

## MODALITY AND AXIOMATIC THEORIES OF TRUTH II: KRIPKE-FEFERMAN

JOHANNES STERN

Munich Center for Mathematical Philosophy, LMU Munich

**Abstract.** In this second and last paper of the two part investigation on “Modality and Axiomatic Theories of Truth” we apply a general strategy for constructing modal theories over axiomatic theories of truth to the theory Kripke-Feferman. This general strategy was developed in the first part of our investigation. Applying the strategy to Kripke-Feferman leads to the theory Modal Kripke-Feferman which we discuss from the three perspectives that we had already considered in the first paper, where we discussed the theory Modal Friedman-Sheard. That is, we first show that Modal Kripke-Feferman preserves theoremhood modulo translation with respect to modal operator logic. Second, we develop a modal semantics fitting the newly developed theory. Third, we investigate whether the modal predicate of Modal Kripke-Feferman can be understood along the lines of a proposal of Kripke, namely as a truth predicate modified by a modal operator.

**§1. Introduction.** This paper is the second of two papers in which we explore a general strategy for constructing modal theories where the modal notion is conceived as a predicate. In the first paper “Modality and Axiomatic Theories of Truth I: Friedman-Sheard” we outlined the general strategy and applied it to the axiomatic theory of truth Friedman-Sheard. This led to the development and the investigation of the theory Modal Friedman-Sheard and a corresponding semantics, modal revision semantics. The basic idea of the strategy is to characterize the modal predicate by its interaction with the truth predicate. This allows us to analyze the modal paradoxes as manifestations of the paradoxicality of the concept of truth. In this paper we continue testing our strategy by applying it to an alternative theory of truth, namely Kripke-Feferman.

Testing our strategy with respect to the axiomatic theory of truth Kripke-Feferman is interesting from several perspectives. First, the theory Kripke-Feferman is a well accepted theory of truth and considered by many to be one of the philosophically more attractive theories available. Therefore, a modal theory which is based on this very theory of truth seems to be philosophically promising. Second, the theories of truth Friedman-Sheard and Kripke-Feferman are very different in character: Friedman-Sheard is a symmetric theory of truth which, on the downside, is  $\omega$ -inconsistent. Kripke-Feferman is a nonsymmetric theory of truth (i.e., the inner and the outer logic of the theory differ), but it is  $\omega$ -consistent and, consequently, has nice models extending the standard model. In a way, the two theories mark two extremes amongst the prominent axiomatic theories of truth and by applying our strategy to these two theories, we hope to get some feedback on how widely applicable our strategy is and which theories of truth fit best with this kind of strategy.

---

Received: August 8, 2013.

The paper is meant to be mostly self-contained and we will therefore recall the essentials of the strategy we introduced in the first part of our investigation. Yet, we frequently omit in-depth philosophical justification and interpretation of certain methodological decisions, assumptions or theorems and refer the reader to the first paper of our investigation for this purpose. We also refer the reader to our first paper for a presentation of our terminological stipulations.

The paper is structured as follows. After quickly reminding the reader of the essentials of the strategy we introduced in the first paper of our study, we apply the strategy to the axiomatic theory of truth Kripke-Feferman and construct a family of theories, which we label modal extensions of Kripke-Feferman. Similar to the investigation of the theory Modal Friedman-Sheard in the first paper, we then discuss modal extensions of Kripke-Feferman from three different perspectives.

We start by showing that modal extensions of Kripke-Feferman are adequate modal theories from the perspective of modal operator logic, that is, these theories preserve theoremhood modulo translation with respect to modal operator logic. We take this to be a minimal criterion, an interesting modal theory, ought to fulfill. In the next section we turn to semantic consideration and follow the outlines of work by Halbach & Welch (2009) in developing modal fixed-point semantics, which combines Kripke's theory of truth, that is, fixed-point semantics, with ideas of possible world semantics for modal operator logic. By reflecting on (modal) fixed-point semantics a shortcoming of modal extensions of Kripke-Feferman becomes apparent, which ultimately leads to an extension of the language by a primitive possibility predicate and the development of the theory Modal Kripke-Feferman. We then show that modal fixed-point semantics leaves us with a nice method of constructing models for Modal Kripke-Feferman over the standard model. Moreover, we show that Modal Kripke-Feferman axiomatizes modal fixed-point semantics in the same way as the theory Kripke-Feferman can be said to axiomatize ordinary fixed-point semantics.

As in the case of the theory Modal Friedman-Sheard we end our technical investigation by discussing a proposal of Saul Kripke. Kripke suggested defining or understanding the modal predicate (' $N$ ') by the truth predicate modified by a modal operator (' $\Box T$ '). We show that, the modal predicate of the theory Modal Kripke-Feferman lends itself to the understanding proposed by Kripke. That is, we show the theory Modal Kripke-Feferman to be reducible to the theory Kripke-Feferman as formulated in a modal operator language. 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: in applying this strategy we haven't pushed the deductive (and expressive) strength of our approach beyond the limits of the theory of truth as formulated in a modal operator language and we allow for an understanding of the modal predicate as defined by the truth predicate and a modal operator. As a consequence, there remains only one potentially paradoxical concept, that is, the concept of truth. We end our paper by an assessment of the theory Modal Kripke-Feferman and a comparison of the two modal theories we have developed in the course of our study on "Modality and Axiomatic Theories of Truth".

**§2. Kripke-Feferman and its modal extensions.** We now start developing the modal theory arising from the theory of truth Kripke-Feferman. To this end we recall the basic idea of the strategy we developed in the first paper of our study on "Modality and Axiomatic Theories of Truth". The key feature of this strategy is to allow the introduction

and elimination of the modal predicate only via the detour of the truth predicate. This necessitates a reformulation of most modal principles and rules of modal operator logic within the predicate setting. Roughly, a truth predicate needs to be introduced whenever the standard formulation of a modal principle introduces or eliminates an occurrence of the modality at stake. By introducing the truth predicate we avoid the (dis)quotation of sentences which is one of the main sources of paradox. The strategy, therefore, has the immediate benefit of rendering the modal notion paradox safe as long as the notion of truth is.

If we apply the strategy in formulating counterparts to the principles of the modal operator logic *S5*

- (*T*)  $\Box\phi \rightarrow \phi$
- (4)  $\Box\phi \rightarrow \Box\Box\phi$
- (*E*)  $\neg\Box\phi \rightarrow \Box\neg\Box\phi$

within the predicate setting, we see that each of these principles requires the introduction of the truth predicate. We arrive at the following modal principles.

- (*T'*)  $\forall x(Sent(x) \rightarrow (Nx \rightarrow Tx))$
- (4')  $\forall t(T \dot{N} t \rightarrow N \dot{N} t)$
- (*E'*)  $\forall t(T \dot{\neg} \dot{N} t \rightarrow N \dot{\neg} \dot{N} t).$

Even though nothing like a truth predicate (or truth operator) is used in the formulation of the operator principles, we think that the principles (*T'*), (4'), and (*E'*) express the same modal properties as the operator principles, as the truth predicate should not make any difference in this respect. Clearly, our strategy makes the modal theory dependent on the theory of truth adopted and thus, the modal theories arising from different underlying theories of truth need to be tested and compared. As we have already mentioned, in this paper we shall be interested in the modal theory arising from the theory of truth Kripke-Feferman.

The modal theory will contain further modal axioms, which were also assumed in constructing the theory Modal Friedman-Sheard. These axioms characterize the interaction of the modal predicate with the logical vocabulary and are modeled after prominent principles of first-order modal operator logic. Importantly, these axioms have the effect “rigidifying” syntax theory, that is, of collapsing arithmetical truth into necessary truth. We want our syntax theory to be rigid, because, we need to guarantee that the names of the syntactical objects and operations 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, e.g., a numeral codes in certain circumstances a closed term but in others a parenthesis, then clearly the axioms (4') and (*E'*) would no longer express the modal properties that we intend them to convey, but would be nonsensical instead.

In addition to these theory-unspecific axiom, which we already adopted in constructing the theory Modal Friedman-Sheard, we introduce some theory-specific axioms. The theory-specific axioms depend on the choice of the underlying theory of truth and may appear plausible for one theory of truth but not for another one.

Finally, we adopt the following modified rule of necessitation

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

where  $\phi$  is a sentence of the language. Obviously, (*T-Nec*) is just the rule of necessitation known from modal operator logic modified to fit our general strategy. We now turn to the discussion of the specific modal theory arising if the theory Kripke-Feferman is adopted. We first recall the essentials of the theory Kripke-Feferman.

**2.1. Kripke-Feferman.** The axiomatic theory “Kripke-Feferman” was introduced by Feferman during a talk in 1983 and was designed to capture Kripke’s purely semantic, that is model-theoretic, theory of truth proof-theoretically. Feferman, however, did not publish a written version of the theory until his 1991 “Reflecting on Incompleteness” (cf. Feferman (1991)). In the mean time, Reinhardt (1985, 1986), Cantini (1989) and McGee (1991) had investigated and discussed Feferman’s proposal in some detail and already used the now common name ‘*KF*’. A thorough discussion of the theory can also be found in the more recent work by Halbach (2011). We shall stick to Halbach in our presentation of the theory.

**DEFINITION 2.1 (Kripke-Feferman).** *The theory  $KF$  consists of all axioms of  $PA$  in the language  $\mathcal{L}_{PAT}$  and the following axioms*

- (KF1)  $\forall s \forall t (Ts \equiv t \leftrightarrow Val(s) = Val(t))$
- (KF2)  $\forall s \forall t (Ts \not\equiv t \leftrightarrow Val(s) \neq Val(t))$
- (KF3)  $\forall x (Sent(x) \rightarrow (T \neg \neg x \leftrightarrow Tx))$
- (KF4)  $\forall x, y (Sent(x \wedge y) \rightarrow (T(x \wedge y) \leftrightarrow Tx \wedge Ty))$
- (KF5)  $\forall x, y (Sent(x \wedge y) \rightarrow (T \neg (x \wedge y) \leftrightarrow T \neg x \vee T \neg y))$
- (KF6)  $\forall v \forall x (Sent(\forall vx) \rightarrow (T(\forall vx) \leftrightarrow \forall t (Tx(t/v))))$
- (KF7)  $\forall v \forall x (Sent(\forall vx) \rightarrow (T(\neg \forall vx) \leftrightarrow \exists t (T \neg x(t/v))))$
- (KF8)  $\forall t (T T t \leftrightarrow T(Val(t)))$
- (KF9)  $\forall t (T \neg T t \leftrightarrow T(\neg Val(t)) \vee \neg Sent(Val(t)))$
- (KF10)  $\forall x (Tx \rightarrow Sent(x))$

*KF*, in contrast to symmetric theories of truth like *FS*, is built on the idea of a positive inductive definition. The axioms (KF8) and (KF9) are distinctive axioms of the theory *KF*.

**THEOREM 2.2.** *KF is consistent.*

*Proof.* For a detailed proof see Halbach (2011). We shall sketch the construction of models for *KF* later in this section. □

Two facts about *KF* will be of some importance. First, *KF* proves the uniform Tarski biconditionals without the truth predicate, and thus, *UTB* is a subtheory of *KF*. Second, the principle of regularity, that is, the formalized substitution principle is derivable in *KF*.

**FACT 2.3.** *The following are theorems of  $KF$  for all  $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PA}$*

- (UTB)  $\forall t_1, \dots, t_n (T[\phi(t_1, \dots, t_n)] \leftrightarrow \phi(Val(t_1), \dots, Val(t_n)))$
- (Reg<sub>T</sub>)  $\forall x, v, s, t (Sent(\forall vx) \rightarrow (Val(s) = Val(t) \rightarrow (Tx(s/v) \leftrightarrow Tx(t/v))))$

In what is to come we employ the notion of a total and consistent sentence:

**DEFINITION 2.4.** *Let  $\phi$  be a sentence of  $\mathcal{L}_{PAT}$  ( $\mathcal{L}_{PATN}$ ). We say that  $\phi$  is a total and consistent sentence relative to  $\Sigma$  iff*

$$\Sigma \vdash \neg T[\phi] \leftrightarrow T[\neg\phi]$$

In the same spirit we define the predicate  $tc(x)$  as follows:

$$tc(x) := \neg Tx \leftrightarrow T \neg x$$

As the truth predicate does not, in general, commute with negation not every sentence of  $\mathcal{L}_{PAT}$  ( $\mathcal{L}_{PATN}$ ) is total and consistent relative to  $KF$ . Moreover, if we required every sentence to be total in  $KF$ , the resulting theory would be inconsistent.<sup>1</sup> Although we cannot let negation commute with the truth predicate we may consistently add the consistency axiom

$$(Cons) \quad \forall x(Sent(x) \rightarrow (T \neg x \rightarrow \neg Tx))$$

to  $KF$ .<sup>2</sup> The resulting theory will be called  $KFC$ . We shall need the following fact concerning  $KFC$ .

FACT 2.5. *The following are theorems of KFC:*

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

$$(T_T) \quad \forall t_1, \dots, t_n(T[\phi(t_1, \dots, t_n)] \rightarrow \phi(Val(t_1), \dots, Val(t_n)))$$

for all  $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PAT}$ .

*Proof.* See Halbach (2011, pp. 213f). □

**2.2. Modal extensions of Kripke-Feferman.** Finally we introduce the family of modal extensions of  $KF$ . A theory is called a modal extension of  $KF$  if some basic modal principles are theorems of the theory. Besides some modal principles, we have already discussed we adopt one  $KF$ -specific axiom

$$(RN) \quad \forall t(T \ N \ t \leftrightarrow N \ Val(t))$$

and one  $KF$ -specific rule

$$(TcR) \quad \frac{tc([\phi])}{tc([N[\phi]])}$$

for all  $\phi \in Sent_{\mathcal{L}_{PATN}}$ .  $(RN)$  has some intuitive appeal and more importantly it fits perfectly with the theoretical outlines of  $KF$  which are to allow for the Tarski-biconditionals of all atomic sentences.  $(TcR)$ , in contrast, plays a rather instrumental role as although it is plausible from a theoretical perspective it seems hard to come up with some convincing intuitive gloss for it. It says that if a sentence is total and consistent then so is its necessitation. Phrased alternatively, the rule asserts that if a sentence  $\phi$  is truth determinate, then so is the necessitation of  $\phi$ ,  $N[\phi]$ .<sup>3</sup>

DEFINITION 2.6 (Modal Extension of  $KF$ ). *Let  $\Sigma$  be a theory of the language  $\mathcal{L}_{PATN}$ . We call  $\Sigma$  a modal extension of  $KF$  iff  $\Sigma$  proves all axioms of  $KF$  in the language  $\mathcal{L}_{PATN}$* <sup>4</sup>

<sup>1</sup> Cf. Halbach (2011, pp. 197/98).

<sup>2</sup> Conversely, we may also consistently add the converse direction to  $KF$ , that is the completeness axiom

$$(Comp) \quad \forall x(Sent(x) \rightarrow (\neg Tx \rightarrow T \neg x))$$

as long as we do not add  $(Cons)$ .

<sup>3</sup> See Feferman (2008) for the notion of truth determinateness.

<sup>4</sup> Thereby we mean the axioms of  $KF$  where the syntactical predicates and functions symbol of  $\mathcal{L}_{PAT}$  have been replaced by the corresponding expressions of  $\mathcal{L}_{PATN}$ . In general, as we pointed

and, additionally, proves the following principles

- (Reg<sub>N</sub>)  $\forall x, v, s, t (Sent(\forall v x) \rightarrow (Val(s) = Val(t) \rightarrow (Nx(s/v) \leftrightarrow Nx(t/v))))$
- (ND)  $\forall s \forall t (Val(s) \neq Val(t) \rightarrow N(s \neq t))$
- (BF)  $\forall v \forall x (Sent(\forall v x) \rightarrow (\forall t Nx(t/v) \rightarrow N \forall v x))$
- (K<sub>N</sub>)  $\forall x, y (Sent(x \rightarrow y) \rightarrow (N(x \rightarrow y) \rightarrow (Nx \rightarrow Ny)))$
- (RN)  $\forall t (T N t \leftrightarrow N Val(t))$

Moreover,  $\Sigma$  is closed under the following rules

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

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

One modal extension of  $KF$  will be of special interest, namely the theory  $MKF^-$  which proves the modal principles  $(T')$ ,  $(4')$  and  $(E')$ .

DEFINITION 2.7 ( $MKF^-$ ). *The smallest modal extension of  $KF$ , which has  $(T')$ ,  $(4')$  and  $(E')$  as theorems is called  $MKF^-$ .*

$MKF^-$  is consistent and, consequently, there are consistent modal extensions of  $KF$ .

THEOREM 2.8.  *$MKF^-$  is consistent.*

*Proof.* The proof works similar to the consistency proof for  $MFS$ . Using the recursion theorem we define a translation function  $g$  from  $\mathcal{L}_{PATN}$  to  $\mathcal{L}_{PAT}$  which translates the modal predicate  $N$  by the truth predicate  $T$  and leaves the remaining vocabulary fixed. It is then straightforward to show that

$$MKF^- \vdash \phi \Rightarrow KFC \vdash g(\phi) \qquad \square$$

COROLLARY 2.9. *There exists a consistent modal extension  $\Sigma$  of  $KF$ .*

For exactly the same reason as in the case of  $MFS$  the interpretation  $g$  will not be faithful. We now move on to showing that syntax theory is rigid in any modal extension of  $KF$ . Consequently, we are guaranteed that the modal axioms convey their intended meaning.

THEOREM 2.10. *For all  $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PA}$  and any modal extension  $\Sigma$  of  $KF$  we have:*

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

*Proof of Theorem 2.10.* By induction over the positive built up of  $\phi$ . See Stern (2012). □

Before we show the adequacy of modal extensions of  $KF$  with respect to modal operator logic, we prove the inconsistency of two possible strengthenings of  $MKF^-$  as this will later prove instructive.

---

out in “Modality and Axiomatic Theories of Truth I: Friedman-Sheard”, when we speak of a theory (or axioms) in an extended language we always assume that the syntactical predicates and functions symbols occurring in the axioms have been replaced by the corresponding expressions of the extended language. We also assume the induction scheme to be extended to the new language.

THEOREM 2.11. *Let  $\Sigma$  be a modal extension of  $KF$  with  $MKF^- \subseteq \Sigma$ . If  $\Sigma$  proves one of the following*

- (i)  $\forall t (T \neg N t \leftrightarrow \neg N Val(t))$
- (ii)  $\forall t (T \neg N t \leftrightarrow \neg T N t)$

then  $\Sigma$  is inconsistent.

*Proof.* We first observe that due to (RN) (i) and (ii) are equivalent in any modal extension  $\Sigma$ . Therefore, we may use both principles in deriving the inconsistency result. By the diagonal lemma we derive in *PATN*

$$(N\lambda) \quad l = [\neg Nl]$$

and we may reason as follows in  $MKF^-$

$$\begin{aligned} T[Nl] &\rightarrow Tl && (RN), (T') \\ &\rightarrow T[\neg Nl] && (N\lambda), Subst \\ &\rightarrow \neg T[Nl] && (ii) \end{aligned}$$

We may thus infer  $\neg T[Nl]$ , which by (RN) yields  $\neg Nl$ . Then, by (Nλ) and substitution, we obtain  $\neg N[\neg Nl]$ . By (E') we get  $\neg T[\neg Nl]$  and by (ii)  $T[Nl]$ . Contradiction.<sup>5</sup> □

**2.3. Modal operator logic and modal extensions of  $KF$ .** We now show that modal extensions of  $KF$  and, in particular,  $MKF^-$  are adequate modal theories from the perspective of modal operator logic. To this end we show that modal extensions of  $KF$  preserve theoremhood with respect to modal operator logic modulo translation. The set up will be exactly as in the first part of our study when we established the parallel result for the theory *MFS*. That is, we consider a multimodal operator language  $\mathcal{L}_{PA}^{\square, \blacksquare}$ , where the sentences of  $\mathcal{L}_{PA}$  act as the propositional atoms of the language. We let  $\tau$  be a translation function that commutes with the logical connectives and which translates  $\square\phi$  ( $\blacksquare\phi$ ) by  $N[\phi]$  ( $T[\phi]$ ).<sup>6</sup>

The following lemma establishes that the Tarski biconditionals for all sentences in the range of the translation function  $\tau$  are provable in every modal extension of  $KF$ . This is the central lemma for proving the adequacy of modal extensions of  $KF$ .

LEMMA 2.12. *Let  $\Sigma$  be a modal extension of  $KF$ . Then for all  $\phi \in \mathcal{L}_{PA}^{\square, \blacksquare}$*

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

*Proof.* By induction on the positive complexity of  $\phi$  one simultaneously shows

- (†)  $\Sigma \vdash T[\tau(\phi)] \leftrightarrow \tau(\phi)$
- (‡)  $\Sigma \vdash tc([\tau(\phi)])$

<sup>5</sup> Notice that for this derivation to work we need to assume that our language contains certain function symbols, for otherwise we could not prove the strong diagonal lemma in *PATN*. We don't know whether the inconsistency result obtains when we dispense of these function symbols and the strong diagonal lemma is no longer available. As, for example, Heck (2007) points out, there are cases in which the strong diagonal lemma proves indispensable in deriving the inconsistency of a theory.

<sup>6</sup> We refer the reader to our first paper “Modality and Axiomatic Theories of Truth I: Friedman-Sheard” for precise definitions of the language and the translation function at stake.

We omit the details of this induction. For the case where  $\phi := \neg \Box \psi$  we make essential use of  $(\ddagger)$  and  $(TcR)$ . □

With Lemma 2.12 at our disposal, the desired adequacy result is easily obtained. We show that  $MKF^-$  preserves theoremhood with respect to the modal operator logic  $S5$ .<sup>7</sup> More precisely, we show  $MKF^-$  to preserve theoremhood with respect to the theory resulting from  $PATN$  by addition of the modal principles of  $S5$  axiomatizing the  $\Box$ -operator and the Tarski biconditionals

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

for the truth operator  $\blacksquare$ .

As in the first paper we call this theory  $PAS5$ .

**THEOREM 2.13.** *Let  $\Sigma$  be a modal extension of  $KF$  extending  $MKF^-$ . Then for all  $\phi \in \mathcal{L}_{PA}^{\Box \blacksquare}$*

$$PAS5 \vdash \phi \Rightarrow \Sigma \vdash \tau(\phi)$$

*Proof.* The theorem is established by an induction on the length of a proof in  $PAS5$ . The base case follows from Lemma 2.12.<sup>8</sup> The induction step follows by Lemma 2.12 and the rule  $(T-Nec)$ . □

The adequacy of  $MKF^-$  depended almost entirely on Lemma 2.12, which is provable for every modal extension of  $KF$ . This suggests that for a modal operator logic other than  $S5$  we may find a suitable modal extension of  $KF$ , which interprets the given logic. We now turn to more semantic aspects of modal extensions of  $KF$  which will reveal a shortcoming of  $MKF^-$ . Ultimately this leads to an extension of  $MKF^-$ . The adequacy of this extension with respect to modal operator logic will be guaranteed, however, because by Theorem 2.13 every extension of  $MKF^-$  will be adequate with respect to  $PAS5$ .

**§3. Modal fixed-point semantics and  $MKF$ .** We have already mentioned that  $KF$  is meant to be an axiomatization of Kripke’s theory of truth, and as a matter of fact, the extensions of the truth predicate we obtain by Kripke’s construction are suitable interpretations of the truth predicate of  $KF$ . This suggests that there might be an interesting semantic story to be told about modal extensions of  $KF$ . Before we tackle this question, we outline the adequacy result for  $KF$ .

**3.1. Fixed-point semantics and  $KF$ .** Kripke’s idea was to work with a partially interpreted truth predicate, and thus, partial models of  $\mathcal{L}_{PAT}$ . Then, by using suitable evaluation schemes he defines a monotone jump operation. Assuming standard set theory, the monotonicity of the jump operation will guarantee the existence of fixed-points. These fixed-points have some desirable properties and are considered as possible extensions of the truth predicate by Kripke’s theory.<sup>9</sup>

Although, Kripke always remained neutral as to which evaluation scheme was preferable, he used the so-called strong Kleene scheme in his initial proposal. This leads to the following jump operation:

<sup>7</sup>  $S5$  is axiomatized by the modal principles  $(K)$ ,  $(T)$ ,  $(4)$  and  $(E)$ , and the rule  $(Nec)$ .

<sup>8</sup> See the proof of the adequacy result of  $MFS$  in “Modality and Axiomatic Theories of Truth I: Friedman-Sheard I” for more details.

<sup>9</sup> Or, this is at least one way to understand Kripke. Another way is to understand him as advocating the minimal fixed-point to be the desirable interpretation of the truth predicate.



DEFINITION 3.1 (Strong Kleene Jump). *Let  $(\mathbb{N}, S)$  be a model for  $\mathcal{L}_{PAT}$ . The strong Kleene jump  $\Theta$  is an operation from sets of natural numbers to sets of natural numbers such that*

$$\Theta(S) = \{\#\phi : (\mathbb{N}, S) \models_{SK} \phi\}$$

Here  $\models_{SK}$  stands for the strong Kleene satisfaction relation.<sup>10</sup> As we have already pointed out  $\Theta$  is a monotone operation which guarantees the existence of fixed-points, i.e. the existence of sets  $S \subseteq \omega$  for which  $\Theta(S) = S$ .

It turns out that the fixed-points of  $\Theta$  are exactly the suitable interpretations of the truth predicate of *KF* in the standard model:

THEOREM 3.2 (Feferman). *For all  $S \subseteq \omega$*

$$\Theta(S) = S \Leftrightarrow (\mathbb{N}, S) \models KF.$$

*Proof.* For a detailed proof we refer the reader to, for example McGee (1991) or Halbach (2011). The left-to-right direction works by verifying that the axioms of *KF* hold in the model. For the converse direction we need to verify that  $\Theta(S) = S$ . It suffices to prove that

$$\#\phi \in S \Leftrightarrow \#\phi \in \Theta(S),$$

which goes by induction over the positive built up of  $\phi$ . As a sample case, we discuss  $\phi \doteq \neg Tt$  as this shall prove instructive. Assume  $\#\neg Tt \in S$ , then  $(\mathbb{N}, S) \models T[\neg Tt]$  and therefore by (*KF9*)  $(\mathbb{N}, S) \models T \neg t$  (for if  $\neg Sent(t)$  we are done). Then  $(\neg t)^{\mathbb{N}} \in S$  which by definition of the Strong Kleene satisfaction relation and  $\Theta$  is just  $\#\neg Tt \in \Theta(S)$ .  $\square$

As a corollary of this theorem we may show that if we consider only consistent extensions of the truth predicate—where a set of (codes of) sentences  $S$  is consistent iff for no sentence  $\phi$ ,  $\#\phi \in S$  and  $\#\neg\phi \in S$ —the fixed-points of  $\Theta$  are precisely the extensions of the truth predicate of *KFC*:

COROLLARY 3.3. *For all consistent  $S \subseteq \omega$ :*

$$\Theta(S) = S \Leftrightarrow (\mathbb{N}, S) \models KFC.$$

**3.2. Modal Kripke-Feferman.** The proof of Theorem 3.2 suggests that we will run into trouble if we try to establish an adequacy result along these lines for modal extensions of *KF* with respect to some modal fixed-point semantics yet to be defined. The problem is that we have no axiom which allows us to deal with a formula of the form  $\neg Nt$  in proving the right-to-left direction of a modified version of Theorem 3.2, that is there seems to be no way to get from  $\#\neg Nt \in S$  to  $\#\neg Nt \in \Theta(S)$  and back. Notice that by Theorem 2.11 we may not add an axiom to *MKF*<sup>-</sup> that allows us to pull the negation out of the scope of the truth predicate.

One might therefore suggest an axiom for the modal predicate which is modeled after (*KF9*), namely

$$\forall t (T \neg Nt \leftrightarrow N \neg Val(t) \vee \neg Sent(Val(t))).$$

<sup>10</sup> For a detailed definition of  $\models_{SK}$  see Halbach (2011). However, the definition can also be extracted from our Definition 3.9.

Whereas, this axiom might well do the job, it is utterly implausible on most modal readings of the predicate ‘*N*’, cf. “it is true that it is not necessary that Peter has brown hair if and only if it is necessary that Peter does not have brown hair”.

However, if we replace the modal predicate on the right hand side of the above biconditional by its dual, for example a possibility predicate, we obtain a principle which is perfectly in line with the modal square of oppositions. Let ‘*P*’ be such a possibility predicate and thus the dual of the predicate ‘*N*’ read as ‘is necessary’. Applying our suggestion, we arrive at the principle

$$\forall t (T \neg N t \leftrightarrow P \neg Val(t) \vee \neg Sent(Val(t)))$$

Fortunately, this strategy proves to be successful and if further suitable axioms for the possibility predicate are added to  $MKF^-$ , we may prove a modal version of Theorem 3.2. Following these outlines we arrive at the theories  $BMKF$  and  $MKF$ . But first we define the language  $\mathcal{L}_{MKF}$ .

DEFINITION 3.4. *The language of  $\mathcal{L}_{MKF}$  is the language  $\mathcal{L}_{PATN}$  augmented by a further one-place predicate ‘*P*’.*

DEFINITION 3.5 ( $BMKF$ ). *The theory  $BMKF$  is the smallest modal extension of  $KF$  in the language  $\mathcal{L}_{MKF}$  (in the sense of Definition 2.6) which besides all axioms of ( $KF$ ), ( $Reg_N$ ), ( $ND$ ), ( $BF$ ), ( $K_N$ ) and ( $RN$ ) in the language  $\mathcal{L}_{MKF}$  contains the axioms:*

- ( $RP$ )  $\forall t (T P t \leftrightarrow P Val(t))$
- ( $DN$ )  $\forall t (T \neg N t \leftrightarrow P \neg Val(t) \vee \neg Sent(Val(t)))$
- ( $DP$ )  $\forall t (T \neg P t \leftrightarrow N \neg Val(t) \vee \neg Sent(Val(t)))$
- ( $Reg_P$ )  $\forall x, v, s, t (Sent(\forall v x) \rightarrow (Val(s) = Val(t) \rightarrow (Px(s/v) \leftrightarrow Px(t/v))))$

*$BMKF$  is closed under ( $TcR$ ) and ( $T-Nec$ ).  $BMKFC$  is  $BMKF$  except that ( $Cons$ ) has been added to the axioms of the theory. The theory  $BX_1, \dots, X_nMKF$  ( $BX_1, \dots, X_nMKFC$ ) is the theory  $BMKF$  ( $BMKFC$ ) except that the principles  $X_1$  to  $X_n$  have been added to the axioms of the theory.*

Clearly, ( $RP$ ), ( $DP$ ), and ( $Reg_P$ ) are just the analogues of ( $RN$ ), ( $DN$ ), and ( $Reg_N$ ) for the possibility predicate, and we take it that the former axioms are acceptable, if the latter are.

DEFINITION 3.6 ( $MKF$ ). *The theory  $MKF$  ( $MKFC$ ) is the theory  $BT'4'E'MKF$  ( $BT'4'E'MKFC$ ).*

Before turning back to semantic questions and the relation between  $MKF$  and what we call modal fixed-point semantics, we point to a general drawback of introducing a primitive possibility predicate which is that we cannot expect the modal square of opposition, that is, the interdefinability of the two modal predicates to hold within  $MKF$ . And indeed the following principles will not be theorems of  $MKF$  or  $MKFC$  nor will they be valid in the modal semantics to be defined:

- $\forall x (Sent(x) \rightarrow (Nx \leftrightarrow \neg P \neg x))$
- $\forall x (Sent(x) \rightarrow (\neg Nx \leftrightarrow P \neg x))$
- $\forall x (Sent(x) \rightarrow (\neg N \neg x \leftrightarrow Px))$
- $\forall x (Sent(x) \rightarrow (N \neg x \leftrightarrow \neg Px))$

Very roughly the failure of the modal square of opposition may be seen as consequence of the fact that “not true” ( $\neg T$ ) and “true that not” ( $T \neg$ ) are not in general equivalent in the case of  $KF$ . We shall come back to this shortcoming in the conclusion of the paper but now turn toward modal fixed-point semantics.

**3.3. Modal fixed-point semantics.** The basic idea behind the semantics is to combine Kripke’s fixed-point semantics (or theory of truth) with ideas from possible world semantics for modal operator logic. To this end we introduce the notion of a modal frame and an evaluation function. The evaluation function assigns a subset of  $\omega$  to each possible world which will serve as the interpretation of the truth predicate at that world. We then define a modal strong Kleene jump on evaluation functions relative to a frame.<sup>11</sup> Like the simple strong Kleene jump the modal strong Kleene jump will have fixed-points and we may obtain standard models for the theory  $BMKFC$  by appeal to these fixed-points.

We start giving the details of the semantics by introducing the notion of a frame, which is exactly the same notion as in modal revision semantics we discussed in our first paper.

**DEFINITION 3.7** (Frame, evaluation function). *Let  $W \neq \emptyset$  be a set of labeled natural number structures 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  $Val_F$ .*

Next we define the notion of a model induced by a frame and an evaluation function relative to a world. Notice that these models may serve as both strong Kleene models and also classical models for  $\mathcal{L}_{MKF}$ .

**DEFINITION 3.8** (Models for  $\mathcal{L}_{MKF}$ ). *Let  $F$  be a frame and  $f$  an evaluation function. Then,  $M_w = \langle w, f(w), Y_w, Z_w \rangle$  is a model of the language  $\mathcal{L}_{MKF}$  induced by  $F$  and  $f$  at a world  $w$ , where  $f(w)$  is the extension of the truth predicate,  $Y_w$  the extension of the necessity predicate and  $Z_w$  the extension of the possibility predicate with*

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

$$Z_w = \bigcup_{v \in \{wR\}} f(v).^{12}$$

$\{wR\}$  is short for  $\{v \in W : wRv\}$ .

Before we can define the modal strong Kleene jump on evaluation functions we need to introduce the modal strong Kleene satisfaction relation:

**DEFINITION 3.9** (Modal Strong Kleene Truth in a Model). *We explain the notion of truth in a model according to the strong Kleene scheme,  $\models_{SK}$ . Let  $M_w$  be a model induced by a frame  $F$  and an evaluation function  $f$  at a world  $w$  in the sense of Definition 3.8:*

- (i)  $M_w \models_{SK} s = t \Leftrightarrow s^{\mathbb{N}} = t^{\mathbb{N}}$
- (ii)  $M_w \models_{SK} s \neq t \Leftrightarrow s^{\mathbb{N}} \neq t^{\mathbb{N}}$
- (iii)  $M_w \models_{SK} Tt \Leftrightarrow t^{\mathbb{N}} \in f(w) \ \& \ t^{\mathbb{N}} \in \text{Sent}_{\mathcal{L}_{PATN}}$
- (iv)  $M_w \models_{SK} \neg Tt \Leftrightarrow (\neg t)^{\mathbb{N}} \in f(w) \ \text{or} \ t^{\mathbb{N}} \in \overline{\text{Sent}_{\mathcal{L}_{PATN}}}$

<sup>11</sup> Basically our construction is a variant of the construction of Halbach & Welch (2009).

<sup>12</sup>  $\bigcap$  is taken to be an operation on  $P(\omega)$ .

- (v)  $M_w \models_{SK} Nt \Leftrightarrow t^{\mathbb{N}} \in Y_w \ \& \ t^{\mathbb{N}} \in \text{Sent}_{\mathcal{L}_{PATN}}$
- (vi)  $M_w \models_{SK} \neg Nt \Leftrightarrow (\neg t)^{\mathbb{N}} \in Z_w \ \text{or} \ t^{\mathbb{N}} \in \overline{\text{Sent}_{\mathcal{L}_{PATN}}}$
- (viii)  $M_w \models_{SK} Pt \Leftrightarrow t^{\mathbb{N}} \in Z_w \ \& \ t^{\mathbb{N}} \in \text{Sent}_{\mathcal{L}_{PATN}}$
- (ix)  $M_w \models_{SK} \neg Pt \Leftrightarrow (\neg t)^{\mathbb{N}} \in Y_w \ \text{or} \ t^{\mathbb{N}} \in \overline{\text{Sent}_{\mathcal{L}_{PATN}}}$
- (x)  $M_w \models_{SK} \neg\neg\psi \Leftrightarrow M_w \models_{SK} \psi$
- (xi)  $M_w \models_{SK} \psi \wedge \chi \Leftrightarrow (M_w \models_{SK} \psi \ \text{and} \ M_w \models_{SK} \chi)$
- (xii)  $M_w \models_{SK} \neg(\psi \wedge \chi) \Leftrightarrow (M_w \models_{SK} \neg\psi \ \text{or} \ M_w \models_{SK} \neg\chi)$
- (xiii)  $M_w \models_{SK} \forall x\psi \Leftrightarrow \text{for all } n \in \omega (M_w \models_{SK} \psi(\bar{n}/x))$
- (xiv)  $M_w \models_{SK} \neg\forall x\psi \Leftrightarrow \text{there exists an } n \in \omega (M_w \models_{SK} \psi(\bar{n}/x))$

$t^{\mathbb{N}}$  is the interpretation of a term  $t$  in the standard model and,  $\text{Sent}_{\mathcal{L}_{PATN}}$  denotes the set of Gödel numbers of sentences and  $\overline{\text{Sent}_{\mathcal{L}_{PATN}}}$  its complement. If  $M_w \models \phi$  we say that  $\phi$  is true in the model induced by  $F$  and  $f$  at  $w$ . If it is important to keep track of the frame and the evaluation function, we write  $F, w \models_{SK}^f \phi$  instead of  $M_w \models_{SK} \phi$ . We say that  $\phi$  is true in the model induced by  $F$  and  $f$ ,  $F, f \models_{SK}$ , if and only if,  $\forall w \in W (F, w \models_{SK}^f \phi)$ .

Classical truth in a  $\mathcal{L}_{MKF}$ -model is defined as one would expect. We also use the same notation as for modal strong Kleene truth in a model but drop the ‘SK’ subscript from the double turnstyle denoting the satisfaction relation. Next we define the modal strong Kleene jump.

**DEFINITION 3.10 (Modal Strong Kleene Jump).** *Let  $F$  be a frame and  $\text{Val}_F$  the set of evaluation functions of  $F$ . The modal Strong Kleene jump  $\Theta_F$  is an operation on  $\text{Val}_F$  relative to  $F$  such that for all  $w \in W$*

$$[\Theta_F(f)](w) = \{\#\phi : F, w \models_{SK}^f \phi\}.$$

Importantly, the modal Strong Kleene jump is again a monotone operation:

**LEMMA 3.11 (Monotonicity).** *Let  $F$  be a frame. The jump  $\Theta_F$  is a monotone operation on  $\text{Val}_F$ , that is, for all  $f, g \in \text{Val}_F$ :*

$$f \leq g \Rightarrow \Theta_F(f) \leq \Theta_F(g),$$

where  $f \leq g :\Leftrightarrow \forall w \in W (f(w) \subseteq g(w))$ .

*Proof.* Clearly for all evaluation functions  $f, g$  with  $f(w) \subseteq g(w)$  for all  $w \in W$  we have  $[\Theta_F(f)](w) \subseteq [\Theta_F(g)](w)$  for all  $w \in W$  and thus  $\Theta_F(f) \leq \Theta_F(g)$ . □

The monotonicity of  $\Theta_F$  implies, assuming standard set theory like  $ZF$ , the existence of fixed-points, that is the existence of an evaluation function  $f \in \text{Val}_F$  such that,

$$\Theta_F(f) = f.$$

With all this at hand, we can give an adequacy result for  $BMKFC$ , but not for  $BMKF$ . The reason for this is that throughout we have assumed the modal axiom  $(K_N)$  to be a theorem of any modal extension of  $KF$  and, especially, of  $BMKF$ . However, for  $(K_N)$  to be valid in modal fixed-point semantics we must only consider consistent extensions of the truth predicate. Indeed, this was the main reason why we could interpret the modal predicate of modal extensions of  $KF$  as the truth predicate of  $KFC$  but not as the truth predicate of  $KF$  simpliciter. Consequently, if we wish to provide a more general adequacy result, we needed

to drop  $(K_N)$  from the axioms of  $BMKF$ . As we take  $(K_N)$  to be a constitutive modal axiom we shall not pursue this line any further.

In what is to come, we frequently use the following lemma that links strong Kleene truth in a model to the classical notion:

LEMMA 3.12. *Let  $F$  be a frame and  $f \in Val_F$  an evaluation function. Then, for all  $t \in Cterms_{\mathcal{L}_{MKF}}$*

- (i)  $F, w \models_{SK}^f Nt \Leftrightarrow F, w \models^f Nt$
- (ii)  $F, w \models_{SK}^f \neg Nt \Leftrightarrow F, w \models^f P \neg t$
- (iii)  $F, w \models_{SK}^f Pt \Leftrightarrow F, w \models^f Pt$
- (iv)  $F, w \models_{SK}^f \neg Pt \Leftrightarrow F, w \models^f N \neg t.$

*Proof.* Immediate from Definitions 3.8 and 3.9. □

We call an evaluation function  $f$  consistent, if for all  $w$ ,  $f(w)$  is consistent.  $BMKFC$  axiomatizes modal fixed-point semantics with respect to consistent evaluation functions.

THEOREM 3.13. *For all frames  $F$  and all consistent evaluation functions  $f \in Val_F$*

$$\Theta_F(f) = f \Leftrightarrow F, f \models BMKFC$$

*Proof.* For the left-to-right direction, we need to show that all theorems of  $BMKFC$  are true in the model induced by the frame  $F$  and the evaluation function  $f$  at an arbitrary world  $w$ . This is a standard induction on the length of a proof in  $BMKFC$  which we shall omit.

For the converse direction, we assume  $\forall w (F, w \models^f BMKFC)$  and need to show that  $f = \Theta_F(f)$ . To this end we show for an arbitrary world  $w$  that  $[\Theta_F(f)](w) = f(w)$ , which we do by induction on the positive built up of a sentence  $\phi$ , since due to  $(KF10)$  we know that only sentences are in  $f(w)$ . Again this induction proves to be routine and we only discuss the cases  $Nt$  and  $\neg Nt$ . In particular, the case  $\neg Nt$  proves to be instructive as it nicely illustrates the importance of the primitive possibility predicate for our construction.

- $\phi \doteq Nt$

$$\begin{aligned} \#Nt \in f(w) &\Leftrightarrow F, w \models^f T[Nt] && \text{Def. 3.8} \\ &\Leftrightarrow F, w \models^f Nt && (RN) \\ &\Leftrightarrow F, w \models_{SK}^f Nt && \text{Lemma 3.12} \\ &\Leftrightarrow \#Nt \in [\Theta_F(f)](w) && \text{Def. 3.10} \end{aligned}$$

- $\phi \doteq \neg Nt$

$$\begin{aligned} \#\neg Nt \in f(w) &\Leftrightarrow F, w \models^f T[\neg Nt] && \text{Def. 3.8} \\ &\Leftrightarrow F, w \models^f P \neg t && (DN) \\ &\Leftrightarrow F, w \models_{SK}^f \neg Nt && \text{Lemma 3.12} \\ &\Leftrightarrow \#\neg Nt \in [\Theta_F(f)](w) && \text{Def. 3.10} \end{aligned}$$

The atomic cases dealing with the possibility predicate work exactly parallel to the cases of the necessity predicate and the remaining cases may be found in any detailed presentation of the proof of Theorem 3.2. □

COROLLARY 3.14. *BMKFC (BMKF) is consistent.*

By Theorem 3.13 we have established a nice connection between *BMKFC* and modal fixed-point semantics, but we still lack such a connection for *MKFC*. Yet, we have all the prerequisites for establishing an adequacy result for *MKFC* with respect to modal fixed-point semantics. We just need to restrict the class of modal frames we consider since the modal axioms ( $T'$ ), ( $4'$ ), and ( $E'$ ) are only guaranteed to be true on frames that have certain properties. We state the connections between the modal axioms and the frame properties after introducing some terminology.

DEFINITION 3.15. *Let  $F$  be a frame and  $\phi \in \text{Sent}_{\mathcal{L}_{MKF}}$ . We write  $F \models \phi$  iff for all consistent evaluation functions  $f$  with  $\Theta_F(f) = f$  we have  $F, f \models \phi$ .*

THEOREM 3.16. *Let  $F = \langle W, R \rangle$  be a frame. Then*

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

*Sketch of a proof.* The proof is identical to the parallel result for modal revision semantics we discussed in the prequel to this paper and we refer the reader to this paper for further detail. The most important observation for the proof is that a truth teller sentence is either in the extension of the truth predicate at a world from the start on or it will never be in the extension of the truth predicate at that world. □

COROLLARY 3.17. *Let  $F = \langle W, R \rangle$  be a frame. Then,*

$$R \text{ is an equiv. relation} \Leftrightarrow F \models MKFC$$

COROLLARY 3.18. *MKFC (MKF) is consistent.*

This ends our discussion of modal fixed-point semantics and its connection to *MKF* or rather to *MKFC*.

**§4. The “Kripke reduction” and *MKF*.** Eventually, we turn our focus toward the so-called “Kripke reduction”, that is, the idea of understanding the modal predicate(s) of *MKF* as a truth predicate modified by a modal operator. Kripke originally proposed to understand the necessity predicate ( $N$ ) as the complex predicate ‘is necessarily true’ ( $\Box T$ ). Since Kripke did not envisage the introduction of a primitive possibility predicate he remained mute on how this predicate ought to be defined in terms of a truth predicate and a modal operator. However, we may understand the possibility predicate as ‘is possibly true’ ( $\Diamond T$ ), that is as a complex predicate obtained by modifying a truth predicate by a possibility operator. In our framework, the question of whether the modal predicates of *MKF* lend themselves to an understanding along the lines of the “Kripke reduction” transforms into the question of whether *MKF* can be reduced to some theory of truth formulated in the language  $\mathcal{L}_{PAT}^\Box$ , that is  $\mathcal{L}_{PAT}$  augmented by a modal operator  $\Box$ ,<sup>13</sup> assuming some reasonable underlying modal logic. Indeed, *MKF* should not be reduced to any theory of truth but to *KF* formulated in  $\mathcal{L}_{PAT}^\Box$  for otherwise the “truth” predicate of the target theory

<sup>13</sup> All formation rules of  $\mathcal{L}_{PAT}^\Box$  are standard.

would not be a truth predicate in the sense of *MKF* and, arguably, we would have failed to “reduce” the modal predicate to the truth predicate modified by the modal operator.

Moreover, as in the previous section we need to assume our underlying truth predicate to be consistent as otherwise it seems unlikely that the (intended) translation of  $K_N$ , that is

$$\forall x, y(\text{Sent}_{\mathcal{L}_{MKF}}(x \rightarrow y) \rightarrow (\Box T(x \rightarrow y) \rightarrow (\Box Tx \rightarrow \Box Ty)))$$

will be derivable in *KF* as formulated in  $\mathcal{L}_{PAT}^\Box$ . If we assume *KFC* as our underlying theory of truth, this principle will follow from ( $K_T$ ) and the principles of modal operator logic (*Nec*) and (*K*). But, as pointed out,  $K_T$  is not a theorem of *KF*, and thus, we cannot use this derivation, if we assume only *KF*. Correspondingly, we take the underlying target theory to be an extension of *KFC*, which suggests carrying out the reduction for *MKFC* rather than for *MKF*.

If we formulate *KFC* in  $\mathcal{L}_{PAT}^\Box$  we need to supplement the theory *KFC* by an axiom describing the interaction of the truth predicate and the modal operator. Indeed, we need two axioms, one for the positive and one for the negative modal case, as *KF* is built on the idea of a positive inductive definition. The resulting extension of *KFC* will be called *KFC* $\Box$ :

DEFINITION 4.1 (*KFC* $\Box$ ). *Let  $\mathcal{L}_{PAT}^\Box$  be  $\mathcal{L}_{PAT}$  augmented by a modal operator. The theory *KFC* $\Box$  consists of all axioms of *KFC* in the language  $\mathcal{L}_{PAT}^\Box$  together with the following principles*

$$\begin{aligned} (IA\Box) \quad & \forall x(\text{Sent}(x) \rightarrow (T \Box x \leftrightarrow \Box Tx)) \\ (IA\Diamond) \quad & \forall x(\text{Sent}(x) \rightarrow (T \neg \Box x \leftrightarrow \Diamond T \neg x)) \end{aligned}$$

Throughout we assume  $\Diamond\phi$  to be a notational abbreviation of,  $\neg\Box\neg\phi$ .<sup>14</sup>

*KFC* $\Box$  gives us no information concerning the modal properties of the  $\Box$ -operator and, again, the set of theorems of *KFC* $\Box$  will depend on the modal logic assumed. Since *MKF* is meant to capture a modal notion related to the one characterized by the modal operator logic *S5* we shall assume the underlying modal operator logic to be *QS5*, that is *S5* in standard first-order logic together with the so-called principles “necessity of distinctness” and the “Barcan formula”

$$\begin{aligned} (ND\Box) \quad & \forall x, y(x \neq y \rightarrow \Box x \neq y) \\ (BF\Box) \quad & \forall x\Box\phi \rightarrow \Box\forall x\phi \end{aligned}$$

We write  $KFC\Box \vdash_{\mathcal{S}} \phi$  iff  $\phi$  is derivable in *KFC* $\Box$  assuming the modal operator logic  $\mathcal{S}$ . In particular we write  $KFC\Box \vdash_{QS5} \phi$  iff  $\phi$  is derivable in *KFC* $\Box$  assuming the underlying modal logic to be *QS5*.

Importantly, we may already derive the rigidity of syntax assuming *QK*—the basic modal system (*K*) in standard first-order logic together with (*ND* $\Box$ ) and (*BF* $\Box$ ):

LEMMA 4.2. *For all  $\phi(x_1, \dots, x_n) \in \mathcal{L}_{PA}$*

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

*Proof.* By induction over the positive built up of  $\phi$ . □

<sup>14</sup> *KFC* $\Box$  is consistent independent of the (consistent) modal operator logic assumed for we can interpret *KFC* $\Box$  in *KFC* if the latter is consistent. Depending on the modal logic under consideration a formula  $\Box\phi$  is either interpreted by  $\phi$  or by  $s = s$  for some closed term  $s$ .

After having specified the target theory of the “Kripke-reduction” we turn our attention toward the translation function we use in carrying out the reduction. We use a slight modification of the translation function employed by Halbach & Welch (2009). As mentioned, we will translate ‘ $N$ ’ to  $\Box T$ ’ and ‘ $P$ ’ to ‘ $\neg\Box\neg T$ ’, but leave the remaining vocabulary fixed. We appeal to the recursion theorem to establish the existence of such a translation function:

LEMMA 4.3 (Halbach and Welch). *There is a translation function  $\eta : \mathcal{L}_{MKF} \longrightarrow \mathcal{L}_{PAT}^\Box$  with the following properties*

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

where  $\eta^\bullet$  represents  $\eta$  in PA.

We have all the prerequisites for stating our main theorem:

THEOREM 4.4. *Let  $\eta$  be a translation function as specified in Lemma 4.3. Then for all  $\phi \in \mathcal{L}_{PATN}$ :*

$$MKFC \vdash \phi \Rightarrow KFC\Box \vdash_{QS5} \eta(\phi)$$

For the proof of Theorem 4.4 we make crucial use of the following lemma:

LEMMA 4.5. *Let QS5 be the underlying modal logic. Then the following are theorems of KF $\Box$ :*

- (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)  $\forall x, v, s, t (Sent(\forall vx) \rightarrow (Val(s) = Val(t) \rightarrow (\Diamond Tx(t/v) \leftrightarrow \Diamond Tx(s/v))))$
- (iii)  $\forall s, t (Val(s) \neq Val(t) \rightarrow \Box Ts \neq t)$
- (iv)  $\forall x, v (Sent(\forall vx) \rightarrow (\forall t \Box Tx(t/v) \rightarrow \Box T \forall vx))$
- (v)  $\forall x, y (Sent(x \rightarrow y) \rightarrow (\Box T(x \rightarrow y) \rightarrow (\Box Tx \rightarrow \Box Ty)))$
- (vi)  $\forall t (T \Box T t \leftrightarrow \Box T Val(t))$
- (vii)  $\forall t (T \Diamond T t \leftrightarrow \Diamond T Val(t))$
- (viii)  $\forall t (T \neg \Box T t \leftrightarrow \Diamond T \neg Val(t) \vee \neg Sent(Val(t)))$
- (ix)  $\forall t (T \neg \Diamond T t \leftrightarrow \Box T \neg Val(t) \vee \neg Sent(Val(t)))$
- (x)  $\forall x (Sent(x) \rightarrow (\Box Tx \rightarrow Tx))$
- (xi)  $\forall t (T \Box T t \rightarrow \Box T \Box T t)$
- (xii)  $\forall t (T \neg \Box T t \rightarrow \Box T \neg \Box T t)$

*Proof.* (i) By induction on the positive complexity of a formula of  $\mathcal{L}_{PAT}^\Box$  we derive ( $Reg_T$ ), then use ( $Nec$ ) and Lemma 4.2. (ii) similar to item (i). (iii), (iv), and (v) follow, respectively, from ( $KF2$ ), ( $KF6$ ) and ( $K_T$ ) and the modal properties of  $QK$  and Lemma 4.2; (vi) by ( $IA_\Box$ ), ( $KF8$ ), ( $Nec$ ), ( $K$ ) and Lemma 4.2; (vii) as item (vi) but using ( $IA_\Diamond$ ) and ( $KF2$ ) instead of ( $IA_\Box$ ). (x) is a straightforward consequence of the modal axiom ( $T$ ). It remains to show (viii), (ix), (xi), and (xii). We reason as follows—again note that  $\phi(t)$



is short for  $Cterm(x) \rightarrow \phi(x)$  and that if we apply  $(Nec)$  due to  $(K)$  and Lemma 4.2 we may derive  $Cterm(x) \rightarrow \Box\phi(x)$ :

(viii) We start with an instance of  $(IA\Diamond)$ :

1.  $T \neg \Box T t \leftrightarrow \Diamond T \neg T t$  (IA $\Diamond$ )
2.  $\Diamond T \neg T t \leftrightarrow \Diamond T \neg Val(t) \vee \neg Sent(Val(t))$  (KF9), (Nec), (K), Lemma 4.2
3.  $T \neg \Box T t \leftrightarrow \Diamond T \neg Val(t) \vee \neg Sent(Val(t))$  1, 2
4.  $\forall t(T \neg \Box T t \leftrightarrow \Diamond T \neg Val(t) \vee \neg Sent(Val(t)))$  3

(ix) We first observe that  $T \neg \Diamond T t$  is only an abbreviation of  $T \neg \neg \Box \neg T t$

1.  $T \neg \neg \Box \neg T t \leftrightarrow T \Box \neg T t$  (KF2)
2.  $T \Box \neg T t \leftrightarrow \Box T \neg T t$  (IA $\Box$ )
3.  $\Box T \neg T t \leftrightarrow \Box T \neg Val(t) \vee \neg Sent(Val(t))$  (KF9), (Nec), (K), Lemma 4.2
4.  $T \neg \neg \Box \neg T t \leftrightarrow \Box T \neg Val(t) \vee \neg Sent(Val(t))$  1, 2, 3
5.  $\forall t(T \neg \neg \Box \neg T t \leftrightarrow \Box T \neg Val(t) \vee \neg Sent(Val(t)))$  4

(xi) We start with an instance of the modal axiom (4):

1.  $\Box T T t \rightarrow \Box \Box T T t$  (4)
2.  $\Box T T t \leftrightarrow T \Box T t$  (IA $\Box$ )
3.  $\Box \Box T T t \leftrightarrow \Box T \Box T t$  2, (Nec), (K), Lemma 4.2
4.  $T \Box T t \rightarrow \Box T \Box T t$  1, 2, 3
5.  $\forall t(T \Box T t \rightarrow \Box T \Box T t)$  4

(xii) The proof starts by an instance of the modal axiom (E):

1.  $\Diamond T \neg T t \rightarrow \Box \Diamond T \neg T t$  (E)
2.  $\Diamond T \neg T t \leftrightarrow T \neg \Box T t$  (IA $\Diamond$ )
3.  $\Box \Diamond T \neg T t \leftrightarrow \Box T \neg \Box T t$  (Nec), (K), Lemma 4.2
4.  $T \neg \Box T t \rightarrow \Box T \neg \Box T t$  1, 2, 3
5.  $\forall t(T \neg \Box T t \rightarrow \Box T \neg \Box T t)$  4

□

Before we start the proof of the main theorem we state some important properties of  $\eta$  that we can prove in  $PATN$  and thus in  $KFC\Box$ .

LEMMA 4.6. *The following is provable in  $PATN$ , and thus in  $KFC\Box$ :*

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

*Proof.* By formalizing the properties of  $\eta$ . □

*Proof of Theorem 4.4.* By a routine induction over the length of a proof in  $MKFC$  assuming  $QS5$ . The start of the induction is established by Lemmas 4.5 and 4.6. For the induction step we have to deal with two cases for a sentence may be derived by  $(T-Nec)$  or  $(TcR)$ . In the case of  $(T-Nec)$  we may assume that  $KFC\Box \vdash_{QS5} \eta(T[\phi])$ . By necessitation we then infer  $KFC\Box \vdash_{QS5} \Box\eta(T[\phi])$  but since  $\rho(N[\phi]) \doteq \Box\eta(T[\phi])$  this ends the case of  $(T-Nec)$ .

For  $(TcR)$  we assume  $PKF\Box \vdash_{QS5} \eta(tc([\phi]))$ , that is  $PKF\Box \vdash_{QS5} T[\neg\eta(\phi)] \leftrightarrow \neg T[\eta(\phi)]$ . Then, by  $(KF8)$  and  $(KF9)$ ,  $PKF\Box \vdash_{QS5} T \neg [T[\eta(\phi)]] \leftrightarrow \neg T[T[\eta(\phi)]]$  and by  $(Nec)$ ,  $(K)$  and propositional logic  $PKF\Box \vdash_{QS5} \Diamond T \neg [T[\eta(\phi)]] \leftrightarrow \neg \Box T [T$

$\lceil \eta(\phi) \rceil$ . Using  $(IA\Box)$  and  $(IA\Diamond)$  we derive the desired  $PKF\Box \vdash_{QS5} T \lceil \lceil \Box T \lceil \eta(\phi) \rceil \rceil \leftrightarrow \lceil \lceil \neg T \lceil \Box T \lceil \eta(\phi) \rceil \rceil \rceil$ .  $\square$

With the proof of this result we end our discussion of the “Kripke reduction” and thereby the third and last theme of our investigation. To conclude, we shall sum up our results, compare the two modal theories *MFS* and *MKF* and give a quick assessment of our general strategy.

**§5. Conclusion.** In this paper we tested a strategy for developing modal theories over theories of truth that we presented in some detail in a prequel to this paper. To this end, we applied the strategy to the axiomatic theory of truth *KF*. Overall the modal theory we constructed over *KF* was fairly attractive.<sup>15</sup> It preserves theoremhood with respect to modal operator logic and, therefore, may be considered as an adequate modal theory from the perspective of modal operator logic. We also showed that *MKFC* can be viewed as an axiomatization of modal fixed-point semantics. That is, *MKFC* stands in the same relation to modal fixed-point semantics as *KF* stands to simple fixed-point semantics, that is, Kripke’s theory of truth. *KF* is often taken to be an axiomatization of Kripke’s theory of truth. We take this result to be a strong point in favor of the theory *MKFC* for, arguably, the fact that *KF* axiomatizes Kripke’s theory of truth is one reason for the popularity of *KF* amongst theorists of truth. Finally, the Kripke reduction gives us a nice explanation of the success of our strategy. Since, the modal predicate can be understood as a truth predicate modified by a modal operator, there remains just one paradoxical concept, namely the concept of truth. This is also reflected within the semantic construction, since the interpretation of the modal predicates is obtained by quantifying over the interpretations of the truth predicate.

However, when we turned to (modal) fixed-point semantics a less-than-ideal feature of the theory of modality arising from *KF* became apparent. In order to establish the adequacy of *MKFC* with respect to modal fixed-point semantics, we had to introduce a primitive possibility predicate for which the modal square of opposition breaks down. The reason for this shortfall may be located in the divergence of the inner and outer logic of *KF*, that is, more precisely, in the fact that the truth predicate of *KF* does not freely commute with negation. Indeed, this fact is nicely illustrated by the properties of our translation function  $\eta$  of Lemma 4.3 which translates ‘ $P[\phi]$ ’ and ‘ $N[\phi]$ ’ to ‘ $\Diamond T[\eta(\phi)]$ ’ and ‘ $\Box T[\eta(\phi)]$ ’, respectively. Consequently, whereas ‘ $\neg N[\neg\phi]$ ’ is translated to ‘ $\neg\Box T[\neg\eta(\phi)]$ ’ the translation of ‘ $P[\phi]$ ’ will be the formula ‘ $\neg\Box\neg T[\eta(\phi)]$ ’, and since the truth predicate of *KF* does not commute freely with negation, these formulas will not, in general, be equivalent. To the rescue of *MKF* (*MKFC*) one should point out that the modal square of opposition will be provable for total and consistent sentences, and thus in particular for all the sentences modal operator logic is concerned with.<sup>16</sup> Similarly, the modal square of opposition

<sup>15</sup> The attentive reader may complain that we haven’t considered the effects of adding contingent vocabulary to the language in testing our strategy and in formulating the theory *MKF*. Contingent vocabulary, however, can be dealt with along the lines we have sketched out in the first paper of our investigation.

<sup>16</sup> Actually, this claim is correct only if the theories *MKF* (*MKFC*) are supplemented by the following rule for the possibility predicate:

$$\frac{tc(\lceil \phi \rceil)}{tc(\lceil P[\phi] \rceil)}$$

will be valid for all grounded sentences within modal fixed-point semantics. The modal predicates of *MKF* (*MKFC*), therefore, behave just like the “ordinary” notions of necessity and possibility on the “normal” sentences but show deviant behavior on the ungrounded sentences. Yet, bearing Kripke’s theory of truth in mind this is just what one would expect as within Kripke’s theory the truth predicate behaves in the intuitive way on the grounded sentences within Kripke’s theory but shows deviant behavior on the ungrounded ones.

Nonetheless, the introduction of a primitive possibility predicate does not count as an attractive feature of a theory of modality and in formulating the theory Modal Friedman-Sheard which we investigated in “Modality and Axiomatic Theories of Truth I”, we could dispense with the introduction of a such a primitive possibility predicate. Friedman-Sheard, the theory of truth underlying Modal Friedman-Sheard, is a symmetric theory of truth. The truth predicate commutes with negation, and is closed under the rules of *T*-introduction and *T*-elimination. This guarantees that the inner and the outer logic are identical. In the case of Friedman-Sheard the inner and the outer logic are both are classical. Our findings suggest that symmetric theories of truth are better suited for the present strategy, which intimately binds the modal theory to the theory of truth assumed. This view is supported by our explanation of the failure of the modal square opposition using the Kripke reduction: as long as the truth predicate commutes with negation the necessity and the possibility predicate will be interdefinable. As the modal predicates of both modal theories we have presented in the course of our investigation may be understood along the lines of Kripke’s proposal, this illustration seems to generalize. Over symmetric theories of truth one modal predicate suffices, but if the underlying theory of truth is nonsymmetric the introduction of a further primitive modal predicate might be required. While these observations suggest that the strategy we developed is particularly well suited for symmetric theories of truth proponents of Kripke-Feferman should not dismiss *MKF* because of the primitive possibility predicate we have assumed. Rather, we take it that the primitive possibility predicate should be considered as nothing but a symptom of the slightly deviant characteristics of the truth predicate of *KF*.

Summing up, our strategy of avoiding introduction and elimination of the modal predicate independently of the truth predicate has proven rather successful since the modal theories “Modal Friedman-Sheard” and “Modal Kripke-Feferman” can be viewed as apt theories of modality, if the underlying theories of truth are accepted. Moreover, the modal predicate was axiomatized by appeal to, arguably, rather plausible modal principles and without resorting to some unprincipled (or principled) typing restriction or by limiting the expressive power of the language and theory. We have, thus, provided a predicate approach to modality without appeal to any syntactic trickery or typing restriction.

**§6. 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

- Cantini, A. (1989). Notes on formal theories of truth. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **35**, 97–130.
- Feferman, S. (1991). Reflecting on incompleteness. *The Journal of Symbolic Logic*, **56**, 1–47.
- Feferman, S. (2008). Axioms for determinateness and truth. *Review of Symbolic Logic*, **1**, 204–217.
- Halbach, V. (2011). *Axiomatic Theories of Truth*. New York: Cambridge University Press.
- 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.
- Heck, R. G. (2007). Self-reference and the languages of arithmetic. *Philosophia Mathematica*, **15**(3), 1–29.
- McGee, V. (1991). *Truth, Vagueness and Paradox*. Indianapolis, IN: Hackett Publishing Company.
- Reinhardt, W. N. (1985). Remarks on significance and meaningful applicability. In L. De Alcantara, editor. *Mathematical Logic and Formal Systems*, Lecture Notes in Pure and Applied Mathematics, vol. 94, pp. 227–242.
- Reinhardt, W. N. (1986). Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of Philosophical Logic*, **15**, 219–251.
- Stern, J. (2012). *Toward Predicate Approaches to Modality*. PhD thesis, University of Geneva.

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