcal{L}_{PA}^{\Box\blacksquare}$, $MFS \vdash N[\tau(\phi)] \rightarrow T[\tau(\phi)]$. Again using Lemma 3.13 we derive $MFS \vdash N[\tau(\phi)] \rightarrow \tau(\phi)$.
3. *(4)*— $MFS \vdash T[N[\tau(\phi)]] \rightarrow N[N[\tau(\phi)]]$ by *(4')* and as before Lemma 3.13 allows us to terminate the proof of this case.
4. *(E)*—As for *(T')* and *(4')*, but using *(E')* and Lemma 3.13.

- (ii) Obviously, MFS is closed under *modus ponens*. Moreover, since MFS is closed under necessitation if $MFS \vdash \tau(\phi)$, then $MFS \vdash N[\tau(\phi)]$ and thus $MFS \vdash \tau(\Box\phi)$. \square

From the perspective of modal operator logic the system MFS therefore proves to be an adequate modal theory and as such MFS satisfies the minimal adequacy condition we expect a modal theory to fulfill.¹⁶ We now turn toward the discussion of semantic aspects of MFS .

§4. Semantics for MFS . In adopting the theory FS as our underlying theory of truth we have excluded, or so it seems, the very possibility of providing an interesting class of models or semantics for our theory MFS . Since FS is a subtheory of MFS , MFS is ω -inconsistent¹⁷ and we therefore cannot interpret the arithmetical vocabulary of the theory in its intended way, that is we cannot construct a model for MFS based on the standard model of arithmetic.

However, at the beginning of this paper we have already pointed to a result which suggests that, nonetheless, there might be an interesting semantic story to be told about MFS . As shown by Halbach (1994), FS axiomatizes revision semantics up to the first transfinite ordinal ω . The hope would be that a similar result holds for MFS and a suitable modal version of revision semantics. Since our investigation will generalize and extend Halbach's result to the modal setting we briefly present his work.

4.1. Revision Semantics and FS . We present some ideas of revision semantics and explain its connection to FS . We closely follow Halbach (1994, 2011). The basic idea of revision semantics is often presented as follows.¹⁸ Starting from an arbitrary interpretation of the truth predicate ("an initial guess") we obtain better and better interpretations of the truth predicate by a suitable series of revisions. From a formal point of view this revision process can be characterized by appeal to a revision operator which we call a classical jump. The classical jump takes a subset of ω , that is, a possible interpretation of the truth predicate, as an argument and provides a new subset of ω , that is, the revised interpretation of the truth predicate, as an output.

DEFINITION 4.1 (Classical Jump). Let Γ be an operation from $P(\omega)$ to $P(\omega)$. Γ is called a classical jump iff for every $S \subseteq \omega$

$$\Gamma(S) := \{\# \phi : (\mathbb{N}, S) \models \phi\}$$

where $\phi \in \text{Sent}_{\mathcal{L}_{PAT}}$.

With this definition at hand one can define iterative applications of the revision operator $\Gamma^n(S)$, by letting $\Gamma^0(S) = S$, $\Gamma^{n+1}(S) = \Gamma(\Gamma^n(S))$, and thereby illustrate the revision process on the finite ordinals.

¹⁶ One might wonder whether τ is in fact a faithful interpretation of $PAS5$ into MFS . The answer is negative for a trivial reason: FS , and as a consequence MFS , is arithmetically stronger than PA , however, $PAS5$ is not. Since τ is the identity function on the arithmetical vocabulary, the existence of arithmetical sentences provable in MFS but not $PAS5$ rules out the possibility of τ being a faithful interpretation. For example, $PAS5 \not\vdash \text{Con}_{PA}$ but $MFS \vdash \tau(\text{Con}_{PA})$ where by definition $\tau(\text{Con}_{PA}) = \text{Con}_{PA}$.

¹⁷ See McGee (1985) for a proof of the ω -inconsistency of FS (and thus MFS).

¹⁸ For an in depth presentation of revision semantics cf. Belnap & Gupta (1993).

Halbach (1994, 2011) suggests a change of perspective with respect to the revision process. Instead of viewing the revision process as a way of obtaining better and better interpretations of the truth predicate, we consider it to be a process of eliminating unsuitable interpretations. The idea is that when we start the revision process every subset of ω is a possible interpretation of the truth predicate. But by application of the jump operation we exclude certain subsets of ω as possible interpretations of the truth predicate. The idea can be captured by defining a revision operator on sets of sets of natural numbers, that is on subsets of $P(\omega)$. To this end let $M \subseteq P(\omega)$ and set

$$\Gamma[M] := \{\Gamma(S) : S \in M\}$$

Iterative applications of Γ to subsets M of $P(\omega)$, $\Gamma^n[M]$, are defined as to be expected: $\Gamma^0[M] = M$, $\Gamma^{n+1}[M] = \Gamma[\Gamma^n[M]]$. With this second revision operator at hand we can already sketch out the connection between FS and revision semantics: Every subset of ω that is a member $\Gamma^n[P(\omega)]$ is a suitable interpretation of the subtheory of FS which results from FS by restricting the number of applications of the rules (T -Intro) and (T -Elim) to $n - 1$. As a consequence these subtheories of FS have nice standard models.

In order to state this semantic adequacy result in a precise way we define the mentioned subtheories of FS :

DEFINITION 4.2 (Subtheories of FS). *The subtheories FS_n of FS are defined as follows. $FS_0 := PAT$, FS_1 is the theory consisting of all axioms of FS together with the reflection principle (Ref_T) for PAT and FS_{n+1} is FS except that the number of applications of (T -Intro) and (T -Elim) in a proof is limited to n .*

We can now specify the link between revision semantics and FS .

THEOREM 4.3 (Halbach). *For all $n \in \omega$ and all $S \subseteq \omega$*

$$S \in \Gamma^n[P(\omega)] \Leftrightarrow (\mathbb{N}, S) \models FS_n$$

Proof. See Halbach (1994, 2011). □

This theorem leaves us with a method for constructing standard models for subsystems of FS . In the next section we provide an analogous result for MFS with respect to modal revision semantics we shall now develop.

4.2. Modal revision semantics. The idea behind modal revision semantics is to combine revision semantics with ideas from possible world semantics for modal operator logic. Ultimately this leads to a relativization of the jump operation to a modal frame. Besides the frame we also introduce an evaluation function into the picture, a function that assigns to every world an interpretation of the truth predicate. A frame and an evaluation function induce a model for \mathcal{L}_{PATN} relative to each world by taking the interpretation of the modal predicate to be the intersection of the interpretations of the truth predicate at all accessible worlds. Implicitly, the jump operation is then applied to these models and we thereby obtain a new interpretation of the truth predicate at each world which then yields a new evaluation function. Accordingly, the modal revision operator works on evaluation functions relative to a frame.

Let us sketch out these introductory remarks and provide the formal definitions.¹⁹

¹⁹ Semantic theories of truth and possible world semantics have been combined by several authors in constructing a semantics for modal predicates. See, for example Asher & Kamp (1989),

DEFINITION 4.4 (Frame, evaluation function). Let $W \neq \emptyset$ be a set of labeled natural number structures²⁰ and $R \subseteq W \times W$ a dyadic relation on W . Then $F = \langle W, R \rangle$ is called a frame. A function $f : W \rightarrow P(\omega)$ is called an evaluation function for a frame F . The set of all evaluation functions of a frame F is denoted by Val_F .

An evaluation function is a function that provides an interpretation (extension) of the truth predicate at every world w .

DEFINITION 4.5 (Models for \mathcal{L}_{PATN}). If F is a frame and f an evaluation function, then F and f induce a model $M_w = \langle w, f(w), Y_w \rangle$ of the language \mathcal{L}_{PATN} relative to every world w . $f(w)$ is the extension of the truth predicate at w and Y_w where

$$Y_w = \bigcap_{v \in [wR]} f(v)$$

is the extension of the necessity predicate.²¹ $[wR]$ is short for the set $\{v \in W : wRv\}$. We often write $F, w \models_f \phi$ to convey the fact that $M_w \models \phi$ whenever M_w is the model induced by the frame F and the evaluation function f relative to w . We write $F, f \models \phi$ if and only if $\forall w (F, w \models_f \phi)$.

DEFINITION 4.6 (Modal Classical Jump). Let Val_F be the set of evaluation functions of a frame F . The modal classical jump Γ_F is an operation on Val_F such that for all $w \in W$

$$[\Gamma_F(f)](w) = \{\# \phi : F, w \models_f \phi\}$$

Also, set $\Gamma_F^0(f) = f$ and $\Gamma_F^{n+1}(f) = \Gamma_F(\Gamma_F^n(f))$.

After these initial definitions we can now basically rerun Halbach's construction that we have sketched in the preceding section. As in the case of standard revision semantics we need to view the modal revision process as a process where more and more evaluation functions get ruled out. We therefore extend the definition of Γ_F and allow for applications of Γ_F to sets of valuation functions of F .

DEFINITION 4.7. Let Γ_F be a jump and $M \subseteq Val_F$ a set of valuation functions on F . Then

$$\Gamma_F[M] = \{\Gamma_F(f) : f \in M\}.$$

As before we define iterative applications of Γ_F by $\Gamma_F^0[M] = M$ and $\Gamma_F^{n+1}[M] = \Gamma[\Gamma^n[M]]$.

The fact that we can view the modal revision process as a process of eliminating unsuitable evaluation functions depends on the antitonicity of the jump operation on Val_F relative to a frame F , that is, the jump eliminates more and more evaluation functions from Val_F . Actually, due to the ω -inconsistency of FS (and therefore MFS and $BMFS$) there will be no evaluation functions left at the stage ω , that is $\Gamma_F^\omega[Val_F] = \emptyset$ (cf. Halbach

Belnap & Gupta (1993), Halbach et al. (2003), Leitgeb (2006), and Halbach & Welch (2009). Asher and Kamp, Belnap and Gupta, and Halbach et al. also work in the setting of the revision theory of truth, although our approach is somewhat different.

²⁰ If we would allow for a contingent predicate P^n , as discussed in section 3.2.2, a world would be a tuple (\mathbb{N}, X) with $X \subseteq \omega^n$. X would serve as the interpretation of P^n at this world.

²¹ We assume \bigcap to be an operation on $P(\omega)$. Thus in particular if $[wR] = \emptyset$, then $\bigcap_{v \in [wR]} f(v) = \omega$.

(1994, 2011) for more on this).²² We now show that the modal jump is indeed antitone on Val_F .

LEMMA 4.8 (Antitonicity). *For all $n, m \in \omega$ with $n \leq m$ and any frame F*

$$\Gamma_F^m[Val_F] \subseteq \Gamma_F^n[Val_F]$$

Proof. The proof is by induction on n . We omit the proof as it consists in adopting Halbach's proof (cf. Halbach, 2011, p. 165) to the modal setting. \square

Further we introduce some notational conventions and define the subsystems of $BMFS$ and MFS analogously to the case of FS .

DEFINITION 4.9. *For all $n \in \omega$ if for all evaluation functions $f \in \Gamma_F^n(F, w \models_f \phi)$ ($n \in \omega$) we write $F, w \models^n \phi$. If $n = 0$ we drop the index and simply write $F, w \models \phi$. Similarly, if $\forall w (F, w \models^n \phi)$ we write $F \models^n \phi$.*

DEFINITION 4.10 (Subsystems of $BMFS$). *The subsystems $BMFS_n$ for $n \in \omega$ of $BMFS$ are defined as follows. We set $BMFS_0$ to be $PATN$ and $BMFS_1$ to consist of all axioms of $BMFS$ and the rule (T -Nec) together with the reflection principle (Ref_T) for $PATN$. Finally we define $BMFS_{n+1}$ to be just like $BMFS$ except that the number of applications of the rules (T -Intro) and (T -Elim) in a proof is limited to n .*

The systems $BX_1 \dots X_m BMFS_n$ are defined correspondingly with $BX_1 \dots X_m MFS_0 = BMFS_0 = PATN$.

We can now establish the central result linking the modal theories arising from FS to modal revision semantics. In its most general form the result links the theory $BMFS$ to modal revision semantics: we show that $BMFS$ axiomatizes modal revision semantics up to the first transfinite ordinal ω . While this result is concerned with $BMFS$ only, the theorem also implies a nice adequacy result for MFS .

THEOREM 4.11. *Let F be a frame and Γ_F a modal classical jump. Then for all $n \in \omega$*

$$f \in \Gamma_F^n[Val_F] \Leftrightarrow F \models_f BMFS_n$$

Proof. The proof is a rather straightforward generalization of Halbach's proof of Theorem 4.3 and goes by induction on n . The left-to-right direction of the claim can be checked by a sub-induction on the length of a proof in $BMFS_n$. For the converse direction, we construct an evaluation function g such that $\Gamma_F(g) = f$. If we can show that g is an evaluation function at the previous level, we can use the induction hypothesis and the definition of Γ_F to conclude to the desired.

We start with the case $n = 0$. In this case there is nothing to show because any structure of \mathcal{L}_{PATN} based on the natural number structure is a model of $PATN$ and thus of $BMFS_0$. Since $\Gamma_F^0[Val_F] = Val_F$ we are done.

For $n = 1$, we first show that if $f \in \Gamma_F^1[Val_F]$, then $\forall w \in W(F, w \models_f BMFS_1)$. This is established by a sub-induction on the length of a derivation in $BMFS_1$, and is left to the reader. For the converse direction, closely following Halbach (2011), we assume $f \in Val_F$ and $\forall w \in W(F, w \models_f BMFS_1)$ and construct an evaluation g such that $\Gamma_F(g) = f$. Since $g \in Val_F = \Gamma_F^0[Val_F]$ it then follows that $f \in \Gamma_F^1[Val_F]$. We set g

²² We take $\Gamma_F^\omega[M]$ to be $\bigcap_{n \in \omega} \Gamma_F^n[M]$.

to be the evaluation function such that for all $w \in W$:

$$g(w) := \{t^{\mathbb{N}} : \#Tt \in f(w)\}^{23}$$

Obviously $g \in Val_F$, that is, g is an evaluation function on F . We now claim that $\Gamma_F(g) = f$. The latter claim is established by an induction on the built up of sentences of \mathcal{L}_{PATN} . Again this is routine and follows the outlines of Halbach (2011). Nonetheless we present the case of ' Nt ' as this illustrates the important role of the axiom (IA).

$$\begin{aligned} \#Nt \in f(w) &\Leftrightarrow F, w \models_f T[Nt] && \text{Def. 4.5} \\ &\Leftrightarrow F, w \models_f N[Tt] && (IA) \\ &\Leftrightarrow \forall v(wRv \Rightarrow \#Tt \in f(v)) && \text{Def. 4.5} \\ &\Leftrightarrow \forall v(wRv \Rightarrow t^{\mathbb{N}} \in g(v)) && \text{Def. } g \\ &\Leftrightarrow F, w \models_g Nt && \text{Def. 4.5} \\ &\Leftrightarrow \#Nt \in [\Gamma_F(g)](w) && \text{Def. 4.6} \end{aligned}$$

This concludes the proof for the case $n = 1$.

It remains to show the induction step, that is the case $n + 1$. This again consists in basically rewriting the proof of Theorem 3.12 in Halbach (1994, 2011). As induction hypothesis we may assume that

$$f \in \Gamma_F^n[Val_F] \Leftrightarrow \forall w \in W(F, w \models_f BMFS_n)$$

For the left-to-right direction assume then that $f \in \Gamma_F^{n+1}[Val_F]$. By antitonicity of the jump on Val_F we know that $f \in \Gamma_F^n[Val_F]$ and thus by induction hypothesis we have $\forall w \in W(F, w \models_f BMFS_n)$. It suffices then to show that one additional application of (*T-Intro*) and (*T-Elim*) preserves validity relative to f . In the case of (*T-Intro*) we know that $BMFS_n \vdash \phi$ and therefore by induction hypothesis $\forall w(F, w \models^n \phi)$. But then by Definitions 4.5 and 4.6, $\forall w \in W(F, w \models^{n+1} T[\phi])$.

For the case of (*T-Elim*) we assume that $BMFS_n \vdash T[\phi]$. By induction hypothesis and Definition 4.5, $\# \phi \in f(w)$ for all $w \in W$ and all $f \in \Gamma_F^n$. By Definition 4.6 we have $\forall w \in W(F, w \models_g \phi)$ with $g \in \Gamma_F^{n-1}[Val_F]$. But by antitonicity $\Gamma_F^{n+1}[Val_F] \subseteq \Gamma_F^{n-1}[Val_F]$ and thus for all $f \in \Gamma_F^{n+1}[Val_F]$, $\forall w \in W(F, w \models_f \phi)$.

For the converse direction we copy our construction of step $n = 1$. The proof goes through unaltered except that we need to show that the newly construed evaluation function g is a member of $\Gamma_F^n[Val_F]$. If we show that $\forall w \in W(F, w \models_g BMFS_n)$, we can infer to the desired using the induction hypothesis. To this end we observe that if $BMFS_n \vdash \phi$, then $BMFS_{n+1} \vdash T[\phi]$ and thus $\forall w \in W(F, w \models_f T[\phi])$. But since $\Gamma_F(g) = f$ it follows that $\forall w \in W(F, w \models_g \phi)$. This ends our proof. \square

Theorem 4.11 nicely ties $BMFS$ and modal revision semantics together and, as we have already mentioned, it also a nice adequacy result for MFS itself. It turns out that MFS axiomatizes modal revision semantics up to ω relative to equivalence frames, that is frames where the accessibility relation is an equivalence relation. This observation turns out to be a corollary of a more general phenomenon which should appear familiar from possible world semantics for modal operator logic:

²³ $t^{\mathbb{N}}$ stands for the denotation (the value) of t in the natural number structure.

THEOREM 4.12. *Let F be a frame and $n \geq 1$. Then*

- (i) F is reflexive $\Leftrightarrow F \models^n BT' MFS_n$
- (ii) F is transitive $\Leftrightarrow F \models^n B4' MFS_n$
- (iii) F is Euclidean $\Leftrightarrow F \models^n BE' MFS_n$.

Sketch of a proof. The left-to-right direction follows the outlines of the standard proofs for modal operator logic (see e.g., Blackburn et al., 2001). For the converse we sketch the proof of item (i): suppose for *reductio* that there is a $w \in W$ with $\neg wRw$. Let κ be the truth teller that is, the sentence saying of itself that it is true. Now let $f \in Val_F$ be an evaluation with $\kappa \in f(v)$ for all $v \in W$ with wRv but $\neg \kappa \in f(w)$. By properties of the jump Γ_F we know that for all evaluation functions g , all worlds w and all $n \in \omega$

- (a) $\kappa \in [\Gamma_F^n(g)](w) \Rightarrow \kappa \in [\Gamma_F^{n+1}(g)](w)$
- (b) $\neg \kappa \in [\Gamma_F^n(g)](w) \Rightarrow \neg \kappa \in [\Gamma_F^{n+1}(g)](w)$

Now, clearly $\Gamma_F^n(f) \in \Gamma_F^n[Val_F]$ for $n \in \omega$ and $F, w \models_{\Gamma_F^n(f)} N[\kappa]$. But on the other hand we have $F, w \not\models_{\Gamma_F^n(f)} T[\kappa]$ which contradicts the assumption. The converse direction of items (ii) and (iii) works similarly. \square

COROLLARY 4.13. *Let F be a frame and $n \geq 1$. Then*

$$F \text{ is an equivalence frame} \Leftrightarrow F \models^n MFS_n.$$

This corollary concludes the discussion of semantic aspects of MFS and thus the second theme of our paper.²⁴ It shows that MFS nicely ties to modal revision semantics—a semantics that combines revision semantics and possible world semantics. Unsurprisingly, there is no hope for strengthening the foregoing results in the direction of results for $BMFS$ and MFS itself rather than for their subtheories as both theories are ω -inconsistent. We move on to the last theme of our investigation, the “Kripke reduction”.

§5. The “Kripke reduction” and MFS . In this section we work out the details of the “Kripke reduction” for the modal predicate of MFS , that is we investigate whether the modal predicate of MFS can be understood as a truth predicate modified by a modal operator. In the present proof-theoretic setting, the question of whether the “Kripke reduction” is feasible transforms into the question of whether MFS can be reduced to some theory of truth formulated in the language $\mathcal{L}_{PAT}^\square$, that is \mathcal{L}_{PAT} augmented by a modal operator \square ,²⁵ assuming some reasonable underlying modal logic. Indeed, MFS should not be reduced

²⁴ If we have a contingent predicate in our language, the relation between the contingent modal theory and revision semantics may be stated as follows (terminology is adopted from section 3.2.2). Let F be a frame and Γ_F a modal classical jump. Moreover, let $w^* \in W$ be a distinguished world for which $w^* \models \Sigma$. Then for all $n \in \omega$

$$f \in \Gamma_F^n[Val_F] \Leftrightarrow (F, f) \models BMFS_n^C \ \& \ F, w^* \models_f BMFS_n^C + \Sigma$$

Similarly, the result for MFS carries over. Let F be a frame and $n \geq 1$. Then

$$F \text{ is an equivalence frame} \Leftrightarrow F \models^n MFS_n^C \ \& \ F, w^* \models MFS_n^C + \Sigma$$

²⁵ All formation rules are standard, that is, no restriction on the formation rules as in the case of $\mathcal{L}_{PA}^{\square\blacksquare}$ are in place. For this reason we have omitted giving a full fledged definition of the language.

to any theory of truth but to FS formulated in $\mathcal{L}_{PAT}^\square$ for otherwise the “truth” predicate of the target theory would not be a truth predicate in the sense of MFS and, arguably, we would have failed to “reduce” the modal predicate to the truth predicate modified by the modal operator.

If we formulate FS in $\mathcal{L}_{PAT}^\square$ we need to provide a further axiom which deals with the interaction of the modal operator and the truth predicate. We denote this modified version of FS by FS^\square .

DEFINITION 5.1 (FS^\square). *The theory FS^\square consists of all axioms and rules of FS in the language $\mathcal{L}_{PAT}^\square$ together with the following axiom:*

$$(IA^\square) \quad \forall x (Sent(x) \rightarrow (\Box Tx \leftrightarrow T \Box x))$$

FS^\square does not give us any information concerning the modal properties of the \Box -operator as they depend on the underlying modal logic assumed.²⁶ To be sure, this was how we intended it to be for the “logical” properties of the boolean connectives or the quantifier are also determined by the underlying logic, that is classical first-order logic, and not by the theory of truth. The set of theorems of FS^\square will therefore depend on the modal logic assumed. As MFS is meant to capture a modal notion somewhat related to the modal operator logic $S5$ we assume the modal principles and rules of this very system in addition to the axioms of standard first-order logic. In addition, we stipulate the axioms of “necessity of distinctness” and the “Barcan formula”

$$(ND^\square) \quad \forall x, y (x \neq y \rightarrow \Box x \neq y)$$

$$(BF^\square) \quad \forall x \Box \phi \rightarrow \Box \forall x \phi$$

and call the resulting modal logic $QS5$, that is, “quantified $S5$ ”.²⁷ We write $FS^\square \vdash_{\mathcal{S}} \phi$ iff ϕ is derivable in FS^\square assuming the modal operator logic \mathcal{S} . Thus, in particular $FS^\square \vdash_{QS5} \phi$ denotes the fact that ϕ is derivable in FS^\square assuming the underlying modal logic to be $QS5$.

Already in the modal operator logic QK , K in first-order logic plus the modal principles (ND^\square) and (BF^\square) we can prove the rigidity of syntax:

LEMMA 5.2. *For all $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PA}$*

$$FS^\square \vdash_{QK} \forall x_1, \dots, x_n (\phi(x_1, \dots, x_n) \rightarrow \Box \phi(x_1, \dots, x_n))$$

Proof. By induction over the positive built up of ϕ . □

Amongst other things, the latter proof hinges on the fact the following principles

$$(NI^\square) \quad \forall x, y (x = y \rightarrow \Box x = y)$$

$$(CBF^\square) \quad \Box \forall x \phi \rightarrow \forall x \Box \phi$$

are theorems of QK .

²⁶ FS^\square is consistent in all consistent normal modal logics. Depending on the normal modal logic at stake we may give an interpretation of FS^\square in FS by translating $\Box \phi$ by ϕ or by $s = s$ for some term s .

²⁷ Actually, (ND^\square) and (BF^\square) are derivable in $S5$ iff standard first-order logic with identity is assumed. See Fitting & Mendelsohn (1998) for proofs of these facts. We have nonetheless decided to explicitly mention (ND^\square) and (BF^\square) as these are the only additional principles we need to add to the smallest normal modal logic K iff we wish to reduce $BMFS$ to FS^\square . But in K , assuming standard first-order logic, neither (ND^\square) nor (BF^\square) are derivable.

In order to reduce or interpret MFS in $FS\Box$ assuming the underlying modal operator logic to be QSS we need to provide a translation function from \mathcal{L}_{PATN} to \mathcal{L}_{PAT} where this translation should interpret the modal predicate ‘ N ’ as ‘ $\Box T$ ’ and keep the remaining vocabulary fixed. It turns out that we can use the translation function provided by Halbach & Welch (2009), if we add a further clause taking care of the truth predicate of \mathcal{L}_{PATN} .²⁸

As Halbach and Welch point out defining the translation function by a simple recursion over ω will not work in this case because within self-referential or “ungrounded” sentences the necessity and the truth predicate, implicitly, occur infinitely many times. Thus there will be no starting point from where to start a simple recursion. Fortunately, due to the recursion theorem we know that a translation function of the desired kind exists:

LEMMA 5.3 (Halbach and Welch). *There is a translation function ρ from \mathcal{L}_{PATN} to \mathcal{L}_{PAT}^{\Box} with the following properties*

$$\rho(\phi) := \begin{cases} \phi & , \text{ if } \phi \doteq (s = t) \\ T\rho^{\bullet}(s) & , \text{ if } \phi \doteq Ts, \text{ for some term } s \\ \Box T\rho^{\bullet}(s) & , \text{ if } \phi \doteq Ns, \text{ for some term } s \\ \neg\rho(\psi) & , \text{ if } \phi \doteq \neg\psi \\ \rho(\psi) \wedge \rho(\chi) & , \text{ if } \phi \doteq \psi \wedge \chi \\ \forall x\rho(\psi) & , \text{ if } \phi \doteq \forall x\psi \\ 0 & , \text{ if } \text{ otherwise} \end{cases}$$

Here, the function symbol ρ^{\bullet} represents the primitive recursive function ρ and thus if s is the name of a formula ϕ , then $\rho^{\bullet}(s)$ will be the name of the formula $\rho(\phi)$. As an example consider the formula $N[T[0 = 0]]$. By definition of ρ , $\rho(N[T[0 = 0]]) = \Box T\rho^{\bullet}([T[0 = 0]])$ and since $\rho^{\bullet}([T[0 = 0]]) = [\rho(T[0 = 0])] = [T\rho^{\bullet}([0 = 0])]$ and $\rho^{\bullet}([0 = 0]) = [\rho(0 = 0)] = [0 = 0]$ we get $\rho(N[T[0 = 0]]) = \Box T[T[0 = 0]]$.

With this translation function at hand we can carry out the reduction of MFS to $FS\Box$ assuming QSS .

THEOREM 5.4. *For all $\phi \in \mathcal{L}_{PATN}$, if $MFS \vdash \phi$, then $FS\Box \vdash_{QSS} \rho(\phi)$.*

The main work in the proof of Theorem 5.4 is done by the following lemma:

LEMMA 5.5. *Let QSS be the underlying modal operator logic. Then the following are theorems of $FS\Box$:*

- (i) $\forall x, v, s, t (Sent(\forall vx) \rightarrow (Val(s) = Val(t) \rightarrow (\Box Tx(t/v) \leftrightarrow \Box Tx(s/v))))$
- (ii) $\forall s, t (Val(s) \neq Val(t) \rightarrow \Box Ts \neq t)$
- (iii) $\forall x, v (Sent(\forall vx) \rightarrow (\forall t \Box Tx(t/v) \rightarrow \Box T\forall yx))$
- (iv) $\forall x, y (Sent(x \rightarrow y) \rightarrow (\Box T(x \rightarrow y) \rightarrow (\Box Tx \rightarrow \Box Ty)))$
- (v) $\forall t (\Box T \Box T t \leftrightarrow T \Box T t)$
- (vi) $\forall x (Sent(x) \rightarrow (\Box Tx \rightarrow Tx))$
- (vii) $\forall t (T \Box T t \rightarrow \Box T \Box T t)$
- (viii) $\forall t (T \neg \Box T t \rightarrow \Box T \neg \Box T t)$

²⁸ As we have mentioned in passing in the introduction, Halbach & Welch (2009) have recently carried out the “Kripke-reduction” semantically. They show that the translation function we provide below is truth preserving with respect to two intended semantics.

Proof. We hint at the proofs of item (i)–(vi) but provide detailed proofs of items (vii)–(viii). (i) follows from Lemma 5.2, the modal properties of $QS5$ and the fact that

$$(Reg_T) \quad \forall x, v, s, t (Sent(\forall v x) \rightarrow (Val(s) = Val(t) \rightarrow (Tx(t/v) \leftrightarrow Tx(s/v))))$$

is a theorem of $(FS\Box)$. (ii) by $(FS1)$, Lemma 5.2 and $QS5$; (iii) by Lemma 5.2, $(FS4)$ and $QS5$. (iv) is a direct consequence of (K_T) , Lemma 5.2 and $QS5$. (v) is an immediate consequence of $(IA\Box)$. (vi) follows directly from the modal axiom (T) .

(vii) By generalization we may derive from an instance of (4), $\forall x(\Box Tx \rightarrow \Box\Box Tx)$. This serves as the starting point of our derivation:

1. $\forall x(\Box Tx \rightarrow \Box\Box Tx)$ (4)
2. $\forall t(T(\Box \neg T t \rightarrow \Box \Box \neg T t))$ 1, $(T-Intro)$, $(FS4)$
3. $\forall t(T\Box \neg T t \rightarrow T\Box \Box \neg T t)$ (K_T)
4. $T\Box \Box \neg T t \leftrightarrow \Box T\Box \neg T t$ $(IA\Box)$
5. $\forall t(T\Box \neg T t \rightarrow \Box T\Box \neg T t)$ 3, 4

(viii) By first-order logic and (E) we derive $\forall t(\neg\Box T \neg T t \rightarrow \Box\neg\Box T \neg T t)$ and then reason as follows:

1. $\forall t(\neg\Box T \neg T t \rightarrow \Box\neg\Box T \neg T t)$
2. $\forall t(\neg\Box T \neg T t \leftrightarrow T \neg \Box \neg T t)$ $(IA\Box)$, $(FS2)$
3. $\Box\forall t(\neg\Box T \neg T t \leftrightarrow T \neg \Box \neg T t)$ 2, (Nec)
4. $\forall t\Box(\neg\Box T \neg T t \leftrightarrow T \neg \Box \neg T t)$ 3, $(CBF\Box)$, Lemma 5.2
5. $\forall t(\Box\neg\Box T \neg T t \leftrightarrow \Box T \neg \Box \neg T t)$ 4, (K)
6. $\forall t(T \neg \Box \neg T t \leftrightarrow \Box T \neg \Box \neg T t)$ 1, 2, 5

□

Before we give the proof of Theorem 5.4 we need one further lemma which tells us that we can prove certain properties about ρ in $FS\Box$:

LEMMA 5.6. *The following is provable in PA and thus in $FS\Box$:*

- (i) $\forall x(Sent_{\mathcal{L}_{PATN}}(x) \rightarrow Sent_{\mathcal{L}_{PAT}^{\Box}}(\rho^{\bullet}(x)))$
- (ii) $\forall x(Sent_{\mathcal{L}_{PA}}(x) \rightarrow \rho^{\bullet}(x) = x)$

Proof. By formalizing the properties of ρ . □

We provide the proof of the main theorem:

Proof of Theorem 5.4. The proof is an induction on the length of a proof in MFS. Lemma 5.5 together with Lemma 5.6 establishes the base of the induction since it is easily verified that the eight axioms of MFS extending FS translate into the modal principles (i)–(viii) of Lemma 5.5. As a sample case we discuss the axiom (T') , that is, we show the translation of (T') to be provable, if we assume the underlying modal logic to be $QS5$:

1. $\forall x(Sent_{\mathcal{L}_{PAT}^{\Box}}(x) \rightarrow (\Box Tx \rightarrow Tx))$ Lemma 5.5: (vi)
2. $\forall x(Sent_{\mathcal{L}_{PAT}^{\Box}}(\rho^{\bullet}(x)) \rightarrow (\Box T\rho^{\bullet}(x) \rightarrow T\rho^{\bullet}(x)))$ 1
3. $\forall x(Sent_{\mathcal{L}_{PATN}}(x) \rightarrow (\Box T\rho^{\bullet}(x) \rightarrow T\rho^{\bullet}(x)))$ 2, Lemma 5.6: (i)
4. $\forall x(Sent_{\mathcal{L}_{PATN}}(x) \rightarrow (\rho(Nx) \rightarrow \rho(Tx)))$ 3, Lemma 5.3
5. $\rho(\forall x(Sent_{\mathcal{L}_{PATN}}(x) \rightarrow (Nx \rightarrow Tx)))$ 4, Lemma 5.3

For the induction step we observe that since $FS\Box$ is closed under (*T-Intro*) and (*Nec*) is a rule of proof of the underlying modal logic $QS5$ if $FS\Box \vdash_{QS5} \rho(\phi)$, then $FS\Box \vdash_{QS5} \Box T\rho^\bullet([\phi])$ and hence $FS\Box \vdash_{QS5} \rho(N[\phi])$. (*T-Intro*) and (*T-Elim*) do not cause any trouble and we conclude to the desired. \square

With this proof ends our discussion of the Kripke-reduction as we have shown the theory MFS to be “Kripke-reducible” to the theory $FS\Box$ assuming the modal logic $QS5$. As we have pointed out in the introduction this should not be understood as a result favoring operator approaches to modality, but rather to provide a conceptual insight into why our strategy for constructing modal theories has proven to be successful in the present case.²⁹

§6. Conclusion. In this paper we have presented a general strategy of constructing modal theories over theories of truth and we have tested this strategy with respect to the axiomatic theory of truth Friedman-Sheard. This led to the development and the evaluation of the theory of Modal Friedman-Sheard. The theory proved to be adequate from the perspective of modal operator logic as it was shown that MFS preserves theoremhood with respect to the modal operator logic $S5$ modulo translation. We also presented modal revision semantics and showed that MFS axiomatizes this semantic theory in a certain way: Modal revision semantics provides us with the suitable interpretations for the truth and the modal predicate in standard models of the subtheories of MFS . A third theme we discussed in our paper was the “Kripke reduction” which suggests an understanding of the modal predicate in terms of the truth predicate and a modal operator. This third theme should not be viewed as part of the evaluation of the theory MFS but as an attempt of understanding why the general strategy we have proposed worked so well. The success of the reduction suggests that the way we have axiomatized the modal predicate leaves us with but one potentially paradoxical concept, namely the concept of truth.

We take our findings to show that Modal Friedman-Sheard is a viable modal theory which should be rather attractive to the modal theorist, so long as she accepts the underlying theory of truth. This indicates that the general strategy for developing modal theories over theories of truth is promising strategy to pursue. However, there are many theories of truth available and it needs to be seen how widely applicable the strategy is. Therefore, it seems fitting to test the strategy we have developed using alternative theories of truth. In particular, this seems interesting because the theory Friedman-Sheard is often dismissed due to its ω -inconsistency. We will take up this task in the second part of our investigation and test our strategy by applying it to the axiomatic theory of truth Kripke-Feferman. Testing our strategy using the theory of truth Kripke-Feferman should give us some feedback on how widely applicable our strategy is since the theories of truth Friedman-Sheard and Kripke-Feferman are very different in character. While Friedman-Sheard is a symmetric but ω -inconsistent theory, Kripke-Feferman is nonsymmetric but ω -consistent. Therefore, if it turns out that the strategy we have presented works in both cases, we think there is

²⁹ The “Kripke reduction” remains feasible, if we allow for a contingent predicate in our language. Of course, $FS\Box$ needs to be supplemented with the axiom (*FSC*) we discussed in section 3.2.2. Call this theory $FS\Box^C$. Moreover, we need to guarantee that the underlying modal logic treats the contingent theory Σ as such, that is the rule of (*Nec*) will no longer be applicable once a Σ -axiom or rule has been used in a proof. The translation function ρ will act as the identity function on sentences of the form $P^n t_1, \dots, t_n$. It is then straightforward to derive that for all $\phi \in \mathcal{L}_C$

$$MFS^C + \Sigma \vdash \phi \Rightarrow FS\Box^C + \Sigma \vdash_{QM5} \rho(\phi)$$

good reason to believe that we have developed a widely applicable strategy for constructing modal theories.

§7. Acknowledgments. The research leading to this paper was carried out within the research project “Syntactical Treatment of Interacting Modalities”, which is funded by the Deutsche Forschungsgesellschaft (DFG). The research project is hosted by the Munich Center for Mathematical Philosophy (MCMP), which, in turn, is supported by the Alexander von Humboldt Foundation. We want to thank Catrin Campbell-Moore, Fabrice Correia, Martin Fischer, Volker Halbach, Leon Horsten, Hannes Leitgeb, and Karl-Georg Niebergall for helpful discussions of this material and several anonymous referees for comments that helped improving this paper.

BIBLIOGRAPHY

- Asher, N., & Kamp, H. (1989). Self-reference, attitudes and paradox. In Chierchia, G., Partee, B. H., and Turner, R., editors. *Properties, Types, and Meaning. Vol. 1: Foundational Issues*, Dordrecht, Netherlands: Kluwer, pp. 85–158.
- Belnap, N., & Gupta, A. (1993). *The Revision Theory of Truth*. Cambridge, MA: The MIT Press.
- Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal Logic*. Cambridge, UK: Cambridge University Press.
- des Rivières, J., & Levesque, H. J. (1986). The consistency of syntactical treatments of knowledge. In Halpern, J. Y., editor. *Theoretical Aspects of Reasoning about Knowledge*. San Mateo, CA: Morgan Kaufmann.
- Fitting, M., & Mendelsohn, R. L. (1998). *First-Order Modal Logic*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Friedman, H., & Sheard, M. (1987). An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, **33**, 1–21.
- Gupta, A. (1982). Truth and paradox. *Journal of Philosophical Logic*, **11**, 1–60.
- Halbach, V. (1994). A system of complete and consistent truth. *Notre Dame Journal of Formal Logic*, **35**(3), 311–327.
- Halbach, V. (2001). Disquotational truth and analyticity. *The Journal of Symbolic Logic*, **66**, 1959–1973.
- Halbach, V. (2002). Modalized disquotationalism. In Halbach, V., and Horsten, L., editors. *Principles of Truth*. Frankfurt, Germany: Ontos Verlag, pp. 75–101.
- Halbach, V. (2011). *Axiomatic Theories of Truth*. New York: Cambridge University Press.
- Halbach, V., Leitgeb, H., & Welch, P. (2003). Possible-worlds semantics for modal notions conceived as predicates. *Journal of Philosophical Logic*, **32**, 179–222.
- Halbach, V., & Welch, P. (2009). Necessities and necessary truths: A prolegomenon to the use of modal logic in the analysis of intensional notions. *Mind*, **118**, 71–100.
- Herzberger, H. (1982). Naive semantics and the liar paradox. *The Journal of Philosophy*, **79**, 479–497.
- Horsten, L. (2002). An axiomatic investigation of provability as a primitive predicate. In Halbach, V., and Horsten, L., editors. *Principles of Truth*. Frankfurt, Germany: Ontos Verlag, pp. 203–220.
- Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy*, **72**, 690–716.
- Leitgeb, H. (2006). Towards a logic of type-free modality and truth. In C. Dimitracopoulos, L. Newelski, D. Normann, and J. R. Steele, editors. *Logic Colloquium 05, Lecture Notes in Logic*. New York: Cambridge University Press, pp. 68–84.

- McGee, V. (1985). How truthlike can a predicate be? A negative result. *The Journal of Philosophical Logic*, **14**(4), 399–410.
- Monk, J. D. (1976). *Mathematical Logic*. New York: Springer Verlag.
- Montague, R. (1963). Syntactical treatments of modality, with corollaries on reflexion principles and finite axiomatizability. *Acta Philosophica Fennica*, **16**, 153–167.
- Peacocke, C. (1976). Truth definitions and actual languages. In Evans, G., and McDowell, J., editors. *Truth and Meaning*. Oxford: Oxford University Press, pp. 162–188.
- Quine, W. V. O. (1977). Intensions revisited. *Midwest Studies in Philosophy*, **2**, 5–11.
- Sheard, M. (1994). A guide to truth predicates in the modern era. *The Journal of Symbolic Logic*, **59**, 1032–1054.
- Sheard, M. (2001). Weak and strong theories of truth. *Studia Logica*, **68**, 89–101.
- Stern, J. (2012). *Toward Predicate Approaches to Modality*. PhD thesis, University of Geneva, 2012.
- Stern, J. (forthcoming). Montague's theorem and modal logic. *Erkenntnis*, DOI 10.1007/s10670-013-95237.

MCMP, FAKULTÄT FÜR PHILOSOPHIE

WISSENSCHAFTSTHEORIE UND RELIGIONSWISSENSCHAFTEN

LMU MÜNCHEN, GESCHWISTER-SCHOLL-PLATZ 1

D-80539 MÜNCHEN, GERMANY

E-mail: johannes.stern@lrz.uni-muenchen.de