Modal-Epistemic Arithmetic and the Problem of Quantifying In

Jan Heylen

Abstract

The subject of this article is Modal-Epistemic Arithmetic, a theory introduced by Horsten to interpret Epistemic Arithmetic, which in turn was introduced by Shapiro to interpret Heyting Arithmetic. I will show how to interpret Modal-Epistemic Arithmetic in Epistemic Arithmetic and I will show how to get rid of a particular Platonist assumption. Then I will discuss models for Modal-Epistemic Arithmetic in light of the problems of logical omniscience and logical competence. Awareness models, impossible worlds models and syntactical models have been introduced to deal with the first problem. Certain conditions on the accessibility relations are needed to deal with the second problem. I go on to argue that those models are subject to the problem of quantifying in, for which I will provide a solution.

1 Introduction

In this article I will describe and discuss Modal-Epistemic Arithmetic. In the first section I will present Shapiro's Epistemic Arithmetic, which was introduced to give a classical interpretation of Heyting Arithmetic. I will explain why its classical models are either inadequate or result in arithmetical omniscience and why it is preferable to analyse Shapiro's notion of knowability as a complex notion.

In the second section I will describe and discuss Horsten's Modal-Epistemic Arithmetic, which was introduced to interpret Epistemic Arithmetic and to show that analysing the complex notion of knowability into an epistemic and a modal component yields a philosophical pay-off. I will show that Modal-Epistemic Arithmetic can be interpreted in Epistemic Arithmetic as well.

In the third and final section I turn to the problem of finding adequate models for Modal-Epistemic Arithmetic and, indirectly, for Epistemic Arithmetic. This problem will be approached by discussing: first, Modal-Epistemic Sentential Logic (**MESL**); second, Modal-Epistemic First-Order Logic (**MEFOL**); third, the extension of the latter with identity, individual constants and function symbols; fourth and finally, Modal-Epistemic Arithmetic (**MEA**). A modal-epistemic logic is confronted with the related problems of logical and arithmetical omniscience on the one hand and the problem of logical incompetence on the other hand. Different types of models have been proposed to solve the first kind of problem, viz. awareness models, impossible worlds models and syntactical models. Horsten's models for **MEA** are of the syntactical variety and as such they can be used to solve the problem of logical omniscience. Moreover, his models provide a solution to the problem of logical competence as well. It will be shown how to implement the same type of solution to awareness models and impossible worlds models. A modal-epistemic first-order logic is confronted with the problem of quantifying in, a problem shared by syntactical models, awareness models and impossible worlds models. I will introduce a solution to this problem. After briefly discussing identity and complex terms, I will critically discuss Horsten's models and provide an alternative based on the previous discussions.

2 Shapiro's Epistemic Arithmetic

In the first subsection I will briefly describe Shapiro's Epistemic Arithmetic. In the second subsection I will explain the difficulty of finding adequate but non-trivial models for his theory.

2.1 Theory

The language of Peano Arithmetic ($\mathcal{L}_{\mathbf{PA}}$) contains the individual constant $\mathbf{0}$, the one-place function symbol $\mathbf{s}$, the two-place function symbols $+$, $\cdot$, and the two-place relation symbol $<$ (Boolos, Burgess, and Jeffrey 2003, p. 103). The set of well-formed terms and the set of well-formed formulas are given their usual recursive definitions. The axiomatic theory of $\mathbf{PA}$ is described in (Boolos, Burgess, and Jeffrey 2003, p. 208).

The language of Epistemic Arithmetic $\mathbf{EA}$, denoted $\mathcal{L}_{\mathbf{EA}}$, is $\mathcal{L}_{\mathbf{PA}} \cup \{\mathbf{K}\}$. The definition of the set of well-formed formulas contains one additional clause: if ψ is a well-formed formula, then so is $\mathbf{K}\psi$. The theory $\mathbf{EA}$ contains $\mathbf{PA}$ and, in addition, the following axiom schemes and rule:[1]

E1 $\mathbf{K}(\phi \rightarrow \psi) \rightarrow (\mathbf{K}\phi \rightarrow \mathbf{K}\psi)$

E2 $\mathbf{K}\phi \rightarrow \phi$

E3 $\mathbf{K}\phi \rightarrow \mathbf{K}\mathbf{K}\phi$

E4 $\vdash_{\mathbf{EA}} \phi \quad \Rightarrow \quad \vdash_{\mathbf{EA}} \mathbf{K}\phi$

The theory $\mathbf{EA}$ is the closure of $\mathbf{PA}$ and the above axiom schemes and rule under classical first-order logic with identity.

The philosophical relevance of $\mathbf{EA}$ depends partly on the fact that there is a faithful translation from the language of Heyting Arithmetic ($\mathbf{HA}$) to $\mathbf{EA}$ which mimics the so-called 'Brouwer-Heyting-Kolmogorov'-interpretation of $\mathbf{HA}$ ((Shapiro 1985), (Goodman 1984), (Flagg and Friedman 1986)). See (Smoryński 1991) and (Horsten 1998) for a philosophical discussion.

2.2 Models for Epistemic Arithmetic

The classical model theory for epistemic logic was developed by Hintikka.[2] A Kripke model $\mathcal{M}$ for epistemic sentential logic is a triple $\langle W, R, V \rangle$, with W a non-empty set, whose elements will be called 'possible worlds', R a binary relation on W, which will be called 'accessibility relation', and V a function from W and the set of sentence letters to truth values. The notion of truth of a formula ϕ at a world w in a Kripke model $\mathcal{M}$ ($\mathcal{M}, w \models \phi$) is defined recursively as follows: if ϕ is a proposition letter p, then $\mathcal{M}, w \models \phi$ if and only if $V(p, w) = 1$; the truth clauses for the negations and conjunctions are as expected; if ϕ is a well-formed formula of the form $\mathbf{K}\psi$, then $\mathcal{M}, w \models \phi$ if and only if $\mathcal{M}, w' \models \psi$ for all w' such that wRw'. The axiom scheme $\mathbf{E1}$ is valid and the rule $\mathbf{E4}$ is validity-preserving. If R is reflexive and transitive, then $\mathbf{E2}$ and $\mathbf{E3}$ are also valid. Soundness and completeness can be proved for the epistemic sentential logical fragment of $\mathbf{EA}$.

The simplest Kripke-style model for epistemic first-order logic without identity is a quadruple $\langle W, R, D, V \rangle$, with W and R as before and with D a non-empty set and V a function from the set of n-place predicate symbols and W to $\wp(D^n)$ (Hughes and Cresswell 1996, ch. 13). An

---

[1] In (Shapiro 1985) the logical fragment of $\mathbf{EA}$ is presented as a natural deduction system. However, in (Horsten 1994) the logical fragment of Modal-Epistemic Arithmetic ($\mathbf{MEA}$) is presented as an axiomatic system. In order to facilitate comparisons, I will choose one proof-style, viz. the Hilbert style of proof theories.

[2] See (Hughes and Cresswell 1996) for a textbook presentation.

assignment function a is a function from individual variables to elements of D. An x-alternative of a, written as $a' \sim_x a$, is an assignment exactly like a except for the fact that it possibly assigns a different element to x. The notion of satisfaction at a world w in a model $\mathcal{M}$ given an assignment a ($\mathcal{M}, w, a \models$) is similar to the notion defined before, but $\mathcal{M}, w(') \models$ is uniformly replaced by $\mathcal{M}, w('), a$. The clause for atomic well-formed formulas is replaced by: if ϕ is of the form $P(x_1, \ldots, x_n)$ (with P an n-place predicate symbol from the language), then $\mathcal{M}, w, a \models \phi$ if and only if $\langle a(x_1), \ldots, a(x_n) \rangle \in V(P, w)$. Moreover, the following clause is added: if ϕ is of the form $\forall x \psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w, a' \models \psi$ for all $a' \sim_x a$. Soundness and completeness can be proved for the epistemic first-order logical fragment of **EA**. One can also introduce individual constants, function terms and the identity relation (Hughes and Cresswell 1996, ch. 17).

It was argued in (Shapiro 1985, p. 23) that, if one uses Kripke-models for **EA**, then either those models are inadequate or one gets arithmetical omniscience. Let ψ be an arithmetical formula that is neither knowable to be true nor knowable to be false at w. It follows that there must be an accessible possible world w' where ψ is false and an accessible possible world w'' where ψ is true. Either ψ or $\neg \psi$ is true in the standard model of arithmetic (Boolos, Burgess, and Jeffrey 2003, p. 104). If ψ is true in the standard model, then it cannot be the case that ψ gets a standard interpretation at w'. But according to Shapiro that means that w' is not a real epistemic alternative. Mathematicians who do not know whether ψ are not uncertain about the interpretation of ψ but about the truth of ψ. In other words, if one cannot model lack of knowledge of arithmetical formulas as they are standardly interpreted, then the interpretation is inadequate. Of course, one could stipulate that the arithmetical vocabulary gets a standard interpretation at every possible world. But then all arithmetical truths will be true at every possible world and, hence, all arithmetical truths will be knowable. If one were to drop the assumption that there is an arithmetical formula such that it and its negation are unknowable and if one accepts that knowability is factive, then one arrives once again at the conclusion that all arithmetical truths are knowable. One can then prove for all $\phi \in \mathcal{L}_{\mathbf{EA}}$, all possible worlds w, and all assignments a that $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w', a \models \phi$. It follows for all $\phi \in \mathcal{L}_{\mathbf{EA}}$, all possible worlds w, and all assignments a that $\mathcal{M}, w, a \models \Box \phi$ if and only if $\mathcal{M}, w, a \models \phi$. In other words, the models become trivial. To sum up, the interpretation is either inadequate or trivial.

The notion of knowability is a primitive notion in the language $\mathcal{L}_{\mathbf{EA}}$. Yet it is clear that it is, in effect, a complex notion: it has a modal component, viz. the *possibility* of having knowledge, and it has an epistemic component, viz. having *knowledge*. It would be better to analyse the concept of knowability into those two components, so as to be able to distinguish and compare modality and knowledge and to study their interaction. We will see that this is philosophically fruitful.

In the next section I will describe and discuss Modal-Epistemic Arithmetic (**MEA**), a theory developed to investigate the interaction between modality and knowledge. The language of **EA** can be faithfully translated into the language of **MEA**. In section 4, I will discuss the non-triviality and adequacy of the models of **MEA**. Given non-trivial and adequate models of **MEA**, the latter can be used to give a non-trivial and adequate interpretation of the language of **EA**.

3   Horsten's Modal-Epistemic Arithmetic: Theory

Horsten developed **MEA** in (Horsten 1993) and (Horsten 1994). The language of **MEA** is $\mathcal{L}_{\mathbf{PA}} \cup \{\Diamond, \mathbf{P}\}$, where $\mathcal{L}_{\mathbf{PA}}$ is as before the language of arithmetic, $\Diamond$ is a metaphysical possibility operator and $\mathbf{P}$ is the informal proof or knowledge operator. The theory **MEA** consists of

classical first-order logic with identity, **PA**, the modal system **S5** and the following axiom schemes and rule:[3]

ME1 $\Diamond\phi \rightarrow \phi$ for all $\phi \in \mathcal{L}_{\mathbf{PA}}$

ME2 $\mathbf{P}\phi \rightarrow \phi$

ME3 $\mathbf{P}\phi \rightarrow \mathbf{PP}\phi$

ME4 $\vdash \phi \quad \Rightarrow \quad \vdash \Diamond\mathbf{P}\phi$

ME5 $\Diamond\mathbf{P}(\phi \rightarrow \psi) \rightarrow (\Diamond\mathbf{P}\phi \rightarrow \Diamond\mathbf{P}\psi)$

This concludes the description of the proof theory of **MEA**.[4]

There is a natural and faithful translation of the language of **EA** into the language of **MEA**.

Definition 1. *Let F be a translation from $\mathcal{L}_{\mathbf{EA}}$ to $\mathcal{L}_{\mathbf{MEA}}$ as follows:*

1. *if ϕ is an atomic formula, then $\phi^F = \phi$;*

2. *if ϕ is $\neg\psi$, then $\phi^F = \neg\psi^F$;*

3. *if ϕ is $\psi \rightarrow \theta$, then $\phi^F = \left(\psi^F \rightarrow \theta^F\right)$;*

4. *if $\phi = \exists x\psi$, then $\phi^F = \exists x\psi^F$;*

5. *if ϕ is $\mathbf{K}\psi$, then $\phi^F = \Diamond\mathbf{P}\psi^F$.*

The above translation is natural, since the **K**-operator was interpreted as a knowability operator. It was noted that the notion of knowability has an epistemic component and a modal component. The F-translation makes those two components visible. The diamond operator is interpreted as a possibility operator and the **P**-operator is interpreted as a knowledge operator. Moreover, the F-translation is faithful, which will be proved in what follows. The following lemmas can be found in (Horsten 1993, ch. 4) and (Horsten 1994, p. 286-287).

Lemma 1 (Horsten)**.** *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, $\vdash_{\mathbf{MEA}} \Diamond\phi^F \rightarrow \phi^F$*

Proof. The proof is by induction on ϕ^F. The atomic case can be proved with the help of **ME1**. The negation case follows from the induction hypothesis and the modal axiom schemes **T** and **5**. The conjunction case follows from the induction hypothesis and the modal axiom scheme **K**. The case of $\phi = \forall x\psi^F$ follows from the induction hypothesis and quantified **K**. The case of $\phi = \Diamond\mathbf{P}\psi^F$ follows from the modal axiom scheme **4**. $\square$

Lemma 2 (Horsten)**.** *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, $\vdash_{\mathbf{MEA}} (\Box\phi \rightarrow \phi)^F$, i.e. $\vdash_{\mathbf{MEA}} \Diamond\mathbf{P}\phi^F \rightarrow \phi^F$*

Proof. The lemma follows from axiom scheme **ME2**, modal axiom scheme **K**, and lemma 1. $\square$

Lemma 3 (Horsten)**.** *For all $\phi \in \mathcal{L}_{\mathbf{MEA}}$, $\vdash_{\mathbf{MEA}} \Diamond\mathbf{P}\phi \rightarrow \Diamond\mathbf{P}\Diamond\mathbf{P}\phi$*

Proof. $\mathbf{P}\phi \rightarrow \mathbf{PP}\phi$ is an axiom scheme. By a derived rule of modal system **K**, it follows that $\Diamond\mathbf{P}\phi \rightarrow \Diamond\mathbf{PP}\phi$. It is a theorem of modal system **T** that $\mathbf{P}\phi \rightarrow \Diamond\mathbf{P}\phi$. By **ME4** and **ME5**, it follows that $\Diamond\mathbf{PP}\phi \rightarrow \Diamond\mathbf{P}\Diamond\mathbf{P}\phi$. The lemma follows from the introduction rule for conjunction and hypothetical syllogism. $\square$

---

[3]There is some discussion about the adequacy of **S5** as a logic of metaphysical possibility – see (Williamson 2008).

[4]Horsten added the Barcan formula to the formulas, but that formula is derivable in **S5** together with the classical quantifier principles (Hughes and Cresswell 1996, p. 247).

Corollary 1. *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, $\vdash_{\mathbf{MEA}} \left(\Box\phi \to \Box\Box\phi\right)^F$, i.e.*

$$\Diamond\mathbf{P}\phi^F \to \Diamond\mathbf{P}\Diamond\mathbf{P}\phi^F$$

The above lemmas are used in the proof of the following theorem:

Theorem 1 (Horsten). *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, if $\vdash_{\mathbf{EA}} \phi$, then $\vdash_{\mathbf{MEA}} \phi^F$.*

Proof. The proof is by induction on the length of proofs in **EA**. **MEA** contains classical first-order logic with identity and **PA** as well, so those cases are trivial. In the case of **E1** one can use **ME5**. In the case of **E2** one can use lemma 2. In the case of **E3** one can use lemma 1. Finally, in the case of **E4** one can use **ME4**. $\square$

Horsten did not prove the converse of the above theorem, but he conjectured it was correct. I will prove his conjecture in what follows. One needs the following definition:

Definition 2. $G : \mathcal{L}_{\mathbf{MEA}} \to \mathcal{L}_{\mathbf{EA}}$, *defined as follows:*

1. *if ϕ is an atomic formula, then $\phi^G = \phi$;*

2. *if $\phi = \neg\psi$, then $\phi^G = \neg\psi^G$;*

3. *if $\phi = \psi \wedge \theta$, then $\phi^G = \psi^G \wedge \theta^G$;*

4. *if $\phi = \forall x\psi$, then $\phi^G = \forall x\psi^G$;*

5. *if $\phi = \Diamond\psi$, $\phi^G = \psi^G$;*

6. *if $\phi = \mathbf{P}\psi$, $\phi^G = \mathbf{K}\psi^G$.*

One can then prove the following two lemmas:

Lemma 4. *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, $\left(\phi^F\right)^G = \phi$.*

Proof. By induction on the complexity of ϕ. Suppose that ϕ is atomic. Then $\phi^F = \phi^G = \phi$. Therefore, $\left(\phi^F\right)^G = \phi^G = \phi$. Suppose that $\phi = \mathbf{P}\psi$. Then $\left(\left(\mathbf{K}\psi\right)^F\right)^G = \left(\Diamond\mathbf{P}\psi^F\right)^G = \left(\mathbf{P}\psi^F\right)^G = \mathbf{K}\left(\psi^F\right)^G = \mathbf{K}\psi$. The last equality holds because of the induction hypothesis. I leave the other cases to the reader. $\square$

Lemma 5. *For all $\phi \in \mathcal{L}_{\mathbf{MEA}}$, if $\vdash_{\mathbf{MEA}} \phi$, then $\vdash_{\mathbf{EA}} \phi^G$.*

Proof. By induction on the complexity of proof in **MEA**. The modal axiom schemes and rules become the trivial axiom scheme $\psi^G \to \psi^G$ and the trivial rule $\psi^G \Rightarrow \psi^G$. The modal-epistemic axiom schemes and rules become $\mathbf{K}\psi^G \to \psi^G$, $\mathbf{K}\psi^G \to \mathbf{KK}\psi^G$, $\left(\mathbf{K}\psi^G \wedge \mathbf{K}\left(\psi^G \to \theta^G\right)\right) \to \theta^G$, and $\psi^G \Rightarrow \mathbf{K}\psi^G$, all of which are restricted versions of axioms schemes and rules of **EA**. I leave the other cases to the reader. $\square$

One can then prove the converse of Horsten's theorem:

Theorem 2. *For all $\phi \in \mathcal{L}_{\mathbf{EA}}$, if $\vdash_{\mathbf{MEA}} \phi^F$, then $\vdash_{\mathbf{EA}} \phi$.*

Proof. Suppose that $\phi \in \mathcal{L}_{\mathbf{EA}}$ and $\vdash_{\mathbf{MEA}} \phi^F$. Since ϕ^F belongs by definition to $\mathcal{L}_{\mathbf{MEA}}$, it follows by lemma 5 that $\vdash_{\mathbf{EA}} \left(\phi^F\right)^G$. So by lemma 4, $\vdash_{\mathbf{EA}} \phi$. $\square$

So there is a faithful and natural translation of the language of **EA** into the language of **MEA**.

Let us briefly consider whether progress has been made by analysing the complex notion of knowability into a modal and an epistemic component. The answer is affirmative. Lemma's 1 and 2 are philosophically interesting. The proof of lemma 1 involves **ME1**, which is the Platonic thesis that all arithmetical truths are necessary truths. However, this thesis is not strictly needed, since **ME1** is a theorem of **MEA** minus **ME1** in its axiomatic base.

Lemma 6. $\vdash_{\mathbf{MEA}^-} \quad \forall x \forall y \, (y < x \to \Box y < x)$, with **MEA**$^-$ *identical to* **MEA**, *except that* **ME1** *does not belong to its axiomatic base.*

Proof. The proof is by the induction principle.

First, one needs to prove that:

$$\forall y \, (y < \mathbf{0} \to \Box y < \mathbf{0})$$

The above is a trivial consequence of the theorem of **PA** that $\forall y \neg y < \mathbf{0}$.

Second, one needs to prove that:

$$\forall x \, (\forall y \, (y < x \to \Box y < x) \to \forall y \, (y < \mathbf{s}\,(x) \to \Box y < \mathbf{s}\,(x)))$$

Suppose that $y < \mathbf{s}\,(x)$. By an axiom of **PA**, it follows that $y < x \lor x = y$. It follows from the first disjunct and $\forall y \, (y < x \to \Box y < x)$ that $\Box y < x$. It follows from the second disjunct, the axiom schemes for identity and the rule of necessitation that $\Box x = y$ (Hughes and Cresswell 1996, p. 313). Therefore, it follows from the supposition that $\Box y < x \lor \Box x = y$. It is a consequence in the modal system **K** that $\Box \, (y < x \lor x = y)$ (Hughes and Cresswell 1996, p. 31). Given the proof principles of **K**, it follows from the latter and the mentioned axiom of **PA** that $\Box y < \mathbf{s}\,(x)$. $\quad \Box$

Theorem 3. *For all* $\phi \in \mathcal{L}_{\mathbf{PA}}$, $\vdash_{\mathbf{MEA}^-} \Diamond \phi \to \phi$, *with* **MEA**$^-$ *identical to* **MEA**, *except that* **ME1** *does not belong to its axiomatic base.*

Proof. The proof is by induction on the complexity of ϕ. Suppose that ϕ is an atomic formula. Then ϕ is of the form $t_i = t_j$ or of the form $t_i < t_j$. Suppose that $\Diamond t_i = t_j$. Then it follows that $t_i = t_j$. See (Hughes and Cresswell 1996, p. 314). Suppose that $\Diamond t_i < t_j$. Lemma 6 and reasoning analogous to the one in the previous case yields the conclusion that $t_i < t_j$.

Suppose that $\phi = \neg \psi$. The induction hypothesis is that $\vdash_T \Diamond \psi \to \psi$. This is equivalent to $\vdash_T \neg \psi \to \neg \Diamond \psi$. It follows that $\Diamond \neg \psi \to \Diamond \neg \Diamond \psi$. Assume that $\Diamond \neg \psi$. Then it follows that $\Diamond \neg \Diamond \psi$. It is an easy **S5**-consequence that $\neg \Diamond \psi$, whence it follows that $\neg \psi$. See (Hughes and Cresswell 1996, p. 35, 42, 58) for the relevant results from modal logic.

Suppose that $\phi = \psi \land \theta$. Assume that $\Diamond \phi$. It follows that $\Diamond \psi \land \Diamond \theta$ (Hughes and Cresswell 1996, p. 35). The conclusion follows by the induction hypothesis.

Suppose that $\phi = \forall x \psi$. Assume that $\Diamond \forall x \psi$. The induction hypothesis is that $\vdash_T \Diamond \psi \to \psi$. It follows by universal generalisation that $\forall x \Diamond \psi \to \forall x \psi$ and it follows from the consequent by universal instantiation that ψ. Since $\vdash_T \forall x \psi \to \psi$, it is also the case that $\vdash_T \Diamond \forall x \psi \to \Diamond \psi$. Finally, since x does not occur freely in $\Diamond \forall x \psi$, one can deduce $\vdash_T \Diamond \forall x \psi \to \forall x \Diamond \psi$. $\quad \Box$

The principle of the factivity of knowability is stronger than the principle of the factivity of knowledge, since the first implies the second given system **T**, whereas the second implies the first given system **S5**, and even then only in a restricted form: the factivity of knowability is restricted to what are, in effect (see lemma 1), necessary truths. Moreover, the principle of the factivity of knowledge is generally deemed correct,[5] whereas the *unrestricted* version of the principle of

---

[5]See (Williamson 2000), although some disagree (Spicer 2008).

the factivity of knowability is not. By way of informal counterexample, consider a theorem of **MEA** ϕ. By **ME4** it follows that $\Diamond\mathbf{P}\phi$. Next, by **ME3** it follows that $\Diamond\mathbf{PP}\phi$. Assume the unrestricted version of the principle of the factivity of knowability. Then it follows that $\mathbf{P}\phi$, for any theorem of **MEA**. In other words, the unrestricted version of the principle entails logical and mathematical omniscience. To be sure, this argument depends on other principles as well, notably **ME3** and **ME4**. The latter seems to be a correct principle, if one recalls that the relevant notion is (informal) proof, a kind of *a priori* knowledge. Given that **MEA** is recursively axiomatizable, one can have a proof of each of its theorems.

To conclude, not only can one in a natural and faithful way translate the language of **EA** to the language of **MEA**, but one also gains some philosophical insight by doing so, as witnessed by lemma's 1 and 2 and, to a lesser extent, lemma 3.

4 Horsten's Modal Epistemic-Arithmetic: Models

In the subsequent subsections I will discuss models for Modal-Epistemic Sentential Logic (**MESL**), Modal-Epistemic First-Order Logic (**MEFOL**) and, finally, Modal-Epistemic Arithmetic (**MEA**). Each level gives rise to problems that carry over to the higher levels (if any). By solving those problems at the lowest level possible one can avoid unnecessary complications.

4.1 Models for Modal-Epistemic Sentential Logic

For any language that contains the language of epistemic logic it is a challenge to give an interpretation that avoids making all the arithmetical truths knowable, as discussed in the previous section. There is an additional problem for *modal-epistemic* languages. Suppose one uses the standard Hintikka-Kripke models for **MESL**,[6] i.e. quadruples $\langle W, R_M, R_E, V \rangle$, with R_M a binary relation on W, called the 'modal accessibility relation', with R_E a reflexive binary relation on W, called the 'epistemic accessibility relation', and with V a function from the set of sentence letters and W to $\{1, 0\}$. The definition of truth of a formula at a world in a given model contains among others the following two clauses:

1. $\mathcal{M}, w \models \Diamond\psi$ if and only if $\mathcal{M}, w' \models \psi$ for some $w' \in W$ such that $wR_M w'$;

2. $\mathcal{M}, w \models \mathbf{P}\psi$ if and only if $\mathcal{M}, w' \models \psi$ for all $w' \in W$ such that $wR_E w'$.

Now that the epistemic and modal component of knowability have been disentangled, one is faced with the problem of *logical omniscience*, i.e. every logical truth is *known* at each possible world and every logical consequence of what is *known* at a world is also *known* at that world. (The *knowability* of logical truths and the logical consequences of *knowable* truths are part and parcel of the modal-epistemic logical fragment of **MEA**. More about that later.) Moreover, if arithmetical sentences get the same truth-value at each world, then all arithmetical truths are known at each possible world as well.

To avoid the problems of logical (and arithmetical) omniscience one could turn to awareness models, which are developed in (Fagin and Halpern 1988), or to impossible worlds models, which are developed in (Rantala 1982), or to syntactical models (Fagin et al. 1995). These types of models are defined as follows:

Definition 3. *An awareness model $\mathcal{M}$ for $\mathcal{L}_{\mathbf{MESL}}$ is a quintuple*

$$\langle W, R_M, R_E, V^a, A \rangle,$$

---

[6]See also the paper by Martin Fischer in this volume.

with W a non-empty set, R_M a binary relation on W, R_E a binary reflexive relation on W and V^a a function from the set of n-place predicate symbols and W to $\wp(D^n)$ and with A, called the 'awareness function', a function from W to $\wp(FORM(\mathcal{L}_{\textbf{MESL}}))$.

Definition 4. *If $\mathcal{M}$ is an awareness model, w a world and ϕ a well-formed formula, then $\mathcal{M}, w \models \phi$ is defined recursively as follows:*

1. *if ϕ is a sentence letter p, then $\mathcal{M}, w \models \phi$ if and only if $V^a(p) = 1$;*

2. *the clauses for $\neg$, $\wedge$ are expected;*

3. *if $\phi = \Diamond\psi$, then $\mathcal{M}, w \models \phi$ if and only if $\mathcal{M}, w' \models \psi$ for some $w' \in W$ such that $wR_M w'$;*

4. *if $\phi = \textbf{P}\psi$, then $\mathcal{M}, w \models \phi$ if and only if $\mathcal{M}, w' \models \psi$ for all $w' \in W$ such that $wR_E w'$ and $\psi \in A(w)$.*

The first conjunct of the truth-condition for formulas of the form $\textbf{P}\psi$ determines the formulas that are *implicitly* known at w, while the second conjunct of the same truth-condition acts as a sieve that filters out the formulas that are also *explicitly* known at w.

Definition 5. *An impossible worlds model for $\mathcal{L}_{\textbf{MESL}}$ is a quintuple*

$$\langle W, W^*, R_M, R_E, V^i \rangle,$$

with W a non-empty set, W^* a set, whose elements are called 'impossible worlds', R_M a binary relation on W, R_E a binary reflexive relation on $W \cup W^*$, V^i a function (i) from the set sentence letters and W to truth-values and (ii) from $FORM(\mathcal{L}_{\textbf{MESL}})$ and W^* to $\{1, 0\}$.

The clauses of the definition of truth of a formula at a *possible world* in a model are standard

Definition 6. *If $\mathcal{M}$ is an impossible worlds model, w a possible world and ϕ a well-formed formula, then $\mathcal{M}, w \models \phi$ is defined recursively as follows:*

1. *if ϕ is a sentence letter p, then $\mathcal{M}, w \models \phi$ if and only if $V^a(p) = 1$;*

2. *the clauses for $\neg$, $\wedge$ are as expected;*

3. *if $\phi = \Diamond\psi$, then $\mathcal{M}, w \models \phi$ if and only if $\mathcal{M}, w' \models \psi$ for some $w' \in W$ such that $wR_M w'$;*

4. *if $\phi = \textbf{P}\psi$, then $\mathcal{M}, w \models \phi$ if and only if $\mathcal{M}, w' \models \psi$ for all $w' \in W \cup W^*$ such that $wR_M w'$.*

Definition 7. *If $\mathcal{M}$ is an impossible worlds model, w^* an impossible world and ϕ a well-formed formula, then $\mathcal{M}, w^* \models \phi$ if and only if $V^i(\phi, w^*) = 1$.*

Validity is defined with respect to truth at *possible* worlds in a model.

Definition 8. *A syntactical model for $\mathcal{L}_{\textbf{MESL}}$ is a triple $\langle W, R_M, V^s \rangle$, with W a non-empty set, R_M a binary relation on W and with V^s a function from $FORM(\mathcal{L}(\textbf{MESL}))$ and W to $\{1, 0\}$ such that:*

1. *if $\phi = \neg\psi$, then $V^s(\phi, w) = 1$ if and only if $V^s(\psi, w) = 0$;*

2. *if $\phi = \psi \wedge \theta$, $V^s(\phi, w) = 1$ if and only if $V^s(\psi, w) = 1$ and $V^s(\theta, w) = 1$;*

3. *if $\phi = \Diamond\psi$, then $V^s(\phi, w)$ if and only if $V^s(\psi, w')$ for some $w' \in W$ such that $wR_M w'$;*

4. if $\phi = \mathbf{P}\psi$, then if $V^s(\phi, w) = 1$, then $V^s(\psi, w) = 1$.

Definition 9. *$\mathcal{M}, w \models \phi$ if and only if $V^s(\phi, w) = 1$.*

One can prove point-wise equivalence results for these types of models.[7] All three types of models provide solutions for the problem of logical omniscience, as the reader can verify for him- or herself. I will sketch the proof for a partial point-wise equivalence result. Consider any awareness model $\mathcal{M}^a = \langle W, R_M, R_E, V^a, A \rangle$, world w and formula ϕ. Then define a syntactical model $\mathcal{M}^s$ as follows: $\mathcal{M}^s = \langle W^a, R_M^a, V^s \rangle$, with V^s such that:

1. for any sentence letter p, $V^s(p, w) = 1$ if and only if $V^a(p, w) = 1$;

2. $V^s(\mathbf{P}\psi, w) = 1$ if and only if

$$\psi \in \{\phi \mid \phi \in A(w) \wedge \forall w'(wR_E w' \rightarrow M, w' \models \phi)\}.$$

One can prove by induction on the complexity of ϕ that $\mathcal{M}^a, w \models \phi$ if and only if $\mathcal{M}^s, w \models \phi$.

One remarkable feature of syntactical models is that they contain a truth-at-a-world-in-the-model function, which is normally defined separately. The reason the definition of truth-at-a-world-in-a-model is integrated in the definition of what a model is, is that one would otherwise have two circular definitions, since the truth-values of $\mathbf{P}\psi$-formulas are primitive (i.e., they are given directly by the model), but the truth of $\mathbf{P}\psi$ has to imply the truth of ψ, which may not be primitive but may be compositional in nature. This circularity is avoided in syntactical models as defined above, but it is not always easy to construct those models, since one essentially hard codes knowledge and thereby one has to respect the condition that what is known is true, which is not always easy to verify. This problem does not arise for awareness models or impossible worlds models, since the reflexivity of the R_E-relation guarantees that, if $\mathcal{M}, w \models \mathbf{P}\phi$, then $\mathcal{M}, w \models \phi$. If one wants to know whether there is a syntactical counter-model to an inference of ϕ from a set of premises Γ, then one can use the following strategy: first, construct an awareness model in which all of Γ are true and ϕ false; second, use the partial point-wise equivalence result provided earlier to prove that there is a syntactical model that is a counter-model as well. Of course, impossible worlds models are equally useful in this respect.

Awareness models, impossible worlds models and syntactical models can be used to solve the problem of logical omniscience. These kinds of models bring forth a new problem, however. This is the problem of *logical incompetence*. Nothing in these models guarantees that one can come to know logical truths or that one can come to know the logical consequences of knowable truths. In other words, any theory formulated in the language of modal-epistemic sentential logic that includes **ME4** and **ME5** is unsound with respect to these models. The modal-epistemic sentential logical fragment of **MEA** is such a theory. Let us call it **MESL**.

For soundness one has to restrict the class of awareness models to those in which R_E is a reflexive, transitive relation and in which R_M is an equivalence relation (since **MESL** contains **S5**) and in which it has the following properties:

1. for all ϕ, if $\mathcal{M} \models \phi$ for all $\mathcal{M}$ of the subclass, then for all $w \in W$, there is a $w' \in W$ such that $wR_M w'$ and $\mathcal{M}, w' \models \mathbf{P}\phi$;

---

[7]See (Fagin et al. 1995, ch. 9, exercise 9.45) and (Sillari 2008). The first gives the result as an exercise, the second gives proofs. Neither of them give results for awareness models and impossible worlds models for *modal-epistemic* languages; their results pertain to models for *epistemic* languages. The first bases awareness models and impossible worlds models on Kripke-style relational structures, as in the main text of this paper, whereas the second bases those models on Scott-Montague-style neighbourhood structures. The first included syntactical models in the exercise, whereas the second did not include them in his result. Nevertheless, it is easy to prove the point-wise equivalence result for awareness models, impossible worlds models and syntactical models for sentential modal-epistemic languages.

2. for all ϕ, ψ and for all $w_i, w_j, w_k \in W$, if $w_i R_M w_j$ and $\mathcal{M}, w_j \models \mathbf{P}(\phi \to \psi)$ and if $w_i R_M w_k$ and $\mathcal{M}, w_k \models \mathbf{P}\phi$, then there is a $w_l \in W$ such that $w_i R_M w_l$ and $\mathcal{M}, w_l \models \mathbf{P}\psi$.

The above conditions cannot be used to define the subclass of awareness models that validate **ME4** and **ME5**, since that would create the following circle: the definition of the subclass of models makes use of the notion of truth at a world in a model of that subclass, which in turn depends on the definition of the subclass of models. In others words, there is a circularity problem here. In the case of syntactical models, the situation is slightly different. The following conditions on R_M are sufficient to prove the soundness of **ME4** and **ME5**:[8]

1. for all ϕ, if $V^s(\phi, w) = 1$ for all V^s belonging to syntactical models and worlds w, then for all $w \in W$ there is a $w' \in W$ such that $w R_M w'$ and $V^s(\mathbf{P}\phi, w') = 1$;

2. for all ϕ, ψ, and $w_i, w_j, w_k \in W$, if $w_i R_M w_j$ and $V^s(\mathbf{P}\phi, w_i) = 1$ and if $w_i R_M w_k$ and $V^s(\mathbf{P}(\phi \to \psi), w_k) = 1$, then there is a $w_l \in W$ such that $w_i R_M w_l$ and $V^s(\mathbf{P}\psi, w_l) = 1$.

Since V^s belongs to the definition of a syntactical model, there is, strictly speaking, no circle here. However, a kind of holism has sneaked in: in order to know whether a syntactical model belongs to the subclass one has to know something about all the other members of the subclass (if any). So it will not always be a trivial affair to verify that a syntactical model satisfies the condition.

Fortunately, one can get around these problems. The solution is due to Horsten (1993, 1994).[9] Consider the following two conditions on R_M in syntactical models:

1s for all ϕ, if $\vdash_{\mathbf{MESL}} \phi$, then for all $\mathcal{M}$, for all $w \in W$, there is a $w' \in W$ such that $w R_M w'$ and $\mathcal{M}, w' \models \mathbf{P}\phi$;

2s for all ϕ, ψ and for all $w_i, w_j, w_k \in W$ if $w_i R w_j$ and $V^s(\mathbf{P}(\phi \to \psi), w_j) = 1$ and if $w_i R_M w_k$ and $V^s(\mathbf{P}\phi, w_k) = 1$, then there is a $w_l \in W$ such that $w_i R_M w_l$ and $V^s(\mathbf{P}\psi, w_l) = 1$.

Corresponding conditions for awareness models are the following:

1a for all ϕ, if $\vdash_{\mathbf{MESL}} \phi$, then for all $\mathcal{M}$, for all $w \in W$, there is a $w' \in W$ such that $w R_M w'$ and $\phi \in A(w')$;

2a for all ϕ, ψ and for all $w_i, w_j, w_k \in W$ if $w_i R_M w_j$ and $w_i R_M w_k$, $\phi \in A(w_j)$ and $\phi \to \psi \in A(w_k)$, then there is a $w_l \in W$ such that $w_i R_M w_l$, $w_j R_E w_l$, $w_k R_E w_l$, $\psi \in w_l$ and for all $w' \in W$ such that $w_l R_E w'$, it holds that $w_j R_E w'$ and $w_k R_E w'$.

And these are the corresponding conditions for impossible worlds models:

1i for all ϕ, if $\vdash_{\mathbf{MESL}} \phi$, then for all $\mathcal{M}$, for all $w \in W$, there is a $w' \in W$ such that $w R_M w'$ and for all $w'' \in W^*$, if $w' R_E w''$, then $V^i(\phi, w'') = 1$;

2i for all ϕ, ψ and for all $w_i, w_j, w_k \in W$ if $w_i R_M w_j$ and $w_i R_M w_k$, then there is a $w_l \in W$ such that $w_i R_M w_l$, $w_i R_M w_l$, $w_j R_E w_l$, $w_k R_E w_l$ and for all $w' \in W$ such that $w_l R_E w'$, $w_j R_E w'$ and $w_k R_E w'$ and, if $w' \in W^*$, then if $V^i(\phi, w') = 1$ and if $V^i(\phi \to \psi, w') = 1$, then $V^i(\psi, w') = 1$.

---

[8]Strictly speaking, *theories* are (un)sound under interpretations, but I will allow myself this lax terminology. Note that I do not want to use the notions of validity or validity-preservation at this point. More about that later.

[9]As we will see in subsection 4.4, Horsten used *one* condition rather than two, but in order to present this material in a step-by-step way, I will use two conditions.

Let us discuss these conditions one by one. Condition 1s on syntactical models guarantees that **ME4** is sound.[10] Suppose one has used **ME4** to derive $\Diamond\mathbf{P}\psi$ on line n in a proof. Then there is a line $m < n$ that contains ψ, which has been derived from the empty set of premises. It follows that, for any syntactical model $\mathcal{M}$ and world w there is a world w' such that $wR_M w'$ and $V^s(\mathbf{P}\psi) = 1$. Therefore $\models \Diamond\mathbf{P}\psi$.

Condition 1a ensures that **ME4** is sound relative to awareness models. If one proceeds from the same assumption as above, then for any $w \in W$, there is a $w' \in W$ such that $wR_M w'$ and $\psi \in A(w')$. Moreover, it follows from the induction hypothesis and $\vdash_{\mathbf{MESL}} \psi$ that $\mathcal{M} \models \psi$ for all awareness models $\mathcal{M}$ in the subclass. Therefore, $\mathcal{M}, w'' \models \psi$ for all w' such that $w'R_E w''$. Together these two facts imply that $\mathcal{M}, w' \models \mathbf{P}\psi$. Given that $wR_M w'$, it is also the case that $\mathcal{M}, w \models \Diamond\mathbf{P}\psi$. This is the case for any $\mathcal{M}$ (that satisfies condition 1a) and any w, so $\models \Diamond\mathbf{P}\psi$.

Finally, condition 1i enforces that **ME4** is sound. If there is a line $m < n$ in the proof such that $\vdash_{\mathbf{MESL}} \psi$, then for all impossible worlds models $\mathcal{M}$ and $w \in W$ there is a $w' \in W$ such that $wR_M w'$ and for all $w^* \in W^*$, if $w'R_E w^*$, then $V^i(\phi, w^*) = 1$. It follows by the induction hypothesis from $\vdash_{\mathbf{MESL}} \psi$ that $\models \psi$ and, therefore, that for all $w'' \in W$, if $w'R_E w''$, then $\mathcal{M}, w'' \models \psi$. Hence, $\mathcal{M}, w' \models \mathbf{P}\psi$, whence it follows that $\mathcal{M}, w \models \Diamond\mathbf{P}\psi$ and, since this holds for any $\mathcal{M}$ or w, $\models \mathbf{P}\psi$.

Condition 2s on syntactical models clearly does the job it is designed to do. After all, it is nothing but the result of hard coding what one wants. Condition 2a on awareness models also delivers what we need. For if $\mathcal{M}, w_i \models \Diamond\mathbf{P}\phi$ and $\mathcal{M}, w \models \Diamond\mathbf{P}(\phi \rightarrow \psi)$, then there is a world w_j and a world w_k with $w_i R_M w_j$ and $w_i R_M w_k$ such that $\mathcal{M}, w_j \models \mathbf{P}\phi$ and $\mathcal{M}, w_k \models \mathbf{P}(\phi \rightarrow \psi)$. Therefore, $\phi \in A(w_j)$ and $\phi \rightarrow \psi \in A(w_k)$. It then follows that there is a w_l such that $w_i R_M w_l$, $w_j R_E w_l$, $w_k R_E w_l$ and $\psi \in A(w_l)$. By the clause for $\mathbf{P}$, it follows that $\mathcal{M}, w_l \models \phi$ and $\mathcal{M}, w_l \models (\phi \rightarrow \psi)$. Consequently, $\mathcal{M}, w_l \models \psi$. Moreover, for all $w' \in W$ such that $w_l R_E w'$, $\mathcal{M}, w' \models \psi$, since $\mathcal{M}, w' \models \phi$ (because $w_j R_E w'$ and $\mathcal{M}, w_j \models \mathbf{P}\phi$) and $\mathcal{M}, w' \models \phi \rightarrow \psi$ (because $w_k R_E w'$ and $\mathcal{M}, w_k \models \mathbf{P}(\phi \rightarrow \psi)$). Therefore, $\mathcal{M}, w_l \models \mathbf{P}\psi$ and $\mathcal{M}, w_i \models \Diamond\mathbf{P}\psi$. So, **ME4** is valid on all models that satisfy condition 2a.

Finally, condition 2i on impossible worlds models makes **ME4** valid. Indeed, even if there is a $w' \in W^*$ such that $w_l R_E w'$, then

$$V^i(\phi, w') = V^i(\phi \rightarrow \psi, w') = 1,$$

since $w_j R_E w'$ and $w_k R_E w'$. So by condition 2i $V^i(\psi, w') = 1$.

Conditions 1s, 1a and 1i are not so easy to work with, since one needs to know in advance whether certain formulas are provable-in-**MESL** from a set of premises. Fortunately, there are sufficient, non-circular conditions that do the job as well. For instance, one can restrict the class of awareness models as follows: for every $w \in W$, there is a $w' \in W$ such that $wR_M w'$ and $A(w) = \mathcal{L}_{\mathbf{MESL}}$. This condition guarantees that **ME4** is truth-preserving at every world in a model that satisfies that condition. If $\mathcal{M} \models \phi$ for all $\mathcal{M}$ (in the subclass), then for any $w \in W$ there is a $w' \in W$ such that $wR_M w'$ and $A(w) = \mathcal{L}_{\mathbf{MESL}}$ and, in particular, $\phi \in A(w')$.

---

[10]The following is based on the proof in (Horsten 1993, ch. 4). In his proof Horsten proceeds from the assumption that ϕ, which is of the form $\Diamond\mathbf{P}\psi$, has been proved using **ME4** to the conclusion that ϕ is true at a possible world accessible from any given world. This is correct, but what follows is that $\mathcal{M}, w \models \Diamond\mathbf{P}\Diamond\mathbf{P}\psi$. Given **ME2** and a few modal principles, which are all assumed to be sound, it follows that $\mathcal{M}, w \models \Diamond\mathbf{P}\psi$, but Horsten does not reason that way. Probably he meant to say that a line m (with $m < n$) contains ψ, and that ψ has been derived from the empty set of premises. The argument given in the main text shows how to proceed from there. There is a conceptually important point to be made here. What is proved here is *soundness*, but not *validity-preservation*. From the assumption that $\Gamma \vdash \phi$ has been derived on the basis of **ME4**, it is inferred that ϕ is of the form $\Diamond\mathbf{P}\psi$ for some ψ and that on a line m (with $m < n$) one has ψ, which has been derived from the empty set of premises. It follows that $\Gamma \models \Diamond\mathbf{P}\phi$. What has not been proved is that, if $\models \psi$, then $\models \Diamond\mathbf{P}\psi$. To prove this, one needs to prove the completeness theorem first.

Moreover, if $\mathcal{M} \models \phi$ for all $\mathcal{M}$ (in the subclass), then $\mathcal{M}, w'' \models \phi$ for all w' such that $w' R_E w''$. Together these two facts imply that $\mathcal{M}, w' \models \mathbf{P}\phi$. Given that $w R_M w'$, it is also the case that $\mathcal{M}, w \models \Diamond \mathbf{P}\phi$. So there is a sufficient, non-circular condition. It is not a necessary condition, however. The effect of letting the awareness set of a possible world be the entire language is that the awareness set does no longer function as a sieve. But the holes of the sieve have been made wider than is strictly speaking necessary. Note that $\{\phi \mid \vdash_{\mathbf{MESL}} \phi\} \subset \mathcal{L}_{\mathbf{MESL}}$. Similarly, the class of impossible world models can be restricted as follows so as to make $\mathbf{ME4}$ truth-preserving: for every $w \in W$, there is a $w' \in W$ such that $w R_M w'$ and, for every w'' such that $w' R_E w''$, $w'' \in W$. The former is a sufficient, non-circular condition. (I leave the proof to the reader.) It is again not a necessary condition, however.

To conclude, there are subclasses of awareness models, impossible worlds models and syntactical models of modal-epistemic sentential languages that do not succumb to the problems of logical omniscience and logical incompetence.

4.2 Models for Modal-Epistemic First-Order Logic

So far I have only considered models for modal-epistemic *sentential* logic. I have not provided definitions of awareness models, impossible world models or syntactical models for *first-order* languages. In order not to complicate things from the start I will focus on first-order logic without the identity symbol, individual constants and function symbols.

Recall that the satisfaction clause for formulas ϕ of the form $\forall x \psi$ is the following: if ϕ is of the form $\forall x \psi$, then $\mathcal{M}, w, a \models \forall x \psi$ if and only if $\mathcal{M}, w, a' \models \psi$ for all $a' \sim_x a$.[11] Sillari (2008) studied awareness models and impossible worlds models for first-order languages. He used the satisfaction clause above. There is a problem that he failed to notice. Suppose that ϕ is of the form $\exists x \mathbf{P} \psi(x)$. Then $\mathcal{M}, w, a \models \phi$ only if for some x-variant a' of a, $\mathcal{M}, w', a' \models \psi(x)$ for all $w R w'$ and $\psi(x) \in A(w)$. There is a problem, however, namely that the Principle of Replacement ($\mathbf{PR}$) fails.[12] As a reminder, here are two definitions and a statement of the principle:

Definition 10. *For any variables x, y, z and variable assignment function a, $a[x \mapsto a(y)]$ is a function such that:*

1. *if $z \neq x$, then $a[x \mapsto a(y)](z) = a(z)$;*

2. *if $z = x$, then $a[x \mapsto a(y)](z) = a(y)$.*

Definition 11. *For any well-formed formula ϕ, variables x, y, $\phi(y/x)$ is identical to ϕ except that all free occurrences of x (if any) are uniformly replaced by free occurrences of y.*

Theorem 4 (Principle of Replacement). *For any well-formed formula ϕ, variables x and y and for any model $\mathcal{M}$ for a first-order language and variable assignment function a, $\mathcal{M}, a[x \mapsto a(y)] \models \phi$ if and only if $\mathcal{M}, a \models \phi(y/x)$. (Hughes and Cresswell 1996, p. 241, 242)*

Corollary 2. *For any model $\mathcal{M}$ for a first-order language and variable assignment function a, if $\mathcal{M}, a \models \forall x \phi$, then $\mathcal{M}, a \models \phi(y/x)$.*[13]

---

[11]Rantala (1982, p. 47) opted for substitutional quantification rather than objectual quantification. The problem of quantifying in as described in this article does not arise for impossible worlds models with substitutional quantification. Since objectual quantification is the standard form of quantification, I will explore what happens if one combines (semi-)syntactical models with objectual quantification.

[12]There is a variation on awareness models that avoids the problem described above. Awareness models of the kind defined above are sometimes referred to as *general awareness* models. Another variety are *local awareness* models. In (Sillari 2008) one can find a definition of local awareness models for first-order epistemic logic. $\mathbf{PR}$ holds in these models. The problem, however, is that it does not work for general awareness models. For it would lead to the unwelcome result that if one knows a truth about certain objects, then one knows all the truths about those objects.

[13]The crucial step of the proof is that $a[x \mapsto a(y)] \sim_x a$.

But there is nothing built into awareness models and impossible worlds that forces those models to make **PR** true. For instance, it can happen that in an awareness model $\mathcal{M}$ and at a world w for an assignment a that $\mathcal{M}, w, a[x \mapsto a(y)] \models \phi$ but $\mathcal{M}, w, a \not\models \phi(y/x)$. Let ϕ be $\mathbf{P}Px$, $W = \{w\}$, $R = \{\langle w, w \rangle\}$, $D = \{d\}$, $V^a(P, w) = d$, $A(w) = \{Px\}$ and let $a(x) = a(y) = d$. Then $\mathcal{M}, w, a[x \mapsto a(y)] \models \mathbf{P}Px$ and $\mathcal{M}, w, a \not\models \mathbf{P}Py$, since $Py \notin A(w)$. Note that it is also true that $\mathcal{M}, w, a \models \forall x \mathbf{P}Px$, so universal instantiation fails as well. Similarly, let ϕ be $\mathbf{P}Px$, $\mathcal{M}$ an impossible worlds model, with $W = \{w\}$, $W^* = \{w^*\}$, $R = \{\langle w, w \rangle, \langle w^*, w^* \rangle, \langle w, w^* \rangle\}$, $D = \{d\}$, $V^i(P, w) = d$, $V^i(Px, w^*) = 1$, $V^i(Px, w^*) = 1$ and let $a(x) = a(y) = d$. Given that $\mathcal{M}, w^*, a' \models \phi$ if and only if $V^i(\phi, w) = 1$, it follows that $\mathcal{M}, w, a[x \mapsto a(y)] \models \mathbf{P}Px$ but $\mathcal{M}, w, a \not\models \mathbf{P}Py$. Note that it is also true that $\mathcal{M}, w, a \models \forall x \mathbf{P}Px$, so universal instantiation fails as well. An example involving syntactical models will be given below.

There is an easy solution to the problem, but I think it is unsatisfactory. One might reason as follows: any theorem of **MEFOL** has to be known at some world w' which is modally accessible from a given world w; $\Box(\forall x \phi \rightarrow \phi(y/x))$ is a theorem of **MEFOL**; so it is knowable at w'; if a formula is known at w', then it is true at w'; consequently, $\forall x \phi \rightarrow \phi(y/x)$ is true at every world w'' that is modally accessible from w'; R_M is a symmetrical relation; therefore $\forall x \phi \rightarrow \phi(y/x)$ is true at w as well; the choice of w was arbitrary, so $\forall x \phi \rightarrow \phi(y/x)$ is true at any world. This solution is all right as far as it goes, but it does not help with awareness models, impossible world models or syntactical models in which R_M does not have the very particular set of properties the above argument depends on. It would be better to have a solution that works for those three types of (semi-)syntactical models in general.

A solution that works for impossible worlds models, awareness models and syntactical models alike consists making A, V^i and V^s relative to assignment functions a. Subsequently, one can hard code **PR** in those (semi-)syntactical elements. The following definitions spell this out.

Definition 12. *An awareness model $\mathcal{M}^a$ for $\mathcal{L}_{\mathbf{MEFOL}}$ is a tuple*

$$\langle W, R_M, R_E, D, V^a, S, A \rangle,$$

with W, R_M, R_E as in definition 3, with D a non-empty set, V^a a function from the set of n-place predicate symbols and W to $\wp(D^n)$, with S the set of functions from variables to elements of D, and A a function from W and S to $\wp(FORM(\mathcal{L}_{\mathbf{MEFOL}}))$ such that for any variables x, y, assignments a, worlds w and formulas ϕ, $\phi(y/x) \in A(w, a)$ if and only if $\phi \in A(w, a[x \mapsto a(y)])$.

Definition 13. *If $\mathcal{M}$ is an awareness model, w a world, a an assignment function and ϕ a well-formed formula, then $\mathcal{M}, w, a \models \phi$ is defined recursively as follows:*

1. *if ϕ is of the form $P(x_1, \ldots, x_n)$ (with P an n-place predicate symbol), then $\mathcal{M}, w, a \models \phi$ if and only if $\langle a(x_1), \ldots, a(x_n) \rangle \in V^a(R, w)$;*

2. *the clauses for $\neg$, $\wedge$ are as expected;*

3. *if $\phi = \Diamond \psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w', a \models \psi$ for some $w' \in W$ such that $w R_M w'$;*

4. *if $\phi = \mathbf{P}\psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w', a \models \psi$ for all $w' \in W$ such that $w R_E w'$ and $\psi \in A(w, a)$;*

5. *if $\phi = \forall x \psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w, a' \models \phi$ for all $a' \sim_x a$.*

Definition 14. *An impossible worlds model $\mathcal{M}^i$ for $\mathcal{L}_{\mathbf{MEFOL}}$ is a septuple $\langle W, W^*, R_M, R_E, D, S, V^i \rangle$, with W, W^*, R_M, R_E as in definition 5, with D a non-empty set, with S the set of functions*

from D to the set of variables, and with V^i a function from (i) the set of n-place predicate letters and W to $\wp(D^n)$ and (ii) from $FORM(\mathcal{L}_{\textbf{MEFOL}})$ and W^* and S to $\{1,0\}$ such that for any variables x, y, assignments a, worlds w and formulas ϕ, $V^i(\phi(y/x), w^*, a) = 1$ if and only if $V^i(\phi, w^*, a[x \mapsto a(y)]) = 1$.

Definition 15. *If $\mathcal{M}$ is an impossible worlds model, w a possible world, a an assignment and ϕ a well-formed formula, then $\mathcal{M}, w, a \models \phi$ is defined recursively as follows:*

1. *if ϕ is of the form $P(x_1, \ldots, x_n)$ (with P an n-place predicate symbol), then $\mathcal{M}, w, a \models \phi$ if and only if $\langle a(x_1), \ldots, a(x_n) \rangle \in V^i(R, w)$;*

2. *the clauses for $\neg$, $\wedge$ are as expected;*

3. *if $\phi = \Diamond\psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w', a \models \psi$ for some $w' \in W$ such that $w R_M w'$;*

4. *if $\phi = \mathbf{P}\psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w', a \models \psi$ for all $w' \in W \cup W^*$ such that $w R_M w'$;*

5. *if $\phi = \forall x \psi$, then $\mathcal{M}, w, a \models \phi$ if and only if $\mathcal{M}, w, a' \models \phi$ for all $a' \sim_x a$.*

Definition 16. *If $\mathcal{M}$ is an impossible worlds model, w^* an impossible world, a an assignment and ϕ a well-formed formula, then $\mathcal{M}, w^*, a \models \phi$ if and only if $V^i(\phi, w^*, a) = 1$.*

Definition 17. *A syntactical model for $\mathcal{L}_{\textbf{MEFOL}}$ is a quintuple*

$$\langle W, R_M, D, S, V^s \rangle,$$

with W a non-empty set, R_M a binary relation on W, S the set of functions from D to the set of variables, and with V^s a function (i) from the set of n-place predicate letters and W to $\wp(D^s)$ and (ii) from $FORM(\mathcal{L}_{\textbf{MEFOL}})$, W and S to $\{1,0\}$ such that:

1. *if ϕ is of the form $P(x_1, \ldots, x_n)$ (with P an n-place predicate symbol), then $V^s(\phi) = 1$ if and only if $\langle a(x_1), \ldots, a(x_n) \rangle \in V^s(P, w)$;*

2. *if $\phi = \neg\psi$, then $V^s(\phi, w, a) = 1$ if and only if $V^s(\psi, w, a) = 0$;*

3. *if $\phi = \psi \wedge \theta$, then $V^s(\phi, w, a) = 1$ if and only if $V^s(\psi, w, a) = 1$ and $V^s(\theta, w, a) = 1$;*

4. *if $\phi = \Diamond\psi$, then $V^s(\phi, w, a)$ if and only if $V^s(\psi, w', a)$ for some $w' \in W$ such that $w R w'$;*

5. *if $\phi = \mathbf{P}\psi$, then if $V^s(\phi, w, a) = 1$, then $V^s(\psi, w, a) = 1$*

6. *if $\phi = \forall x \psi$, then $V^s(\phi, w, a) = 1$ if and only if $V^s(\psi, w, a') = 1$ for all $a' \sim_x a$;*

7. *for any variables x, y, assignments a, worlds w and formulas ψ,*

$$V^s(\mathbf{P}\psi(y/x), w, a) = 1$$

if and only if

$$V^s(\mathbf{P}\psi, w, a[x \mapsto a(y)]) = 1.$$

Given the above definitions, it is easy to prove the following:

Theorem 5. *For any well-formed formula ϕ, variables x and y, and for any awareness model, impossible worlds model or syntactical model $\mathcal{M}$ for a first-order modal-epistemic language, world w and variable assignment function a, $\mathcal{M}, w, a \models \phi(y/x)$ if and only if $\mathcal{M}, w, a[x \mapsto a(y)] \models \phi$.*

The problem identified at the beginning of this subsection has been solved. Given Theorem 5 one can prove that the axiom scheme of Universal Instantiation is valid for any of the three types of models as defined in 12-17:

UI $\forall x\phi \to \phi\left(y/x\right)$

It can easily be checked that Universal Generalisation is validity-preserving.

UG If $\Gamma \vdash \phi\left(x\right)$ and x does not occur in Γ, then $\Gamma \vdash \forall x\phi\left(x\right)$

If one adds **UI** and **UG** to **MESL**, then one obtains Modal-Epistemic First-Order Logic or **MEFOL**. In order to prove the relevant soundness theorems, one has to impose conditions on R_M, as discussed in subsection 4.1.

In this subsection the problem of quantifying in that poses itself for awareness models, impossible worlds models and syntactical models for **MEFOL** has been discussed. **MEFOL** is Modal-Epistemic First-Order Logic without identity, individual constants or function symbols. In the next subsection I will briefly discuss models for modal-epistemic languages with identity and terms other than variables.

4.3 Identity and terms

The language $\mathcal{L}_{\mathbf{MEFOL}_=}$ is the extension of $\mathcal{L}_{\mathbf{MEFOL}}$ with the identity symbol, individual constants and function symbols. Now the question arises how to interpret those new symbols. For present purposes it will do to opt for the standard interpretation given in (Hughes and Cresswell 1996, p. 327-328). V is no longer only a function from the set of n-place predicate symbols and W to $\wp\left(D^n\right)$ but also from the set of individual constants to D and from the set of n-place function symbols to the set of functions from D^n to D. Next, one can define the denotation of a term t given a model $\mathcal{M}$ and an assignment a:

Definition 18. *Given a model $\mathcal{M}$, variable assignment function a and term t, the denotation of a term $den_{\mathcal{M},a}\left(t\right)$ is defined inductively as follows:*

1. *if t is an individual constant c, then $den_{\mathcal{M},a}\left(t\right) = V\left(c\right)$;*

2. *if t is an individual variable x, then $den_{\mathcal{M},a}\left(t\right) = a\left(x\right)$;*

3. *if t is of the form $f\left(t_1,\ldots,t_n\right)$ with f a n-place function symbol and $t_1,\ldots,t_n$ terms, then $den_{\mathcal{M},a}\left(t\right) = V\left(f\right)\left(den_{\mathcal{M},a}\left(t_1\right),\ldots,den_{\mathcal{M},a}\left(t_n\right)\right)$.*

Subsequently, one has to modify the definitions of satisfaction of a formula at a world in a model. The clause for atomic formulas now reads as follows: if ϕ is a formula of the form $P\left(t_1,\ldots,t_n\right)$, then $\mathcal{M},w,a \models \phi$ if and only if $\langle den_{\mathcal{M},a}\left(t_1\right),\ldots,den_{\mathcal{M},a}\left(t_n\right)\rangle \in V\left(P,w\right)$ and, if ϕ is of the form $t_1 = t_2$, then $\mathcal{M},w,a \models \phi$ if and only if $den_{\mathcal{M},a}\left(t_1\right) = den_{\mathcal{M},a}\left(t_2\right)$.

The first-order logic with identity and complex terms that is contained in **MEA**, call it **MEFOL$_=$**, is the extension of **MEFOL** with the following axiom schemes:

UIt $\forall x\phi \to \phi\left(t/x\right)$ (if t is substitutable for x in ϕ)

SI $x = x$

SUBS $x = y \to \left(\phi \to \phi'\right)$, with ϕ' identical to ϕ except perhaps that ϕ' has y free in zero or more places where ϕ has free x

The latter axiom scheme replaces **UI**, which is deducible from it. It is easy to verify that **SI** is valid, but **UIt** and **SUBS** are not.

One can make **UIt** and **SUBS** valid in awareness models by stipulating that, for any formula ϕ, variable x, term t, world w and assignment a, $\phi \in A(w, a[x \mapsto den_{\mathcal{M},a}(t)])$ if and only if $\phi' \in A(w, a)$ with ϕ' identical to ϕ except perhaps that ϕ has t in zero or more places where ϕ has x (if t is substitutable for x in ϕ). Note that $\phi(t/x)$ is a special case: *all* occurrences of x in ϕ have been replaced by t. Also, if $a(x) = a(y)$, then $a = a[x \mapsto a(y)]$. So, if $\phi \in A(w, a)$, then $\phi' \in A(w, a)$. Similarly, to make **UIt** and **SUBS** valid in impossible worlds models, one can stipulate that, for any formula ϕ, variable x, term t, assignment a and for any $w^* \in W$, $V^i(\phi, w^*, a[x \mapsto den_{\mathcal{M},a}(t)]) = 1$ if and only if $V^i(\phi', w^*, a) = 1$. Finally, in order to obtain the same goal for syntactical models, one can stipulate that for any formula ϕ, variable x, term t, world w and assignment a, $V^s(\phi, w, a[x \mapsto den_{\mathcal{M},a}(t)]) = 1$ if and only if $V^s(\phi', w, a) = 1$.

4.4   Models for Modal-Epistemic Arithmetic

So far the focus has been on models for the logical fragments of **MEA**. Let us have a look at models for **MEA** itself. First one needs to consider standard models of **PA**. Let $\mathcal{N} = \langle \omega, T \rangle$ be the standard interpretation of $\mathcal{L}_{\mathbf{PA}}$, with ω the standard set of natural numbers and with T giving the standard interpretation of the arithmetical vocabulary. Here is the definition of Horsten models for the language of **MEA**:

Definition 19 (Horsten Models for MEA). *A model of $\mathcal{L}_{\mathbf{MEA}}$ of is an ordered triple $\langle W, R_M, I \rangle$, with:*

1. *W a set of sets of sentences of $\mathcal{L}_{\mathbf{MEA}}$ such that:*

   (a) *for all $w \in W$ and for every sentence $\phi \in \mathcal{L}_{\mathbf{MEA}}$, if $\phi \in w$, then $\mathbf{P}\phi \in w$;*

   (b) *for all $w \in W$ and for every sentence $\phi \in \mathcal{L}_{\mathbf{MEA}}$, if $\phi \in w$, then $\mathcal{M}, w \models \phi$;*

2. *R_M a reflexive, symmetrical, transitive binary relation on W such that for any w_i, w_j, $w_k \in W$, if $w_i R_M w_j$ and $w_i R_M w_k$, then there is a $w_l \in W$ such that $w_i R_M w_l$ and $\{\phi | w_j \cup w_k \vdash_{\mathbf{MEA}} \phi\} \subseteq w_l$;*

3. *I is a function (i) that maps each constant c and function symbol f of $\mathcal{L}_{\mathbf{MEA}}$ and W to $T(c)$ and $T(f)$ respectively and (ii) from $FORM(\mathcal{L}_{\mathbf{MEA}})$ and W to $\{1, 0\}$ such that:*

   (a) *if $\phi = P(t_1, \ldots, t_n)$, then $I(w, \phi) = 1$ if and only if*

   $$\langle I(w, t_1), \ldots, I(w, t_1) \rangle \in I(w, P);$$

   (b) *if $\phi = \psi \rightarrow \theta$, then $I(w, \phi) = 1$ if and only if $I(w, \psi) = 0$ or $I(w, \theta) = 1$;*

   (c) *if $\phi = \neg\psi$, then $I(w, \phi) = 1$ if and only if $I(w, \psi) = 0$;*

   (d) *if $\phi = \exists x \psi(x)$; then $I(w, \phi) = 1$ if and only if*

   $$I_{[a/x]}(w, \psi(x)) = 1$$

   for some $a \in \omega$ (with $I_{[a/x]}$ identical to I with the possible exception that it assigns the number a to x);[14]

   (e) *if $\phi = \mathbf{P}\psi$, then $I(w, \phi) = 1$ if and only if $\psi \in w$;*

---

[14]It is curious that that Horsten talks of interpretations assigning numbers to variables. It might be that he means one of the two following things. First, it could be that he meant to relativise the interpretation functions

(f) if $\phi = \Diamond\psi$, then $I(w, \phi) = 1$ if and only if $I(w', \psi) = 1$ for some $w' \in W$ such that $wR_M w'$.

I will discuss the above definition in three stages. First, I will comment on what the above definition implies for **MESL**. Then I will proceed to what it implies for **MEFOL**. Finally, I will have a look at the arithmetical fragment.

One thing about the above definition that stands out is that I, which is an element of a Horsten model, has the same role as a truth-at-a-world-in-a-model function. In other words, I is very much like the V^s function of syntactical models. It follows from the clause for **P**-sentences and the conditions on W that, if $I(w, \mathbf{P}\psi) = 1$, then $I(w, \psi) = 1$. This is basically the clause for **P**-formulas in syntactical models. In order to avoid the circularity between the definitions of what a model is and what truth-at-a-world-in-a-model, one puts a truth function for complex formulas in the definition of a model itself.

The difference between syntactical models and Horsten models consists in the fact that in the latter the truth of $\mathbf{P}\psi$-sentences at a world w depends on whether $\psi \in w$, whereas there is no such dependence in syntactical models. However, the distinctive feature of Horsten models gives rise to a difficulty with respect to the canonical model proof of completeness. In the latter one takes the worlds w to be maximally consistent sets of formulas. **MESL**-theorems are elements of every maximally **MESL**-consistent set. If one went along Horsten's lines, then the result would be that for all models $\mathcal{M}$, worlds w and **MESL**-theorems ϕ, $\mathcal{M}, w \models \mathbf{P}\phi$, since $\mathcal{M}, w \models \mathbf{P}\phi$ if *and only if* $\phi \in w$. We would be back to the problem of logical omniscience. Horsten recognizes this difficulty. His solution consists in making a detour via another type of models, in which W is a set of ordered pairs of worlds (Horsten 1993, p. 97-98). I will not provide the details. An easier solution is to explicitly redefine his models as syntactical models, because the truth condition for formulas of the form $\mathbf{P}\phi$ is not a bi-conditional but a conditional: *if $V^*(\mathbf{P}\phi, w) = 1$, then $V^*(\phi, w) = 1$.* So, even if one has $V^*(\phi, w) = 1$ if and only if $\phi \in w$, one cannot deduce from the latter that $V^*(\mathbf{P}\phi, w) = 1$.

So not only are Horsten models a variation on syntactical models, they are best reformulated as syntactical models. They share with the latter the property of logical non-omniscience, since the set of formulas that are elements of a given world can be a proper subset of the language. So if ϕ is a theorem of **MEA** and if $w \in W$, one can stipulate that $\phi \notin w$, thereby guaranteeing that $\mathcal{M}, w \models \neg\mathbf{P}\phi$. Logical competence is guaranteed by the condition on R_M.

So much for Horsten models for the language of **MESL**. A noteworthy property of Horsten models for the language of **MEFOL** is that the worlds are sets of *sentences*. One might wonder why Horsten opted to restrict the extension of W this way. Why did he choose to restrict the elements of W to sets of closed formulas rather than formulas, whether they be open or closed? A clue may be found in the informal description of the models in (Horsten 1994, p. 87):

> Consider an arbitrary mathematician who is occupied solely with proving sentences of L_{MEA}. A possible world of a model can be thought of as a possible situation in which this mathematician has a particular (possibly empty, possibly infinite) collection of statements of L_{MEA} of which she has a demonstration.
>
> No limitations are imposed on the means which our mathematician has at her disposal for proving such sentences. She may in some possible situation have a higher-order demonstration of the ontic sentence of L_{MEA} which expresses the consistency of MEA. She may

---

to variable assignments, as is usually done and was done in subsection 4.2, but that he suppressed the variable assignments in the clauses in which he did not really need them. Second, it could be that he meant to define satisfaction along the lines of (Boolos, Burgess, and Jeffrey 2003, p. 117), namely an object satisfies an open formula 'if we considered the extended language $L \cup \{c\}$ obtained by adding a new constant c [...] to our given language L, and if among all the extensions of our given interpretation $\mathcal{M}$ to an interpretation of this extended language we considered the one $\mathcal{M}_m^c$ that assigns c the denotation m, then $F(c)$ would be true'. Neither of these options seems to be a good reading of what is in (Horsten 1994).

in some possible situation "intuitively" see that a statement which is independent of Peano arithmetic expresses an irreducible truth of arithmetic.

On the one hand, the elements of w include, among other things, formulas that can be proved in **PA**, extensions of **PA** or via higher mathematics. Whichever is the case, a formula is provable in an axiomatic theory if and only if its universal closure is provable and the universal closure of a formula is a sentence. Therefore one might want to restrict provable formulas to sentences. On the other hand, it is less clear whether one intuits only the truth of sentences or one can intuit that a property is true of an object as well.[15]

Whatever the reason that one has to define W as a set of sentences rather than formulas, the effect is that one has to deal again with the problem of quantifying in. Consider the **MEFOL**-theorem $\exists x \Diamond \mathbf{P}(P(x) \vee \neg P(x))$. The latter is true, i.e. $I(w, \exists x \Diamond \mathbf{P}(P(x) \vee \neg P(x))) = 1$ if and only if

$$I_{[a/x]}(w, \Diamond \mathbf{P}(P(x) \vee \neg P(x))) = 1$$

for some $a \in \omega$. The latter condition is fulfilled if and only if

$$I_{[a/x]}(w', \mathbf{P}(P(x) \vee \neg P(x))) = 1$$

for some $a \in \omega$ and for some $w' \in W$ such that wRw'. This, in turn, is true if and only if $P(x) \vee \neg P(x) \in w'$. But W was defined as a set of sets of *sentences*. Therefore $P(x) \vee \neg P(x) \notin w'$ and, hence, $I(w, \exists x \mathbf{P}(P(x) \vee \neg P(x))) = 0$. So, contrary to what Horsten thinks, the first-order logical fragment of **MEA** is not sound under his interpretation due to the fact that only *sentences* can be known. If, on the other hand, one would allow not only sentences or closed formulas to be knowable but *open formulas* as well,[16] then one would be confronted with the problem with respect to quantifying in that we encountered in the previous subsection, viz. the problem that **PR** and **UI** fail.

The situation is even worse than that. Recall that, if R is a symmetrical relation and is such that, if $\vdash_{\mathbf{MEA}} \phi$, then for every $w \in W$, there is a $w' \in W$ such that wRw' and $\phi \in w$, then $\forall x \mathbf{P}\psi(x) \rightarrow \mathbf{P}\psi(y/x)$ is true at every world. The antecedent conditions are satisfied in Horsten models, but at the same time the consequent cannot be true at any world. So the description of Horsten models is, in fact, incoherent. Fortunately, there is a solution to this problem, viz. the solution described in subsections 4.2 and 4.3. As it is best to view Horsten's models as syntactical models, one can import the solution for these models here.[17]

There is one more remark about Horsten models that is worth mentioning, namely that it is interesting to note that *all* of the Horsten models for **MEA** are standard models of arithmetic. The quantifiers range only over standard numbers and the arithmetical vocabulary is about the standard numbers. So Shapiro's worry does not apply. In section 2.2, it was explained that the challenge of finding models of the language of **EA** consists in finding models that

---

[15]In (Leitgeb 2009) the latter option is defended.

[16]In (Horsten 1993, p. 99, 100) the knowledge sets in fact *can* contain open formulas. When in the process of constructing a canonical model $\mathcal{M}'$ Horsten defines $\mathbf{P}^-(\Gamma)$ as 'the set consisting precisely of every formula A for which PA is in Γ', and goes on to stipulate that '[f]or each $\langle w_i, w_i' \rangle \in W' : w_i' = P^-(w_i)$'. There is no restriction to sentences anymore. Here w_i is a maximally **MEFOL**-consistent set of formulas of $\mathcal{L}_{\mathbf{MEFOL}}$ that is also ω-consistent. The satisfaction condition for formulas of the form $\mathbf{P}A$ is the following: $\mathcal{M}', \langle w_i, w_i' \rangle \models \mathbf{P}A$ if and only if $A \in w_i'$. This does not cohere very well with the definition of Horsten models nor with the definition variation on Horsten models in which the elements of W are ordered pairs.

[17]One might think that, if one understands Horsten models along the lines of (Boolos, Burgess, and Jeffrey 2003, p. 117), then there is no problem of quantifying in, since the satisfaction of open formulas is defined in terms of the truth of sentences. The problem is that these sentences belong to an extension of $\mathcal{L}_{\mathbf{MEA}}$ and, therefore, none of these sentences belongs to a world in the model, since worlds contain, by definition, only sentences of $\mathcal{L}_{\mathbf{MEA}}$.

allow the arithmetical vocabulary to be interpreted in the standard way without ending up with arithmetical omniscience. In other words, the problem was to show that *there are some* models of that kind. But now we have ended up with *only* models of that kind. This has undesirable consequences. Let us consider only the arithmetical fragment of the language and restrict the models accordingly. On the one hand, with Horsten models it is a consequence of $\{\phi(t) \mid t \in \mathcal{L}_{\mathbf{PA}} \wedge \phi \in \mathcal{L}_{\mathbf{PA}}\}$ that $\forall x \phi(x)$, since every $a \in \omega$ is the denotation of a Peano term. On the other hand, $\forall x \phi(x)$ is not deducible from $\{\phi(t) \mid t \in \mathcal{L}_{\mathbf{PA}} \wedge \phi \in \mathcal{L}_{\mathbf{PA}}\}$, since it is not deducible from any finite subset of it. To avoid this mismatch between logical consequence and deducibility, it is best to make the notion of Horsten models more liberal and allow models with non-standard numbers in their domain as well.

To conclude, Horsten's models for **MEA** are of the syntactical variety. They do not fall prey to the problem of logical omniscience nor to the problem of logical competence. It does face the problem of quantifying in, but we have seen how one that problem can be solved. Finally, Horsten's models show that one can assign a standard interpretation to the arithmetical vocabulary without trivializing the models, but it is wrong to allow only models based on the standard interpretation of arithmetic.

5   Conclusion

Even though this article focussed on Modal-Epistemic Arithmetic, I have touched on a number of issues that are of wider interest. The problem of logical omniscience is well-known and subject of much research. The topic of logical competence has, to the best of my knowledge, not yet gained equal standing. The two problems are intimately related: on the one hand, one does not want epistemic agents that *know* all logical truths and all logical consequences of truths they already know; on the other hand, one does want epistemic agents to *have the ability to know* logical truths and logical consequences of what they already know. The problem of quantifying in has been ignored by some, partly because much research focuses on sentential logic, and has been missed entirely by others. Starting from existing solutions to the problem of logical competence, I have tried to give solutions to the problems of logical competence and quantifying in.

Of course, this article was also about Modal-Epistemic Arithmetic. Horsten proved that Epistemic Arithmetic is interpretable in Modal-Epistemic Arithmetic: there is a translation of the language of Epistemic Arithmetic into the language of Modal-Epistemic Arithmetic such that the latter theory proves the translation of every theorem of the former theory. One contribution of this article consists in a proof that, if Modal-Epistemic Arithmetic proves the translation of a formula of the language of Epistemic Arithmetic, then that formula itself is a theorem of Epistemic Arithmetic. Horsten (1994, p. 287) argues that, by going from a purely epistemic framework to a modal-epistemic framework, one could find a justification of at least one principle, namely the factivity of knowability, because a restricted version of that principle follows from the simpler epistemic principle of the factivity of knowledge, modal principles and the Platonic thesis that all arithmetical truths are necessary. I showed that the Platonic thesis is not required as an axiom schema, since it can be proved in Modal-Epistemic Arithmetic without the thesis in its axiomatic base.

Moving from proofs to models, I argued that Horsten's models are a variation on syntactical models. Logical and arithmetical omniscience do not constitute problems for this kind of models. Moreover, Horsten defined his models in such a way that logical competence is ensured. But as a kind of syntactical models, Horsten's models inherit the problem of quantifying in. The important challenge Horsten took up was a challenge by Shapiro: find models for Epistemic Arithmetic that assign the standard interpretations to the arithmetical vocabulary but that are

not trivial. I have argued that Horsten may have taken this challenge a bit too seriously by allowing *only* standard models.[18]

References

Boolos, George S., John P. Burgess, and Richard C. Jeffrey (2003). *Computability and Logic*. Fourth. New York: Cambridge University Press.

Fagin, Ronald and Joseph Y. Halpern (1988). "Belief, Awareness, and Limited Reasoning". In: *Artificial Intelligence* 34, pp. 39–76.

Fagin, Ronald et al. (1995). *Reasoning About Knowledge*. MIT Press.

Flagg, Robert and Harvey Friedman (1986). "Epistemic and Intuitionistic Formal Systems". In: *Annals of Pure and Applied Logic* 51, pp. 387–392.

Goodman, Nicholas D. (1984). "Epistemic Arithmetic is a conservative extension of intuitionistic Arithmetic". In: *The Journal of Symbolic Logic* 49.1, pp. 192–203.

Horsten, Leon (1993). "Epistemic and Modal-Epistemic Arithmetic". PhD thesis. Katholieke Universiteit Leuven.

— (1994). "Modal-Epistemic variants of Shapiro's system of Epistemic Arithmetic". In: *Notre Dame Journal of Formal Logic* 35.2, pp. 284–291.

— (1998). "In Defense of Epistemic Arithmetic". In: *Synthese* 116.1, pp. 1–25.

Hughes, George Edward and Maxwell John Cresswell (1996). *A New Introduction to Modal Logic*. London: Routledge.

Leitgeb, Hannes (2009). "On Formal and Informal Provability". In: *New Waves in Philosophy of Mathematics*. Ed. by O. Linnebo and O. Bueno. Palgrave Macmillan.

Rantala, V. (1982). "Impossible worlds semantics and logical omniscience". In: *Acta Philosophica Fennica* 35, pp. 106–115.

Shapiro, Stewart (1985). "Epistemic and Intuitionistic Arithmetic". In: *Intensional Mathematics*. Ed. by Stewart Shapiro. Vol. 113. Studies in Logic and the Foundation of Mathematics. North-Holland, pp. 11–43.

Sillari, Giacomo (2008). "Quantified Logic of Awareness and Impossible Possible Worlds". In: *Review of Symbolic Logic* 1.4, pp. 514–529.

Smoryński, Craig A. (1991). "Review: Stewart Shapiro, Intensional Mathematics". In: *Journal of Symbolic Logic* 56.4, pp. 1496–4199.

Spicer, Finn (2008). "Are there any conceptual truths about knowledge?" In: *Proceedings of the Aristotelian Society* 108.1, pp. 43–60.

Williamson, Timothy (2000). *Knowledge and its limits*. New York: Oxford University Press.

— (2008). *The Philosophy of Philosophy*. Wiley-Blackwell.

---

[18]Thanks to Lorenz Demey, Martin Fischer and Leon Horsten for valuable feedback.