

# A modal logic of a truth definition for finite models

Marek Czarnecki\*

Institute of Philosophy, Warsaw University

Konrad Zdanowski\*

Institute of Mathematics, Polish Academy of Science

## Abstract

The property of being true in almost all finite, initial segments of the standard model for arithmetic is a  $\Sigma_2^0$ -complete property. Thus, it admits a kind of a weak truth definition. We define such an arithmetical predicate. Then we define its modal logic SL and prove a completeness theorem with respect to finite models semantics. The proof that SL is the modal logic of a weak truth definition for finite arithmetical models is based on an extension of SL by a fixpoint construction.

## 1 Introduction

We investigate in this work finite models being initial segments of the standard model for arithmetic. Such models proved to be useful in the context of descriptive complexity and investigations of computational aspects of semantics of languages. In descriptive complexity (see [Imm99] for a nice survey) one tries to capture the strength of various complexity classes by logics in which exactly problems from a given complexity class can be expressed. In order to do so, one usually enriches the structure of finite models with some

---

\*Both authors were partially supported by grant N N201 382234 of the Polish Ministry of Science and Higher Education.

arithmetical predicates. Even relatively weak first order logic corresponds to a natural complexity class called uniform  $AC_0$ .

Here we concentrate on some general properties of first order arithmetic in finite models. Let  $FM(\mathcal{N}) = \{\mathcal{N}_i: i \in \omega\}$  be the family of finite initial segments of the standard model  $\mathcal{N} = (\omega, \leq, R_+, R_\times)$ , where  $R_+$  and  $R_\times$  are relational versions of addition and multiplication, respectively.

It is known that the property of a sentence  $\varphi$  being true in almost all such models,  $FM(\mathcal{N}) \models_{sl} \varphi$ , is  $\Sigma_2^0$ -complete in arithmetical hierarchy (see [MZ05]). The upper bound can be clearly seen just from the arithmetical definition of the property:

$$\exists M \forall n \geq M \mathcal{N}_n \models \varphi.$$

On the other hand, if for a given formula  $\varphi(x)$  we consider a set  $X_\varphi$  defined as

$$X_\varphi = \{n \in \omega: FM(\mathcal{N}) \models_{sl} \varphi(n)\}$$

then for first order arithmetical  $\varphi$  we get exactly the sets in  $\Sigma_2^0$  (see [MZ05]). It follows that there will be an arithmetical formula  $\psi$  such that we have an equivalence for each sentence  $\varphi$ ,

$$FM(\mathcal{N}) \models_{sl} \varphi \text{ if and only if } FM(\mathcal{N}) \models_{sl} \psi(\ulcorner \varphi \urcorner),$$

where  $\ulcorner \varphi \urcorner$  is a Gödel number of  $\varphi$ . In fact we can even require that the formula  $\psi$  commutes with propositional connectives so its behavior resembles, to a certain degree, a behavior of a truth definition. We will call such a formula a weak  $FM$ -truth definition. This is just an approximation of a truth definition in the usual sense since. By undefinability of truth in  $FM(\mathcal{N})$  (see [Mos01]), we cannot have more, that is

$$FM(\mathcal{N}) \models_{sl} (\varphi \equiv \psi(\ulcorner \varphi \urcorner)).$$

We will investigate the properties of such weak  $FM$ -truth definitions which are expressible in a propositional modal logic. That is we consider a modal logic  $L_{Tr}$  defined as follows. Let  $tr$  be a translation function from the set of propositional variables into arithmetical sentences. Then, we extend  $tr$  by requiring that it commutes with propositional connectives and that it translates the necessity operator  $\Box$  as  $\psi$ . So, e.g., the inductive clause for  $\Box$  looks like

$$tr(\Box\varphi) = \psi(\ulcorner tr(\varphi) \urcorner).$$

Then we define  $L_{\text{Tr}}$  as the set of modal formulas  $\varphi$  such that for any function  $\text{tr}$ , the arithmetical formula  $\text{tr}(\varphi)$  is true in almost all finite models  $FM(\mathcal{N})$ . In our main theorem we characterize  $L_{\text{Tr}}$  as an extension of a basic modal logic  $K$  by an axioms scheme:

$$\Box(\neg\varphi) \equiv \neg\Box\varphi.$$

Thus, the modal properties of a weak  $FM$ -truth definition may be contrasted with that of provability predicate which corresponds to, so called, Gödel–Löb modal logic  $GL$  which is an extension of  $K$  by a scheme  $\Box\varphi \Rightarrow \Box\Box\varphi$  (see e.g. [Fra] or [Boo] for two different and interesting surveys) . Indeed, this two logics are incomparable. This fact may be somehow expected since our weak truth predicate approximates certain semantics while  $GL$  is a logic which captures the properties of demonstrability, a very different concept. The proof of our main result shows also that, unlike in the case of  $GL$ , we cannot consistently extend  $SL$  by any axiom scheme.

The method of proving our main result is by extending the modal logic  $K$  by the scheme above, getting a logic we call  $SL$ , and then by extending it with an additional fix point construction, obtaining a logic we call  $SL^*$ . Then, we prove that  $SL^*$  is: a sublogic of  $L_{\text{Tr}}$ , conservative over  $SL$  in the vocabulary of  $SL$  and, at the same time, a maximal logic which cannot be consistently extended to a stronger logic. All together it gives the theorem.

Our result can be seen from two points of view. From the first one it is a contribution to the study of logical properties of finite models. It investigates what fragments of finite model semantics can be expressed within finite models and what are finite models properties of these concepts. From the second point point of view it is a study in modal logics defining a certain natural modal logic and showing that it has a maximality property by considering its extension by a fix point construction. We believe that the two points of view are interesting on their own but we also believe that their interplay gives us an additional scientific value.

## 2 Basic definitions

In this section we fix the notation and introduce the main concepts. We assume some background in the finite models theory and recursion theory. Any introductory textbooks, e.g. [EFT94] and [Sho93] should be sufficient.

We consider the first order arithmetic in a relational language. Moreover, with each predicate we connect its intended meaning e.g.  $R_+$  with the relation of addition,  $R_\times$  with the relation of multiplication, etc. Therefore, we will not distinguish between the signature of the language (vocabulary) and relations in a model. The latter will be always either well known arithmetical relations or its finite models versions.

For a formula  $\varphi$  by  $\ulcorner \varphi \urcorner$  we denote its Gödel number and by  $|\varphi|$  its length. We use a shorthand  $\exists^=1$  for the quantifier “there exists exactly one element”.

An arithmetical formula  $\varphi$  is bounded or  $\Delta_0$  if all quantifiers occurring in  $\varphi$  are of the form  $(Qx \leq t)$ , where  $Q \in \{\exists, \forall\}$  and  $t$  is a term. By  $\Sigma_n$  we denote the set of formulas which begin with a block of existential quantifiers and have  $n - 1$  alternations followed by a bounded formula. Similarly,  $\varphi$  is in  $\Pi_n$  if it begins with a block of universal quantifiers and has  $n - 1$  alternations followed by a bounded formula. Let us observe that  $\Sigma_0$  as well as  $\Pi_0$  formulas are exactly bounded formulas.

For a formula  $\varphi(x_1, \dots, x_n)$  and  $k \in \omega$  by  $\varphi^{\leq k}$  we denote a formula which arises from  $\varphi$  by bounding all quantifiers in  $\varphi$  by  $\bar{k}$  and adding additionally conjunct  $\bigwedge_{1 \leq i \leq n} x_i \leq \bar{k}$ . Of course, if  $\varphi$  is a sentence then the added conjunct is empty.

We use symbols  $\Sigma_n$  and  $\Pi_n$  to denote also the classes of relations in arithmetical hierarchy. (It will be always clear from the context in which sense the notation is used.) A relation  $R \subseteq \mathcal{N}^r$  is  $\Sigma_n$  ( $\Pi_n$ ) if it definable by a  $\Sigma_n^0$  ( $\Pi_n^0$ ) formula in the standard model for arithmetic. A relation  $R$  is  $\Delta_n^0$  if it is  $\Sigma_n^0$  and  $\Pi_n^0$ .

By  $\bar{n}$  we denote the numeral  $n$ . Since we consider relational arithmetical vocabulary we need to express numerals in a somehow complicated manner, by  $\varphi(\bar{n})$  we abbreviate a formula

$$\exists x_0 \dots \exists x_n (x_0 = 0 \wedge \bigwedge_{1 \leq i \leq n} S(x_{i-1}, x_i) \wedge \varphi(x_n)).$$

Equivalently, we can write the above formula as

$$\forall x_0 \dots \forall x_n (x_0 = 0 \wedge \bigwedge_{1 \leq i \leq n} S(x_{i-1}, x_i) \Rightarrow \varphi(x_n)).$$

Thus, one can eliminate all terms in a formula without increasing its quantifier complexity. Whenever we speak of, e.g., a term denoting an integer or that a formula  $\varphi(x)$  is true about the term  $t$  is should be understand by the usual translations of terms into the corresponding relational language.

The quantifier rank of a formula  $\varphi$ ,  $rk(\varphi)$ , is defined in a usual way, i.e.  $rk(\varphi) = 0$  if  $\varphi$  is atomic formula,  $rk(\neg\varphi) = rk(\varphi)$ ,  $rk(\varphi \wedge \psi) = \max\{rk(\varphi), rk(\psi)\}$ , and  $rk(\exists x\varphi) = 1 + rk(\varphi)$ . Similarly for modal formulas we define a modal rank as  $mr(p) = 0$  for every variable  $p$ ,  $mr(\neg\varphi) = mr(\varphi)$ ,  $mr(\varphi \circ \psi) = \max\{mr(\varphi), mr(\psi)\}$ , for  $\circ \in \{\wedge, \vee, \Rightarrow\}$ , and  $mr(\Box\varphi) = 1 + mr(\varphi)$ .

## 2.1 Finite models for arithmetic

Let  $\mathcal{A}$  be a model having as a universe the set of natural numbers, i.e.  $\mathcal{A} = (\omega, R_1, \dots, R_s)$ , where  $R_1, \dots, R_s$  are relations on  $\omega$ . We will consider finite initial fragments of these models. Namely, for  $n \in \omega$ , by  $\mathcal{A}_n$  we denote the following structure

$$\mathcal{A}_n = (\{0, \dots, n\}, R_1^n, \dots, R_s^n, n),$$

where  $R_i^n$  is the restriction of  $R_i$  to the set  $\{0, \dots, n\}$ . We will denote the family  $\{\mathcal{A}_n\}_{n \in \omega}$  by  $FM(\mathcal{A})$ . The signature of  $\mathcal{A}_n$  is an extension of the signature of  $\mathcal{A}$  by one constant. This constant will be denoted by  $MAX$ . We introduce it just for convenience since in all models we consider the maximal element is definable by a fixed arithmetical formula.

Let  $\varphi(x_1, \dots, x_p)$  be a formula and  $b_1, \dots, b_p \in \omega$ . We say that  $\varphi$  is satisfied by  $b_1, \dots, b_p$  in all finite models of  $FM(\mathcal{A})$  ( $FM(\mathcal{A}) \models \varphi[b_1, \dots, b_p]$ ) if for all  $n \geq \max(b_1, \dots, b_p)$   $\mathcal{A}_n \models \varphi[b_1, \dots, b_p]$ . We say that  $\varphi$  is satisfied by  $b_1, \dots, b_p$  in all sufficiently large finite models of  $FM(\mathcal{A})$ , what is denoted by  $FM(\mathcal{A}) \models_{sl} \varphi[b_1, \dots, b_p]$ , if there is  $k \in \omega$  such that for all  $n \geq k$   $\mathcal{A}_n \models \varphi[b_1, \dots, b_p]$ . In what follows we work with one fixed infinite model for arithmetic,  $\mathcal{N} = (\omega, S, R_+, R_\times, \leq)$ . Therefore, instead of writing  $FM(\mathcal{N}) \models_{sl} \varphi[\bar{b}]$  we may just write  $\models_{sl} \varphi[\bar{b}]$ .

Let  $\mathcal{F}$  be a set of sentences of first order logic. By  $Th_{\mathcal{F}}(\mathcal{A})$ , where  $\mathcal{A}$  is a model, we denote the set of all sentences from  $\mathcal{F}$  true in  $\mathcal{A}$ . For a class of models  $\mathcal{K}$ , by  $Th_{\mathcal{F}}(\mathcal{K})$  we denote the set of sentences from  $\mathcal{F}$  true in all models from  $\mathcal{K}$ , that is  $Th_{\mathcal{F}}(\mathcal{K}) = \bigcap_{\mathcal{A} \in \mathcal{K}} Th_{\mathcal{F}}(\mathcal{A})$ . By  $sl_{\mathcal{F}}(FM)$  we denote the set of sentences from  $\mathcal{F}$  true in all sufficiently large finite models of  $FM(\mathcal{A})$ . So, we have

$$Th_{\mathcal{F}}(FM(\mathcal{A})) = \{\varphi \in \mathcal{F} : \forall n \in \omega \mathcal{A}_n \models \varphi\},$$

$$sl_{\mathcal{F}}(FM(\mathcal{A})) = \{\varphi \in \mathcal{F} : \exists k \forall n \geq k \mathcal{A}_n \models \varphi\}.$$

When  $\mathcal{F}$  is the set of all sentences of a given signature we will omit the subscript  $\mathcal{F}$ .

## 2.2 Truth definitions and diagonal lemma

The idea how to represent the relations on  $\omega$  in finite models was formulated in the article of Marcin Mostowski [Mos01]. He defined there the notion of FM-representability. Relation  $R \subseteq \omega^r$  is FM-representable in  $FM(\mathcal{N})$  if and only if there exists a formula  $\varphi(x_1, \dots, x_r)$  such that for all  $a_1, \dots, a_r \in \omega$ ,

$$(a_1, \dots, a_r) \in R \text{ if and only if } FM(\mathcal{N}) \models_{sl} \varphi[a_1, \dots, a_r]$$

and

$$(a_1, \dots, a_r) \notin R \text{ if and only if } FM(\mathcal{N}) \models_{sl} \neg\varphi[a_1, \dots, a_r].$$

For the theory of finite models of arithmetic with addition and multiplication we have the following theorem.

**Theorem 1 ([Mos01])** *Relation  $R \subseteq \omega^r$  is FM-representable in  $FM(\mathcal{N})$  if and only if  $R$  is in  $\Delta_2^0$ .*

One can characterize the relations in  $\Delta_2^0$  as those which are decidable by a Turing machine with a recursively enumerable oracle (see e.g. [Sho93]).

In his investigations of representability in finite models Mostowski was especially interested in the notion of truth. The FM-representability theorem gives us the way of expressing various notions in finite arithmetical models, e.g., the following relations are recursive, thus FM-representable:

- $Name(x, y)$ , abbreviating that  $y$  is the Gödel number of a canonical term  $\bar{x}$  naming  $x$ ,
- $Subst(x, y, z)$ , abbreviating that  $z$  is the Gödel number of the formula obtained by substituting the term with Gödel number  $y$  for a variable  $v_0$  into formula with Gödel number  $x$ .

A possibility to represent various syntactical notions opened a way of representing the whole variety of syntactic and semantical concepts. This resulted with forging the following notion of FM-truth definition: formula  $\psi(x)$  is an FM-truth definition when for every arithmetical sentence  $\varphi$ ,

$$FM(\mathcal{N}) \models_{sl} \varphi \equiv \psi(\ulcorner \varphi \urcorner).$$

In [Mos01] Marcin Mostowski stated and proved the following FM-version of the diagonal lemma.

**Lemma 2 (FM–version of the diagonal lemma)** *For every arithmetical formula  $\psi(x)$  with free variable  $x$ , there exists a sentence  $\varphi$  such that:*

$$FM(\mathcal{N}) \models_{sl} \varphi \equiv \psi(\ulcorner \varphi \urcorner).$$

As a consequence of Lemma 2 Marcin Mostowski proved also an FM–version of the undefinability of truth theorem.

**Theorem 3 (FM–version of the undefinability of truth theorem)** *There is no arithmetical formula which is an FM–truth definition.*

Therefore the notion of FM–truth for arithmetic of addition and multiplication can not be expressed within the finite model framework similarly as it can not be expressed in the standard model. However in the next section we show that we can construct a formula which – in a way – approximates an FM–truth definition i.e. it has several properties we should expect from a truth definition.

### 2.3 A truth definition for almost all finite models

It is a folklore that there is no  $\Delta_0$  truth definition for  $\Delta_0$  formulas and that there exists a  $\Sigma_1$  truth definition  $\text{Tr}_{\Delta_0}(x)$  for  $\Delta_0$ . Thus  $\text{Tr}_{\Delta_0}(x)$  is of the form  $\exists y \vartheta(x, y)$ , where  $\vartheta \in \Delta_0$ . Despite the fact that we cannot get rid of this leading existential quantifier in  $\text{Tr}_{\Delta_0}(x)$  we know how to estimate it (see [HP93]). Let  $\varphi(a_1, \dots, a_n)$  be a  $\Delta_0$  formula, where  $a_i$  are fixed parameters, and let  $h(x, y)$  be a function defined as

$$h(\ulcorner \varphi \urcorner, a) = (a + 2)^{2^{|\varphi|}}.$$

Then the following holds (see [HP93])

$$\exists y \vartheta(\ulcorner \varphi \urcorner, y) \text{ if and only if } \exists y \leq h(\ulcorner \varphi \urcorner, \max \{a_1, \dots, a_n\}) \vartheta(\ulcorner \varphi \urcorner, y).$$

Now let us introduce a  $\Delta_0$  truth definition  $\alpha(x, k, z)$  with two additional parameters  $k$  and  $z$ . The first variable  $x$  is supposed to be a Gödel number of some sentence  $\varphi$ ,  $k$  is supposed to be a bound for quantification in  $\varphi$  and values of free variables and finally  $z$  is a bound for existential quantifier for  $\text{Tr}_{\Delta_0}$ . Thus we define:

$$\alpha(x, k, z) = \exists y \leq z \vartheta(x^{\leq k}, y)^1.$$

---

<sup>1</sup>By  $x^{\leq k}$  we denote the Gödel number of  $\varphi^{\leq k}$ , when  $x$  is the Gödel number of  $\varphi$  – here  $k$  may be considered a variable and its value can be fixed externally.

Let  $f(x, k) = \max\{h(\ulcorner y \leq l \urcorner, k) : y \leq x, l \leq k\}$ . Function  $f$  defined this way is monotone in both arguments and the following are equivalent:

1.  $\mathcal{N}_k \models \varphi$ ,
2.  $\mathcal{N}_{f(\ulcorner \varphi \urcorner, k)} \models \alpha(\ulcorner \varphi \urcorner, k, \max)$ ,
3.  $\forall n \geq f(\ulcorner \varphi \urcorner, k) \mathcal{N}_n \models \alpha(\ulcorner \varphi \urcorner, k, \max)$ .

The part (3) in the equivalence above is essential for our purpose as we investigate asymptotic properties of formulas. Now it is sufficient to take  $F(x) = f(x, x)$  and define the relation  $k = F^{-1}(x) \equiv_{df} x \in [F(k), F(k+1))$ . The notation is slightly abused, yet justified by close correspondence between the relation  $k = F^{-1}(x)$  and the coimage of  $F$ . Observe that  $\forall k \in \omega \exists x k = F^{-1}(x)$  and  $\forall x \geq F(0) \exists^{=1} k (x \in [F(k), F(k+1)))$ . The formula  $\alpha(x, y, z)$  can be written in  $\Delta_0$  form. Similarly, the relation  $z = F^{-1}(x)$  is  $\Delta_0$ -definable. It follows that there is one arithmetical formula which defines in a given finite model  $\mathcal{N}_m$  the restriction of  $z = F^{-1}(x)$  to the universe of  $\mathcal{N}_m$ . Finally the formula  $\text{Tr}_{sl}(x)$  approximating FM-truth is defined as:

$$\text{Tr}_{sl}(x) = \exists k = F^{-1}(\text{MAX})\alpha(x, k, \text{MAX}).$$

By our discussion the above formula is  $\Delta_0$ .

The following theorem explains what we mean saying that  $\text{Tr}_{sl}(x)$  approximates FM-truth.

**Definition 4** *We say that an arithmetical formula  $\tau(x)$  is a weak FM-truth definition if for all quantifier free formulas  $\psi$*

1.  $FM(\mathcal{N}) \models_{sl} (\psi \equiv \tau(\ulcorner \psi \urcorner))$

and for all sentences  $\varphi, \psi$ ,

2.  $FM(\mathcal{N}) \models_{sl} \varphi$  if and only if  $FM \models_{sl} \tau(\ulcorner \varphi \urcorner)$ ,
3.  $FM(\mathcal{N}) \models_{sl} \tau(\ulcorner \neg \varphi \urcorner) \equiv \neg \tau(\ulcorner \varphi \urcorner)$ ,
4.  $FM(\mathcal{N}) \models_{sl} \tau(\ulcorner \varphi \circ \psi \urcorner) \equiv (\tau(\ulcorner \varphi \urcorner) \circ \tau(\ulcorner \psi \urcorner))$ , for  $\circ \in \{\wedge, \vee, \Rightarrow\}$ .

We do not expect a weak  $FM$ -truth definition to commute with quantifiers, because this would give us a regular  $FM$ -truth definition which do not exists. However, we could have a weak  $FM$ -truth definition  $\text{Tr}_{\text{sl}}(x)$  such that it satisfies one of the implications, either

$$FM(\mathcal{N}) \models_{\text{sl}} \text{Tr}_{\text{sl}}(\ulcorner \forall x \varphi \urcorner) \Rightarrow \forall a \text{Tr}_{\text{sl}}(\ulcorner \varphi(a) \urcorner)$$

or

$$FM(\mathcal{N}) \models_{\text{sl}} \forall a \text{Tr}_{\text{sl}}(\ulcorner \varphi(a) \urcorner) \Rightarrow \text{Tr}_{\text{sl}}(\ulcorner \forall x \varphi \urcorner).$$

**Theorem 5**  $\text{Tr}_{\text{sl}}$  is a weak  $FM$ -truth definition.

**Proof.**

1. First observe that for every quantifier-free formula we can eliminate  $MAX$  with respect to finite models semantics. Thus we can assume that  $\varphi(x_1, \dots, x_n)$  is a quantifier-free formula without occurrences of  $MAX$ . For a fixed valuation  $v$ , let  $M = \max\{\max_{1 \leq i \leq n} v(x_i), \ulcorner \varphi \urcorner\}$ . Then in each model  $M_m$  for  $m \geq F(M)$ ,  $N_m \models \psi \equiv \text{Tr}_{\text{sl}}(\ulcorner \psi \urcorner)[v]$  for every atomic formula  $\psi$  and thus also for their boolean combinations with Gödel numbers not exceeding  $M$  – thus also for  $\varphi$ .
2. For all sufficiently large  $k$  and all  $n \in [F(k), F(k+1))$ ,  $\mathcal{N}_k \models \varphi$  is equivalent to  $\mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \varphi \urcorner)$ . Therefore

$$\exists k \forall n > k \mathcal{N}_n \models \varphi \text{ if and only if } \exists k \forall n > k \mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \varphi \urcorner).$$

3. For all  $n \geq F(\ulcorner \neg \varphi \urcorner) > F(\ulcorner \varphi \urcorner)$  and  $k = F^{-1}(n)$ ,  $\mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \neg \varphi \urcorner)$  is equivalent to  $\mathcal{N}_k \models \neg \varphi$ . The latter is equivalent to  $\mathcal{N}_n \not\models \varphi$  and, finally, to  $\mathcal{N}_n \not\models \text{Tr}_{\text{sl}}(\ulcorner \varphi \urcorner)$ .
4. We treat only the case of  $\circ = \wedge$ . For all  $n \geq F(\ulcorner \varphi \wedge \psi \urcorner) > \max\{F(\ulcorner \varphi \urcorner), F(\ulcorner \psi \urcorner)\}$  and  $k = F^{-1}(x)$ ,  $\mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \varphi \wedge \psi \urcorner)$  is equivalent to  $\mathcal{N}_k \models \varphi$  and  $\mathcal{N}_k \models \psi$  which means  $\mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \varphi \urcorner)$  and  $\mathcal{N}_n \models \text{Tr}_{\text{sl}}(\ulcorner \psi \urcorner)$  respectively.

We showed that  $\text{Tr}_{\text{sl}}$  is a weak  $FM$ -truth predicate.  $\square$

## 2.4 Kripke semantics for modal logics

Formulas of modal logics are generated by the following grammar:

$$\varphi \longrightarrow \perp \mid p \mid \varphi \Rightarrow \varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg\varphi \mid \Box\varphi,$$

where  $p$  is an arbitrary element of the set PROP of propositional variables. We introduce the following abbreviation:  $\Diamond\varphi = \neg\Box\neg\varphi$ .

By Kripke frame we call a pair  $(W, R)$ , where  $W \neq \emptyset$  is a set of – so called – possible worlds and  $R \subseteq W^2$  is an accessibility relation. For a fixed Kripke frame  $F = (W, R)$  a valuation in  $F$  is a function  $V: (W \times \text{PROP}) \rightarrow \{0, 1\}$ . We call a triple  $(W, R, V)$  a Kripke model when  $(W, R)$  is a Kripke frame and  $V$  is a valuation on it. If it is clear from the context that we refer to a Kripke model we will call it a model for short. The semantics for modal logics is defined inductively on the construction of formula. For every Kripke model  $M = (W, R, V)$ :

- $\llbracket \perp \rrbracket_M = \emptyset$ ,
- $\llbracket p \rrbracket_M = \{w \in W : V(w, p) = 1\}$ ,
- $\llbracket \varphi \wedge \psi \rrbracket_M = \llbracket \varphi \rrbracket_M \cap \llbracket \psi \rrbracket_M$ ,
- $\llbracket \neg\varphi \rrbracket_M = W - \llbracket \varphi \rrbracket_M$ ,
- $\llbracket \Box\varphi \rrbracket_M = \{w \in W : \forall v \in W (wRv \Rightarrow v \in \llbracket \varphi \rrbracket_M)\}$ .

We say that a formula  $\varphi$  is true at a world  $w \in W$  of a model  $M = (W, R, V)$  when  $w \in \llbracket \varphi \rrbracket_M$ . We denote this fact also by  $M, w \models \varphi$  or if we consider a Kripke frame  $F = (W, R)$  with a valuation  $V$  we may denote it by  $F, w \models \varphi[V]$ . However if the context of a model or a frame and valuation is unambiguous we denote the fact that  $\varphi$  is true in  $w$  plainly by  $w \models \varphi$ .

## 3 Modal logics SL, SL\* and L<sub>Tr</sub>

In this section we present a modal logic SL for which we prove a Solovay style completeness theorem for our arithmetical truth definition. The definition of SL mimics the properties of Tr<sub>sl</sub> introduced in Subsection 2.3. The intention is to interpret  $\Box$  as Tr<sub>sl</sub> and to add appropriate axioms to the system i.e. translate properties of Tr<sub>sl</sub> to the modal language.

**Definition 6 (K)** *The modal logic K is an extension of classical propositional logic by axioms:*

$$\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box(\varphi) \Rightarrow \Box(\psi)),$$

where  $\varphi$  and  $\psi$  are arbitrary modal formulas and by the necessitation rule, i.e., if we proved  $\varphi$  we can write in a proof also  $\Box(\varphi)$ , for any formula  $\varphi$ .

Any modal logic which is closed on modus ponens, necessitation and substitution is called normal. In this work we consider only normal logics with one natural proviso. We allow substitution only for propositional variables of the basic language, not for propositional constants added to logic  $SL^*$ .

**Definition 7 (SL)** *The modal logic SL is an extension of modal logic K with the following axioms, for each formula  $\varphi$ ,*

$$\Box(\neg\varphi) \equiv \neg\Box(\varphi).$$

Let us observe, that adding the above axiom to  $K$  is enough to make  $\Box$  commute with all propositional connectives. One can prove this by using laws of classical propositional calculus. It follows also that for every formula  $\varphi$ ,  $SL \vdash \Box\varphi \equiv \Diamond\varphi$ . For a fixed  $\varphi$  the following are equivalent in  $SL$ :  $\Box\varphi$ ,  $\Box\neg\neg\varphi$  and, by (1),  $\neg\Box\neg\varphi$  which is, by definition,  $\Diamond\varphi$ .

Now, we define a modal logic  $SL^*$  which we prove to be a conservative extension of  $SL$ . Firstly, we extend the language of  $SL$ .

**Definition 8** *A variable  $p$  is guarded in a modal formula  $\varphi(p)$  if each occurrence of  $p$  is within the scope of a modal operator. For each  $p$  and formula  $\varphi(p)$  such that  $p$  is guarded in  $\varphi$  we add a new propositional constant  $q_{\langle\varphi,p\rangle}$ . The logic  $SL^*$  is an extension of  $SL$  by the axioms*

$$q_{\langle\varphi,p\rangle} \equiv \varphi(q_{\langle\varphi,p\rangle}/p),$$

where  $q_{\langle\varphi,p\rangle}$  is a new constant and  $\varphi(q_{\langle\varphi,p\rangle}/p)$  is a result of replacing in  $\varphi(p)$  each occurrence of  $p$  by  $q_{\langle\varphi,p\rangle}$ .

The last logic we are going to define is the modal logic of the truth predicate for  $\models_{sl}$ .

**Definition 9** *Let  $L_{Tr}$  be the set of all formulas of modal logic  $\varphi(p_1, \dots, p_n)$  such that for any translation  $tr$  it holds  $\models_{sl} \varphi^{tr}$ .*

By definition  $L_{\text{Tr}}$  is the modal logic of the truth predicate for  $\models_{sl}$ . It is easy to see that  $L_{\text{Tr}}$  is a consistent normal modal logic.

**Fact 10** *Both SL and  $SL^*$  are contained in  $L_{\text{Tr}}$ .*

**Proof.** The fact that  $SL \subseteq L_{\text{Tr}}$  is obvious since the axioms of SL mimics the properties of  $\text{Tr}_{sl}(x)$  in  $\models_{sl}$ . To show that  $SL^* \subseteq L_{\text{Tr}}$  let us assume that  $SL^* \vdash \varphi$ , where  $\varphi$  is in the language of SL. We show that  $\varphi \in L_{\text{Tr}}$ . Let  $\varphi_1, \dots, \varphi_n$  be a proof of  $\varphi$  in  $SL^*$  and let  $q_{\langle \psi_1, p_1 \rangle}, \dots, q_{\langle \psi_k, p_k \rangle}$  be all additional variables of  $SL^*$  which are used in the proof. Let  $\text{tr}$  be an arbitrary translation of variables occurring in the proof besides  $q_{\langle \psi_1, p_1 \rangle}, \dots, q_{\langle \psi_k, p_k \rangle}$ . Now, for each  $i \leq k$  let  $\gamma_i$  be an arithmetical sentence such that

$$\models_{sl} (\gamma_i \equiv \psi_i^{\text{tr}}(\gamma_i/p_i)).$$

Such  $\gamma_i$ 's exist by the fix point lemma for  $\models_{sl}$  (Lemma 2). Let us extend  $\text{tr}$  by putting  $\text{tr}(q_{\langle \psi_i, p_i \rangle}) = \gamma_i$ . It is easy to see that under this extension the translation of each additional axiom of  $SL^*$  is true in almost all models of  $FM(\mathcal{N})$ . Then, it can be proved by induction on  $i \leq n$ , that  $\models_{sl} \varphi_i^{\text{tr}}$ . Since the translation  $\text{tr}$  was arbitrary it follows that  $\varphi_n \in L_{\text{Tr}}$ .  $\square$

**Corollary 11** *The logic  $SL^*$  is consistent.*

### 3.1 Completeness theorem for SL

We start this section with the following remark on SL's models. Let us consider formula  $\Box \perp$  – this formula is true exactly in those worlds of a given Kripke frame from which there are no accessible worlds – let us call them final. On the other hand  $\Diamond \perp$  is trivially equivalent to  $\perp$ . Since  $SL \vdash \Box \perp \equiv \Diamond \perp$  there are no final points in SL's models.

We call a Kripke frame a *line* if it is of the form  $(\{0, \dots, n\}, S_n \cup \{(n, n)\})$  or  $(\omega, S)$ , where  $S$  is the successor relation and  $S_n$  is its restriction to the set  $\{0, \dots, n\}$ . Thus, a frame is a line if it is a finite initial segment of the successor relation with a loop added at the top or if it is the standard model for arithmetic of the successor relation. We will denote the  $n$ -th finite line with universe  $\{0, \dots, n\}$  by  $L_n$ . The infinite line will be denoted by  $L_\omega$ . Now, we prove a completeness theorem for SL.

**Definition 12** We say that a valuation  $V$  in Kripke frame  $F$  is admissible for a logic  $L$  if each axiom of  $L$  is true in  $F$  under  $V$ , that is  $F \models \varphi[V]$ . We say that a family of Kripke frames  $\mathcal{F}$  is sound and complete with respect to the modal logic  $L$  if the following are equivalent:

1.  $L \vdash \varphi$ ,
2. for any  $F \in \mathcal{F}$  and for any valuation  $V$  in  $F$  admissible for  $L$ ,  $F \models \varphi[V]$ .

If  $\mathcal{F}$  is a singleton  $\{F\}$  we say that  $F$  is sound and complete with respect to  $L$ .

The main tool for proving a completeness theorem is the following lemma.

**Definition 13** Let  $\mathcal{F}$  be a set of modal formulas. By  $\Box^{-1}\mathcal{F}$  we denote the following set

$$\Box^{-1}\mathcal{F} = \{\varphi : \Box\varphi \in \mathcal{F}\}.$$

**Lemma 14** Let  $L$  be a modal logic containing SL and let  $\mathcal{F}$  be a maximal consistent in  $L$  set of formulas. Then,  $\Box^{-1}\mathcal{F}$  is a maximal consistent in  $L$  set of formulas.

**Proof.** Let  $L$  and  $\mathcal{F}$  satisfy the assumptions of the lemma. One can easily see that  $\Box^{-1}\mathcal{F}$  is consistent in  $L$ . Otherwise, there would be a formula  $\varphi \in \Box^{-1}\mathcal{F}$  such that  $L \vdash \neg\varphi$ . Then,  $L \vdash \Box\neg\varphi$  and  $L \vdash \neg\Box\varphi$ . But  $\Box\varphi \in \mathcal{F}$  thus  $\mathcal{F}$  would be inconsistent. Therefore, for the sake of contradiction, let us assume that  $\Box^{-1}\mathcal{F}$  is not maximal with respect to consistency. Thus, there is a formula  $\psi \notin \Box^{-1}\mathcal{F}$  such that  $\Box^{-1}\mathcal{F} \cup \{\psi\}$  is still consistent. We have that  $\Box\psi \notin \mathcal{F}$ . By maximality of  $\mathcal{F}$ , the set  $\mathcal{F} \cup \{\Box\psi\}$  is inconsistent in  $L$ . Thus, there exists a formula  $\varphi \in \mathcal{F}$  such that  $L \vdash \varphi \Rightarrow \neg\Box\psi$ . It follows that  $\neg\Box\psi \in \mathcal{F}$ . Since  $\Box\neg\psi$  is equivalent in SL to this last formula we have that  $\Box\neg\psi \in \mathcal{F}$  and  $\neg\psi \in \Box^{-1}\mathcal{F}$ . We obtain a contradiction since  $\Box^{-1}\mathcal{F} \cup \{\psi\}$  was assumed to be consistent.  $\square$

**Theorem 15** 1. The infinite line  $L_\omega$  is sound and complete with respect to logics SL and SL\*.

2. The family of all finite lines is sound and complete with respect to SL.

**Proof.** Since the soundness part may be easily verified we concentrate on completeness only. Moreover, it is enough to prove the theorem only for  $L_\omega$ . Indeed, if  $L_\omega, 0 \not\models \varphi[V]$  and  $n$  is a model depth of  $\varphi$ , then  $L_n, 0 \not\models \varphi[V_n]$ , where  $V_n$  is a restriction of  $V$  to the worlds  $0, \dots, n$ .

To prove completeness we will show the following implication: for each sentence  $\varphi$ , if  $SL^* \not\models \varphi$ , then there exists a valuation  $V$  admissible for  $SL^*$  such that  $L_\omega, 0 \not\models \varphi[V]$ . As the reader will see, the same construction will work also for  $SL$ . The only difference is that in the case of  $SL$  we may stop the construction after  $n$  steps, where  $n$  is the modal depth of  $\varphi$ .

Thus, let us assume that  $SL^* \not\models \varphi$ . Let  $\mathcal{F}_0$  be a set of formulas which is maximal consistent in  $L$  and contains  $\neg\varphi$ . We construct a sequence of sets  $\mathcal{F}_i$ , for  $i \in \omega$  such that

$$\mathcal{F}_{i+1} = \Box^{-1}\mathcal{F}_i.$$

By Lemma 14 each of  $\mathcal{F}_i$  is maximal consistent in  $L$ . Now, we construct a valuation  $V: \omega \times \text{PROP} \rightarrow \{0, 1\}$  in  $L_\omega$  as follows

$$V(i, p) = 1 \text{ if and only if } p \in \mathcal{F}_i.$$

A straightforward proof by induction on the complexity of a formula shows that for all formulas  $\psi$  and for all  $i$ ,

$$L_\omega, i \models \psi[V] \text{ if and only if } \psi \in \mathcal{F}_i.$$

Moreover, since each  $\mathcal{F}_i$  is maximal consistent in  $SL^*$ , it satisfies the following equivalence, for each  $q_{\langle\psi, p\rangle}$ ,

$$q_{\langle\psi, p\rangle} \in \mathcal{F}_i \text{ if and only if } \psi(q_{\langle\psi, p\rangle}) \in \mathcal{F}_i.$$

Then, we get from the above that  $V$  is admissible for  $SL^*$ . It follows that we constructed a model which falsifies  $\varphi$ .  $\square$

Let us observe, that in the proof above in the case of  $SL^*$  we could not restrict ourselves just to a finite number of  $\mathcal{F}_i$ . To the contrary, it can be shown that all sets  $\mathcal{F}_i$  are different. This is caused by fix point axioms of  $SL^*$ . This is why  $SL^*$  has no finite models while in the case of  $SL$  it would be enough to consider only sets  $\mathcal{F}_0, \dots, \mathcal{F}_n$ , where  $n$  is a modal depth of  $\varphi$  independent from  $SL$ .

## 4 Main theorem

In this section we prove the main result of the article which characterizes the modal logic of the truth predicate in finite models.

**Definition 16** For a formula  $\varphi$ , we define  $(\neg)^0\psi$  as  $\neg\psi$  and  $(\neg)^1\psi$  as  $\psi$ . For a finite set of propositional variables  $P$  and a function  $\varepsilon: \{0, \dots, k\} \times P \rightarrow \{0, 1\}$  we define  $\Phi_\varepsilon$  as a formula

$$\bigwedge_{0 \leq i \leq k} \bigwedge_{p \in P} \Box^i (\neg)^{\varepsilon(i,p)} p.$$

Any function  $\varepsilon$  of the above form we will be called a valuation on variables  $P$  and  $k$  consecutive worlds.

For  $P$  and  $\varepsilon$  as above, if the formula  $\Phi_\varepsilon$  is true at a given world  $a$  of a model  $L$  of SL then it determines completely the values of propositions in  $P$  at  $a$  and worlds which can be accessed from  $a$  in  $k$  steps. Indeed, if  $L, a \models \Phi_\varepsilon[V]$  then, for any  $p \in P$  and any  $0 \leq i \leq k$ ,

$$L, a + i \models p[V] \quad \text{if and only if} \quad \varepsilon(i, p) = 1.$$

It follows that  $\Phi_\varepsilon$  determines also truth values of formulas with modal depth not greater than  $k$  over variables from  $P$ .

**Lemma 17** Let  $k \in \omega$ , let  $P$  be a finite set of variables and let  $\varepsilon: \{0, \dots, k\} \times P \rightarrow \{0, 1\}$ . If  $\varphi$  is a formula of modal depth not greater than  $k$  with all variables from  $P$  then  $\Phi_\varepsilon$  decides  $\varphi$  that is

$$SL \vdash \Phi_\varepsilon \Rightarrow \varphi \quad \text{or} \quad SL \vdash \Phi_\varepsilon \Rightarrow \neg\varphi.$$

Now, we show that any valuation on propositional variables is consistent with  $SL^*$ .

**Lemma 18** For every  $n > 0$  and  $\varepsilon: \{0, \dots, n-1\} \rightarrow \{0, 1\}$  there is a formula  $\psi$  such that

$$SL^* \vdash \bigvee_{0 \leq r < n} \Box^r \left( \bigwedge_{0 \leq i < n} \Box^i (\neg)^{\varepsilon(i)} \psi \right).$$

Moreover,  $SL^* \vdash (\psi \equiv \Box^n \psi)$ .

**Proof.** For  $n = 1$  if  $\varepsilon(0) = 0$  we put  $\psi = \perp$  and  $\psi = \top$  otherwise. For  $n > 1$  and a fixed propositional constant  $p$  consider the following formula

$$\varphi_n = \left( \neg \Box^{n-1} p \wedge \bigwedge_{0 < i < n} (\Box^i p \equiv \Box^{n-1} p) \right).$$

Since  $\varphi_n$  is guarded in  $p$  there is a propositional constant  $q_{\langle \varphi_n, p \rangle}$  such that

$$\text{SL}^* \vdash q_{\langle \varphi_n, p \rangle} \equiv \left( \neg \Box^{n-1} q_{\langle \varphi_n, p \rangle} \wedge \bigwedge_{0 < i < n} (\Box^i q_{\langle \varphi_n, p \rangle} \equiv \Box^{n-1} q_{\langle \varphi_n, p \rangle}) \right).$$

We need the following properties of  $q_{\langle \varphi_n, p \rangle}$ : for each valuation  $V$  admissible for  $\text{SL}^*$ , and for each  $a$ :

1. there exists  $i \leq n - 1$  such that  $L_\omega, a + i \models q_{\langle \varphi_n, p \rangle}[V]$ ,
2. if  $L_\omega, a \models q_{\langle \varphi_n, p \rangle}[V]$  then for each  $1 \leq i \leq n - 1$ ,  $L_\omega, a + i \not\models q_{\langle \varphi_n, p \rangle}[V]$ .

It follows that  $q_{\langle \varphi_n, p \rangle}$  is true exactly in every  $n$ -th world of  $L_\omega$ . For the first point it suffices to observe that if for all  $1 \leq i \leq n - 1$ ,  $L_\omega, a + i \not\models q_{\langle \varphi_n, p \rangle}[V]$  then, by the definition of  $q_{\langle \varphi_n, p \rangle}$ , it has to be true at the world  $a$ . The second point also follows easily from the definition of  $q_{\langle \varphi_n, p \rangle}$ .

Now, for a fixed  $\varepsilon: \{0, \dots, n - 1\} \rightarrow \{0, 1\}$  we define  $\psi$  as follows:

$$\psi = \bigvee_{\substack{0 \leq j \leq n-1 \\ \varepsilon(j)=1}} \Box^{n-j} q_{\langle \varphi_n, p \rangle},$$

where the empty disjunction is  $\perp$ . Since it is easy to see that the lemma holds for  $\psi = \perp$  we may assume that there is  $j \in \{0, \dots, n - 1\}$  such that  $\varepsilon(j) = 1$ . We will show that such defined  $\psi$  has needed properties. By completeness of  $L_\omega$  it is enough to show that for each admissible valuation  $V$  there exists  $r \leq n - 1$  such that

$$L_\omega, r \models \bigwedge_{0 \leq i \leq n-1} \Box^i (\neg)^{\varepsilon(i)} \psi[V].$$

Thus, let  $V$  be an admissible valuation and let  $r \leq n - 1$  be smallest world such that

$$L_\omega, r \models q_{\langle \varphi_n, p \rangle}[V].$$

By the two mentioned above properties of  $q_{\langle\varphi_n,p\rangle}$ , for each  $a$ ,

$$L_\omega, a \models q_{\langle\varphi_n,p\rangle}[V] \text{ if and only if } a = r + kn, \text{ for some } k \in \omega.$$

It follows that for each  $0 \leq i \leq n - 1$

$$\begin{aligned} L_\omega, r \models \Box^i \psi[V] &\iff L_\omega, r \models \Box^i \left( \bigvee_{\substack{0 \leq j \leq n-1 \\ \varepsilon(j)=1}} \Box^{n-i} q_{\langle\varphi_n,p\rangle} \right) \\ &\iff L_\omega, r \models \bigvee_{\substack{0 \leq j \leq n-1 \\ \varepsilon(j)=1}} \Box^i \Box^{n-j} q_{\langle\varphi_n,p\rangle} \\ &\iff \varepsilon(i) = 1. \end{aligned}$$

Similarly, for each  $0 \leq i \leq n - 1$ ,

$$L_\omega, r \models \Box^i \neg \psi[V] \iff \varepsilon(i) = 0.$$

Thus,

$$L_\omega, r \models \bigwedge_{0 \leq i < n} \Box^i (\neg)^{\varepsilon(i)} \psi[V]$$

and since the valuation  $V$  is arbitrary and  $r \leq n - 1$ , we get by completeness of  $L_\omega$  that

$$\text{SL}^* \vdash \bigvee_{0 \leq r < n} \Box^r \left( \bigwedge_{0 \leq i < n} \Box^i (\neg)^{\varepsilon(i)} \psi \right).$$

This completes the proof of the first part of the lemma. To prove the ‘‘Moreover’’ part one needs to observe that the only propositional constant used in  $\psi$  is  $q_{\langle\varphi_n,p\rangle}$  and there are no variables in  $\psi$ . But for this constant we have  $\text{SL} \vdash (q_{\langle\varphi_n,p\rangle} \equiv \Box^n q_{\langle\varphi_n,p\rangle})$  and this property easily transfer over all formulas which uses only this one constant.  $\square$

**Lemma 19** *Let  $k \in \omega$ , let  $P$  be a finite set of variables and let  $\varepsilon: \{0, \dots, k\} \times P \longrightarrow \{0, 1\}$ . If  $L$  is a consistent modal logic such that  $\text{SL}^* \subseteq L$  then  $\Phi_\varepsilon$  is consistent with  $L$ .*

**Proof.** Let  $P = \{p_1, \dots, p_m\}$  and let  $n_1, \dots, n_m$  be pairwise coprime natural numbers greater than  $k$ . We extend  $\varepsilon$  to a function from  $\bigcup_{1 \leq t \leq m} \{0, \dots, n_t\} \times \{p_t\}$  by putting  $\varepsilon(i, p_t) = 0$  for any  $i > k$ . Now, for  $1 \leq t \leq m$ , let  $\psi_t$  be an  $\text{SL}^*$  formula from Lemma 18 such that

$$\text{SL}^* \vdash \bigvee_{0 \leq r < n_t} \bigwedge_{0 \leq i < n_t} \Box^{i+r} (\neg)^{\varepsilon(i, p_t)} \psi_t$$

and

$$\text{SL}^* \vdash \psi_t \equiv \Box^{n_t} \psi_t.$$

Now, let  $V$  be an arbitrary valuation admissible for  $\text{SL}^*$  and, for  $1 \leq t \leq m$ , let  $a_t < n_t$  be such that

$$L_\omega, a_t \models \bigwedge_{0 \leq i < n_t} \Box^i (\neg)^{\varepsilon(i, p_t)} \psi_t.$$

We will replace all  $a_t$ 's by a single world  $b$ . By the second property of  $\psi_t$ 's mentioned above, for each  $k \in \omega$ ,

$$L_\omega, a_t + kn_t \models \bigwedge_{0 \leq i < n_t} \Box^i (\neg)^{\varepsilon(i, p_t)} \psi_t.$$

Since  $n_t$ 's are pairwise coprime, by Chinese Remainder Theorem, there exists  $b$  such that, for each  $1 \leq t \leq m$ , the remainder of  $b$  modulo  $n_t$  is  $a_t$ . For this  $b$  we have

$$L_\omega, b \models \bigwedge_{1 \leq t \leq m} \bigwedge_{0 \leq i < n_t} \Box^i (\neg)^{\varepsilon(i, p_t)} \psi_t.$$

Moreover, we can choose  $b < N = \max \{n_t : 1 \leq t \leq m\} + \prod_{1 \leq t \leq m} n_t$ . Since a valuation  $V$  was arbitrary, the following formula is provable in  $\text{SL}^*$ :

$$\bigvee_{0 \leq j < N} \Box^j \left( \bigwedge_{1 \leq t \leq m} \bigwedge_{0 \leq i < n_t} \Box^i (\neg)^{\varepsilon(i, p_t)} \psi_t \right).$$

It follows that a weaker formula below is also provable in  $\text{SL}^*$

$$\bigvee_{0 \leq j < 2N} \Box^j \left( \bigwedge_{1 \leq t \leq m} \bigwedge_{0 \leq i \leq k} \Box^i (\neg)^{\varepsilon(i, p_t)} \psi_t \right).$$

The last formula is equivalent to

$$\bigvee_{0 \leq j < N} \Box^j \Phi_\varepsilon(\psi_1/p_1, \dots, \psi_m/p_m). \quad (1)$$

Now, if  $\neg\Phi_\varepsilon$  is provable in L then, by necessitation and substitution of  $\psi_i$ 's for  $p_i$ 's, also the following formula would be provable

$$\bigwedge_{0 \leq j < N} \Box^j \neg\Phi_\varepsilon(\psi_1/p_1, \dots, \psi_m/p_m). \quad (2)$$

However, (1) is equivalent to the negation of (2). Since (1) is provable in  $\text{SL}^* \subseteq L_{\text{Tr}}$ , it follows that  $\Phi_\varepsilon$  has to be consistent with L.  $\square$

The last lemma shows that if L is a consistent logic such that  $\text{SL}^* \subseteq L$  then L has to be consistent with any valuation described by a function  $\varepsilon: \{1, \dots, k\} \times P \rightarrow \{0, 1\}$ . We use this fact to show that  $L_{\text{Tr}}$  is conservative over  $\text{SL}^*$  as well as to show that  $\text{SL}^*$  is conservative over SL.

**Lemma 20** *Each consistent normal modal logic L extending  $\text{SL}^*$  is conservative over  $\text{SL}^*$  in the language of SL.*

**Proof.** Let L be a modal logic extending SL and let  $\varphi(p_1, \dots, p_n)$  be a formula in the language of SL such that  $\text{SL}^* \not\vdash \varphi$ . We will show that  $L \not\vdash \varphi$ . Let us assume that all variables of  $\varphi$  are among  $p_1, \dots, p_n$ . Let  $P = \{p_1, \dots, p_n\}$  and let  $V: \omega \times P \rightarrow \{0, 1\}$  be a valuation witnessing that  $\text{SL}^* \not\vdash \varphi$ . So, it holds that  $L_\omega, 0 \not\models \varphi[V]$ . In order to determine the logical value of  $\varphi$  we need only to take a look at  $V$  restricted to a set  $\{0, \dots, k\} \times P$ , for some  $k$  greater or equal to the modal depth of  $\varphi$ . Let  $\varepsilon$  be  $V$  restricted to this set. Since  $\Phi_\varepsilon$  is consistent with  $\neg\varphi$ , by Lemma 17 it also implies  $\neg\varphi$ . By Lemma 19,  $\Phi_\varepsilon$  is consistent with L. Thus the formula  $\neg\varphi$  has to be consistent with L, too.  $\square$

**Lemma 21**  *$\text{SL}^*$  is conservative over SL in the language of SL.*

**Proof.** Let  $\varphi$  be such that  $\text{SL} \not\vdash \varphi$  and let  $V$  be a valuation in  $L_\omega$  witnessing this fact that is

$$L_\omega, 0 \not\models \varphi[V].$$

Let  $P$  be all variables of  $\varphi$  let  $k$  be the modal depth of  $\varphi$ . Let  $\varepsilon$  be  $V$  restricted to the set  $\{0, \dots, k\} \times P$ . The formula  $\Phi_\varepsilon$  is consistent with  $\varphi$  thus, by Lemma 17,  $\text{SL} \vdash \Phi_\varepsilon \Rightarrow \varphi$ . By Lemma 19 the formula  $\Phi_\varepsilon$  is consistent with  $\text{SL}^*$  so  $\varphi$  has to be consistent with  $\text{SL}^*$ , too.  $\square$

**Theorem 22** *The logic SL is the modal logic of the truth predicate for  $\models_{\text{sl}}$ .*

**Proof.** Of course the modal logic  $L_{Tr}$  of a weak truth predicate for  $\models_{sl}$  satisfies the assumption of Lemma 20. So, it is a sublogic of  $SL^*$  which is a conservative extension of  $SL$ . Since  $L_{Tr}$  is in the same language as  $SL$ , it follows, that  $L_{Tr}$  is equivalent to  $SL$ .  $\square$

## References

- [Boo] George Boolos. *The logic of provability*.
- [EFT94] H.-D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic*. Springer–Verlag, 1994. second edition.
- [Fra] Torkel Franzén. *Inexhaustibility. A non-exhaustive treatment*.
- [HP93] P. Hájek and P. Pudlák. *Metamathematics of First–Order Arithmetic*. Springer Verlag, 1993.
- [Imm99] Neil Immerman. *Descriptive Complexity*. Springer, 1999.
- [Mos01] M. Mostowski. On representing concepts in finite models. *Mathematical Logic Quarterly*, 47:513–523, 2001.
- [MZ05] M. Mostowski and K. Zdanowski. Beyond  $FM$ –representability. In Benedikt L. S. Barry Cooper, editor, *New Computational Paradigms. First Conference on Computability in Europe*, 2005.
- [Sho93] J. R. Shoenfield. *Recursion Theory*. Lectures Notes in Logic. Springer–Verlag, 1993.