Smullyan’ s truth and provability

Kurahashi, Taishi

Tominaga, Kohei

(Citation)
Journal of Logic and Computation,35(5):exaf001

(Issue Date)
2025-07

(Resource Type)
journal article

(Version)
Accepted Manuscript

Smullyan's truth and provability

Taishi Kurahashi*† and Kohei Tominaga‡§

Abstract

We revisit Smullyan's paper "Truth and Provability" (2013) for three purposes. First, we introduce the notion of Smullyan models to give a precise definition for Smullyan's framework discussed in that paper. Second, we clarify the relationship between three theorems proved by Smullyan and other newly introduced properties for Smullyan models in terms of both implications and non-implications. Third, we construct two Smullyan models based on arithmetical ideas and show the correspondence between the properties of these Smullyan models and those concerning truth and provability in arithmetic.

1 Introduction

Gödel's First Incompleteness Theorem and Tarski's Undefinability Theorem are major achievements in mathematical logic and have had a great impact on mathematics and other fields. A version of the First Incompleteness Theorem states that every computable sound extension of Peano arithmetic is incomplete. The Undefinability Theorem states that the set of all true sentences in the standard model of arithmetic is not definable in the standard model. These theorems are positioned nowadays as basic results in mathematical logic, but of course understanding the proofs of these theorems requires a reasonable amount of knowledge and experience in mathematical logic.

The structures of the proofs of these theorems are themselves very interesting, and Smullyan wrote a number of books and papers to bringing the essence of these structures to the general reader (e.g. [6, 7, 8, 9, 10, 11]). In particular, the article [11] titled "Truth and Provability" published in *The Mathematical Intelligencer* (2013) provided a concise presentation of the structures of the proofs of these theorems using a very simplified framework dealing with finite strings of symbols. He wrote as follows:

*Email: kurahashi@people.kobe-u.ac.jp

†Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan.

‡Email: tominaga@stu.kobe-u.ac.jp

§Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan.

> The purpose of this article is to provide the general reader, even those readers with no familiarity with the symbolism of mathematical logic, with the essential ideas behind the proofs of the Gödel and Tarski theorems. We do this by constructing a very simple system (an abstraction of part of Reference [1][1]), which, despite its simplicity, has enough power for the Tarski and Gödel arguments to go through. First we address Tarski's theorem, and then Gödel's. ([11, p. 21])

Roughly speaking, every Smullyan's system is specified by determining a set of predicates which are finite strings and determining which set of finite strings each predicate names. In addition, Smullyan's framework employs two prefixes n and r for predicates and special rules regarding the naming relation for the predicates prefixed by these symbols.[2] For such a simple system, Smullyan proved the following three theorems.

Theorem (Theorem F)**.** *Every predicate has a fixed point.*

Theorem (Theorem T)**.** *The set of true sentences is not nameable.*

Theorem (Theorem G)**.** *If a predicate P names a set of true sentences, then there is a sentence X that is undecidable in P.*

We revisit Smullyan's "Truth and Provability" in this paper for three main purposes. First, we introduce the notion of *Smullyan models* to give a precise definition for Smullyan's framework in order to make it easier to discuss the framework mathematically. Second, we clarify the relationship between the symbols n and r and the three theorems stated above. For this purpose, in addition to the above three theorems originally considered by Smullyan, we introduce several other properties that would be expected to hold for Smullyan models and analyze their relationships in terms of both implications and non-implications. Third, we construct Smullyan models based on arithmetical ideas and discuss the correspondence between the properties of these models and those concerning truth and provability in arithmetic.

The organization of the present paper is as follows. In Section 2, we introduce the notion of Smullyan models and reprove above mentioned three Smullyan's theorems according to Smullyan models. In Section 3, we introduce several properties of Smullyan models and then prove some equivalences between these properties. The implications between these properties will be summarized in Figure 1. Section 4 is devoted to proving non-implications between some of these properties by giving several counterexamples of Smullyan models. In Section 5, we construct two specific Smullyan models $M_{\mathbb{N}}$ and M_{PA} based on the standard model of arithmetic and Peano arithmetic, respectively. Among other things, we show that a stronger version of Theorem T for $M_{\mathbb{N}}$ actually yields the original Tarski's Undefinability Theorem.

[1] Reference [7] of the present article.

[2] Smullyan adopted the capital letters N and R, while we use n and r, respectively, in view of the use of lower case letters for symbols.

2 Smullyan models and Smullyan's theorems

In this section, we introduce the notion of Smullyan models and reprove Smullyan's three theorems.

For each non-empty set Σ of symbols, let Σ^* denote the set of all finite strings of the elements of Σ. Let ϵ denote the empty string and we assume $\epsilon \in \Sigma^*$. For any $X, Y \in \Sigma^*$, let XY denote the finite string obtained by concatenating Y after the last element of X. For each $X \in \Sigma^*$ and $i \in \mathbb{N}$, X^i is inductively defined as follows: X^0 is ϵ; and X^{i+1} is $X^i X$. So, X^i is $\underbrace{XX \cdots X}_{i}$.

Definition 2.1 (Smullyan models)**.** A triple $M = (\Sigma, \mathtt{Pred}, \Phi)$ is said to be a *Smullyan model* if it satisfies the following conditions:

- Σ is a non-empty set of symbols.
- $\mathtt{Pred}$ is a subset of Σ^* satisfying the following requirement:
  - For any $H \in \mathtt{Pred}$ and $X \in \Sigma^* \setminus \{\epsilon\}$, we have $HX \notin \mathtt{Pred}$. (†)
- Φ is a function $\mathtt{Pred} \to \mathcal{P}(\Sigma^*)$.

For every Smullyan model $M = (\Sigma, \mathtt{Pred}, \Phi)$, we adopt a convention that $\Sigma_M, \Sigma^*_M, \mathtt{Pred}_M$ and Φ_M denote $\Sigma, \Sigma^*, \mathtt{Pred}$ and Φ, respectively.

Definition 2.2 (Predicates and sentences)**.** Let M be a Smullyan model.

- Every element of $\mathtt{Pred}_M$ is called an *M-predicate*.
- A finite string $Y \in \Sigma^*_M$ is said to be an *M-sentence* if it is of the form HX for some $H \in \mathtt{Pred}_M$ and $X \in \Sigma^*_M$. Let $\mathtt{Sent}_M$ denote the set of all M-sentences.
- Let $\mathtt{Sent}^+_M := \mathtt{Sent}_M \setminus \mathtt{Pred}_M$.

Notice that every M-predicate H is an M-sentence because $H \equiv H\epsilon$. Here $X \equiv Y$ means that the finite strings X and Y are identical. So, the definition of $\mathtt{Sent}^+_M$ makes sense. The following lemma explains why the requirement (†) is imposed on the definition of Smullyan models.

Lemma 2.3. *Let M be a Smullyan model. For each M-sentence S, the unique M-predicate H such that S is of the form HX for some $X \in \Sigma^*_M$ is found.*

Proof. Suppose, towards a contradiction, that M-sentences HX and $H'X'$ are identical for some distinct M-predicates H and H'. Without loss of generality, we may assume that H is a proper initial segment of H'. Then, we find a non-empty $Y \in \Sigma^*_M$ such that $H' \equiv HY$. This violates the requirement (†). □

For any Smullyan model M, we say that an M-predicate H *names* a subset $V \subseteq \Sigma^*_M$ if $V = \Phi_M(H)$. For each Smullyan model M, an M-sentence HX is intended to express the statement that 'X is contained in the set of all strings named by the M-predicate H'. Thus, each M-sentence is determined to be *true* or *false* depending on whether the intended statement actually holds or not, respectively.

Definition 2.4. Let M be a Smullyan model.

- $\mathtt{True}_M := \{HX \in \mathtt{Sent}_M \mid H \in \mathtt{Pred}_M \text{ and } X \in \Phi_M(H)\}$.
- $\mathtt{True}^+_M := \{HX \in \mathtt{Sent}^+_M \mid H \in \mathtt{Pred}_M \text{ and } X \in \Phi_M(H)\}$.
- $\mathtt{False}_M := \mathtt{Sent}_M \setminus \mathtt{True}_M$.
- $\mathtt{False}^+_M := \mathtt{Sent}^+_M \setminus \mathtt{True}^+_M$.

For each M-sentence S, we write $M \models S$ if $S \in \mathtt{True}_M$.

Notice that $\mathtt{True}^+_M = \mathtt{True}_M \setminus \mathtt{Pred}_M$ and $\mathtt{False}^+_M = \mathtt{False}_M \setminus \mathtt{Pred}_M$ hold.
In addition to the basic framework described above, Smullyan considered the system equipped with the two special symbols n and r.

Definition 2.5 (n-Smullyan models, r-Smullyan models, and nr-Smullyan models)**.** Let M be a Smullyan model.

- M is called an n-*Smullyan model* if $\mathsf{n} \in \Sigma_M$ and for each $H \in \mathtt{Pred}_M$, we have $\mathsf{n}H \in \mathtt{Pred}_M$ and
$$\Phi_M(\mathsf{n}H) = \Sigma^*_M \setminus \Phi_M(H).$$
- M is called an r-*Smullyan model* if $\mathsf{r} \in \Sigma_M$ and for each $H \in \mathtt{Pred}_M$, we have $\mathsf{r}H \in \mathtt{Pred}_M$ and
$$\Phi_M(\mathsf{r}H) = \{K \in \mathtt{Pred}_M \mid KK \in \Phi_M(H)\}.$$
- M is said to be an nr-*Smullyan model* if it is both an n-Smullyan model and an r-Smullyan model.

For each n-Smullyan model M, the symbol n behaves as the negation, that is, it is easily shown that for any M-sentence S, we have that $M \models \mathsf{n}S$ if and only if $M \not\models S$. The symbol r was used by Smullyan "to suggest the word *repeat*".

Definition 2.6. Let M be a Smullyan model. We say that $S \in \mathtt{Sent}^+_M$ is an M-*fixed point* of an M-predicate H if the following equivalence holds:
$$M \models S \iff M \models HS.$$

Theorem 2.7 (Fixed Point Theorem [11, Theorem F])**.** *Let M be an r-Smullyan model. For any M-predicate H, there exists an M-fixed point of H.*

Proof. For each $H \in \mathtt{Pred}_M$, the following equivalences show that $\mathsf{r}H\mathsf{r}H \in \mathtt{Sent}^+_M$ is an M-fixed point of H:
$$M \models \mathsf{r}H\mathsf{r}H \iff \mathsf{r}H \in \Phi_M(\mathsf{r}H) \iff \mathsf{r}H\mathsf{r}H \in \Phi_M(H) \iff M \models H\mathsf{r}H\mathsf{r}H.$$
□

Theorem 2.8 (Tarski's Undefinability Theorem [11, Theorem T])**.** *For any* nr*-Smullyan model* M*, there is no* M*-predicate that names* $\mathtt{True}_M$*.*

Proof. Let H be any M-predicate. By the Fixed Point Theorem, we find an M-fixed point S of the M-predicate $\mathsf{n}H$. Then, we have

$$\begin{aligned} S \in \mathtt{True}_M &\iff M \models S \iff M \models \mathsf{n}HS \\ &\iff M \not\models HS \iff S \notin \Phi_M(H). \end{aligned}$$

These equivalences show that H does not name $\mathtt{True}_M$. □

Theorem 2.9 (Gödel's First Incompleteness Theorem [11, Theorem G])**.** *Let* M *be an* nr*-Smullyan model. For any* M*-predicate* H *satisfying* $\Phi_M(H) \subseteq \mathtt{True}_M$*, there exists an* M*-sentence* S *such that* $S \notin \Phi_M(H)$ *and* $\mathsf{n}S \notin \Phi_M(H)$*.*

Proof. Suppose that an M-predicate H satisfies $\Phi_M(H) \subseteq \mathtt{True}_M$. By the Fixed Point Theorem, we find an M-fixed point S of the M-predicate $\mathsf{n}H$. We have

$$\begin{aligned} S \in \mathtt{True}_M &\iff M \models S \iff M \models \mathsf{n}HS \\ &\iff M \not\models HS \iff S \notin \Phi_M(H). \end{aligned}$$

By combining these equivalences with the supposition $\Phi_M(H) \subseteq \mathtt{True}_M$, we get $S \in \mathtt{True}_M$ and $S \notin \Phi_M(H)$. Then, we have $\mathsf{n}S \notin \mathtt{True}_M$, which implies $\mathsf{n}S \notin \Phi_M(H)$. We have shown that the M-sentence S witnesses the theorem. □

3 Properties of Smullyan models

In this section, we introduce several properties of Smullyan models. We then prove some equivalences between these properties. We would like to mention here that the equivalence of the Fixed Point Theorem, Tarski's theorem, and Gödel's theorem in arithmetic was discussed by Salehi [3, 4].

Definition 3.1 (Properties of Smullyan models)**.** We consider the following properties of Smullyan models M:

- (FPT) Every M-predicate has an M-fixed point.
- (T-Tarski) There is no M-predicate that names $\mathtt{True}_M$.
- (F-Tarski) There is no M-predicate that names $\mathtt{False}_M$.
- (T-Tarski$^+$) There is no M-predicate H such that $\mathtt{True}^+_M = \Phi_M(H) \cap \mathtt{Sent}^+_M$.
- (F-Tarski$^+$) There is no M-predicate H such that $\mathtt{False}^+_M = \Phi_M(H) \cap \mathtt{Sent}^+_M$.
- (mG1) For any M-predicate H satisfying $\Phi_M(H) \subseteq \mathtt{True}_M$, there exists $S \in \mathtt{Sent}_M$ such that $M \models S$ and $S \notin \Phi_M(H)$.

- (mG1$^+$) For any M-predicate H satisfying $\Phi_M(H) \cap \mathtt{Sent}_M^+ \subseteq \mathtt{True}_M^+$, there exists $S \in \mathtt{Sent}_M^+$ such that $M \models S$ and $S \notin \Phi_M(H)$.

We also consider the following properties of n-Smullyan models M:

- (G1) For any M-predicate H satisfying $\Phi_M(H) \subseteq \mathtt{True}_M$, there exists $S \in \mathtt{Sent}_M$ such that $S \notin \Phi_M(H)$ and $\mathsf{n}S \notin \Phi_M(H)$.

- (G1$^+$) For any M-predicate H satisfying $\Phi_M(H) \cap \mathtt{Sent}_M^+ \subseteq \mathtt{True}_M^+$, there exists $S \in \mathtt{Sent}_M^+$ such that $S \notin \Phi_M(H)$ and $\mathsf{n}S \notin \Phi_M(H)$.

FPT stands for 'Fixed Point Theorem' and Smullyan's Theorem F (Theorem 2.7) states that every r-Smullyan model satisfies FPT. T-Tarski and F-Tarski state that Tarski's Undefinability Theorem holds for M with respect to $\mathtt{True}_M$ and $\mathtt{False}_M$, respectively. Theorem 2.8 states that every nr-Smullyan model satisfies T-Tarski. Although T-Tarski and F-Tarski seem to be equivalent, indeed they are not. In fact, we will prove in the next section that these properties are incomparable even if we consider n-Smullyan models (Propositions 4.9 and 4.10). This incomparability is caused by the reason that, in general, the set named by an M-predicate H may contain finite strings that are not M-sentences, and such H trivially names neither $\mathtt{True}_M$ nor $\mathtt{False}_M$. Thus, we consider T-Tarski$^+$ and F-Tarski$^+$, which are stronger versions of T-Tarski and F-Tarski, respectively, in which this triviality is removed. The reason why we defined these strong properties using $\mathtt{Sent}_M^+$ rather than $\mathtt{Sent}_M$ is to yield meaningful properties of arithmetic in Section 5. In the next section, we will prove that the stronger versions are actually strictly stronger than the original ones. We also prove in this section that T-Tarski$^+$ and F-Tarski$^+$ are equivalent for any n-Smullyan models (Proposition 3.4). Interestingly, this stronger version F-Tarski$^+$ of F-Tarski is equivalent to FPT for any Smullyan model (Proposition 3.2).

Theorem 2.9 states that every nr-Smullyan model satisfies G1, where G1 stands for 'Gödel's 1st Theorem'. However, G1 needs to consider n-Smullyan models to state it, so it is a bit awkward for our purposes of analyzing the general situation of Smullyan models. For this reason, we introduce a new property mG1 which corresponds to a version of Gödel's First Incompleteness Theorem stating that 'every computable sound extension of Peano arithmetic has a true but unprovable sentence'. Here 'm' stands for 'modified'. We prove in this section that G1 and mG1 are equivalent for any n-Smullyan model (Proposition 3.5). Furthermore, we prove that T-Tarski and mG1 are equivalent for every Smullyan model (Proposition 3.3). The properties G1$^+$ and mG1$^+$ are stronger versions of G1 and mG1, respectively, in which the triviality is removed.

3.1 Equivalences between the properties

We show several equivalences between the properties introduced above. The results of this section are summarized in Figure 1. In conclusion of this section,

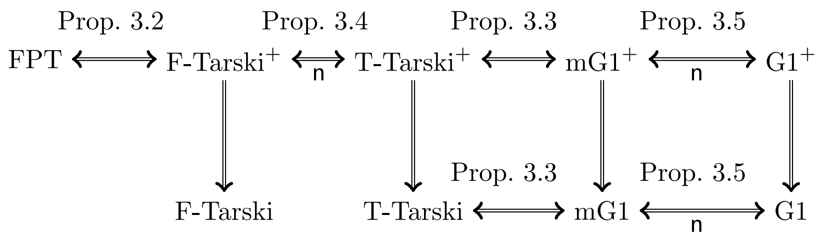

Figure 1: Implications between the properties

it is sufficient to consider the four properties T-Tarski, F-Tarski, T-Tarski$^+$, and F-Tarski$^+$ when dealing with the properties we have introduced.

Notice that what Smullyan achieved can also be understood through the equivalences shown in Figure 1. For, Theorem 2.7 states that every r-Smullyan model satisfies FPT, and hence Figure 1 shows that every nr-Smullyan model satisfies all the properties indicated in the figure.

Proposition 3.2. FPT *and* F-Tarski$^+$ *are equivalent for any Smullyan model.*

Proof. Let M be any Smullyan model. This proposition is proved by the following equivalences:

$$\begin{aligned}
\text{FPT} &\iff ({}^{\forall}H \in \mathtt{Pred}_M)\left({}^{\exists}S \in \mathtt{Sent}_M^+\right)(M \models S \iff M \models HS)\\
&\iff ({}^{\forall}H \in \mathtt{Pred}_M)\left({}^{\exists}S \in \mathtt{Sent}_M^+\right)(S \in \mathtt{True}_M \iff S \in \Phi_M(H))\\
&\iff ({}^{\forall}H \in \mathtt{Pred}_M)\left({}^{\exists}S \in \mathtt{Sent}_M^+\right)\\
&\qquad\qquad\qquad \left(S \in \mathtt{False}_M^+ \iff S \notin \Phi_M(H) \cap \mathtt{Sent}_M^+\right)\\
&\iff ({}^{\forall}H \in \mathtt{Pred}_M)\left(\mathtt{False}_M^+ \neq \Phi_M(H) \cap \mathtt{Sent}_M^+\right)\\
&\iff \text{F-Tarski}^+.
\end{aligned}$$

□

Proposition 3.3.

1. T-Tarski *and* mG1 *are equivalent for any Smullyan model.*
2. T-Tarski$^+$ *and* mG1$^+$ *are equivalent for any Smullyan model.*

Proof. We only prove the first clause. The second clause is proved in the similar way. Let M be any Smullyan model. It is easy to see that for each $H \in \mathtt{Pred}_M$, the condition $\mathtt{True}_M \neq \Phi_M(H)$ is equivalent to

$$\text{“}\Phi_M(H) \subseteq \mathtt{True}_M \Rightarrow {}^{\exists}S \in \mathtt{Sent}_M(M \models S \;\&\; S \notin \Phi_M(H))\text{”}.$$

This shows that T-Tarski and mG1 are equivalent for M. □

Proposition 3.4. F-Tarski$^+$ *and* T-Tarski$^+$ *are equivalent for any* n*-Smullyan model.*

Proof. Let M be any n-Smullyan model. Notice that $\Phi_M(H) = \Phi_M(\mathsf{nn}H)$ holds for any $H \in \mathtt{Pred}_M$. Then, this proposition is shown by the following equivalences:

$$\begin{aligned}
\text{F-Tarski}^+ &\iff (^\forall H \in \mathtt{Pred}_M)\bigl(\mathtt{False}^+_M \neq \Phi_M(H) \cap \mathtt{Sent}^+_M\bigr)\\
&\iff (^\forall H \in \mathtt{Pred}_M)\bigl[\bigl(\mathtt{False}^+_M \neq \Phi_M(H) \cap \mathtt{Sent}^+_M\bigr)\\
&\qquad\qquad \& \ \bigl(\mathtt{False}^+_M \neq \Phi_M(\mathsf{n}H) \cap \mathtt{Sent}^+_M\bigr)\bigr]\\
&\iff (^\forall H \in \mathtt{Pred}_M)\bigl[\bigl(\mathtt{False}^+_M \neq (\Sigma^*_M \setminus \Phi(\mathsf{n}H)) \cap \mathtt{Sent}^+_M\bigr)\\
&\qquad\qquad \& \ \bigl(\mathtt{False}^+_M \neq (\Sigma^*_M \setminus \Phi_M(H)) \cap \mathtt{Sent}^+_M\bigr)\bigr]\\
&\iff (^\forall H \in \mathtt{Pred}_M)\bigl[\bigl(\mathtt{Sent}^+_M \setminus \mathtt{True}^+_M \neq \mathtt{Sent}^+_M \setminus \Phi_M(\mathsf{n}H)\bigr)\\
&\qquad\qquad \& \ \bigl(\mathtt{Sent}^+_M \setminus \mathtt{True}^+_M \neq \mathtt{Sent}^+_M \setminus \Phi_M(H)\bigr)\bigr]\\
&\iff (^\forall H \in \mathtt{Pred}_M)\bigl[\bigl(\mathtt{True}^+_M \neq \Phi_M(H) \cap \mathtt{Sent}^+_M\bigr)\\
&\qquad\qquad \& \ \bigl(\mathtt{True}^+_M \neq \Phi_M(\mathsf{n}H) \cap \mathtt{Sent}^+_M\bigr)\bigr]\\
&\iff (^\forall H \in \mathtt{Pred}_M)\bigl(\mathtt{True}^+_M \neq \Phi_M(H) \cap \mathtt{Sent}^+_M\bigr)\\
&\iff \text{T-Tarski}^+.
\end{aligned}$$

□

Proposition 3.5.

1. mG1 *and* G1 *are equivalent for any* n*-Smullyan model.*
2. mG1^+ *and* G1^+ *are equivalent for any* n*-Smullyan model.*

Proof. We only prove the first clause. The second clause is proved in a similar way. Let M be any n-Smullyan model.

(mG1 $\Rightarrow$ G1): Suppose that M satisfies mG1. Let H be any M-predicate such that $\Phi_M(H) \subseteq \mathtt{True}_M$. By mG1, there is some $S \in \mathtt{Sent}_M$ such that $M \models S$ and $S \notin \Phi_M(H)$. Then, we have $M \not\models \mathsf{n}S$ and hence $\mathsf{n}S \notin \mathtt{True}_M$. Since $\Phi_M(H) \subseteq \mathtt{True}_M$, we obtain $\mathsf{n}S \notin \Phi_M(H)$. We have shown that S witnesses G1 for M.

(G1 $\Rightarrow$ mG1): Suppose that M satisfies G1. Let H be any M-predicate such that $\Phi_M(H) \subseteq \mathtt{True}_M$. By G1, there is some $S \in \mathtt{Sent}_M$ such that $S \notin \Phi_M(H)$ and $\mathsf{n}S \notin \Phi_M(H)$. Depending on whether $M \models S$ or $M \models \mathsf{n}S$, we have that S or $\mathsf{n}S$ is a witness of mG1 for M, respectively. □

4 Non-implications

As mentioned before, the results we presented in the previous section indicate that it suffices to analyze the relationship between the four properties F-Tarski$^+$, F-Tarski, T-Tarski$^+$, and T-Tarski. The situation of the implications between the combinations of these properties is visualized in Figure 2.

In this section, we prove that in general no more arrows can be added into Figure 2. Concretely we prove the following non-implications.

- F-Tarski$^+$ $\overset{\mathsf{r}}{\not\Rightarrow}$ T-Tarski (Proposition 4.4).

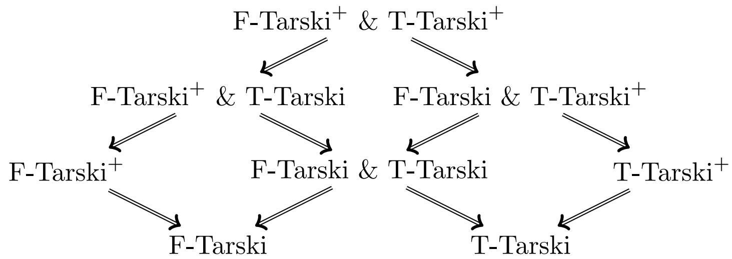

Figure 2: Implications between the properties

- T-Tarski$^+$ $\not\Rightarrow$ F-Tarski (Proposition 4.5).
- F-Tarski & T-Tarski $\overset{\mathsf{n}}{\not\Rightarrow}$ F-Tarski$^+$ (Proposition 4.6).
- F-Tarski & T-Tarski $\overset{\mathsf{n}}{\not\Rightarrow}$ T-Tarski$^+$ (Proposition 4.6).
- F-Tarski$^+$ & T-Tarski $\overset{\mathsf{r}}{\not\Rightarrow}$ T-Tarski$^+$ (Proposition 4.7).
- F-Tarski & T-Tarski$^+$ $\not\Rightarrow$ F-Tarski$^+$ (Proposition 4.8).
- F-Tarski $\overset{\mathsf{n}}{\not\Rightarrow}$ T-Tarski (Proposition 4.9).
- T-Tarski $\overset{\mathsf{n}}{\not\Rightarrow}$ F-Tarski (Proposition 4.10).

The relation $\overset{\mathsf{r}}{\not\Rightarrow}$ (resp. $\overset{\mathsf{n}}{\not\Rightarrow}$) indicates that the non-implication is shown by giving a counter-model which is an r-Smullyan model (resp. n-Smullyan model). As a consequence of these non-implications, no more arrows can be added into Figure 1 as well.

Our counter-models presented in this section are restricted and tractable versions of the general Smullyan models, and we first introduce the notion of these *simple models*. We fix a symbol $\sharp$, which is different from both n and r.

Definition 4.1 (Simple models)**.** A tuple $M = (\Sigma, \Phi)$ is said to be a *simple model* if it satisfies the following conditions:

- Σ is a non-empty set of symbols such that $\sharp \in \Sigma$.
- Let $\mathtt{Pred}_M := \{X\sharp \mid X \in \Sigma^* \text{ and } X \text{ does not contain } \sharp\}$.
- Φ is a function $\mathtt{Pred}_M \to \mathcal{P}(\Sigma^*)$.

It is easy to show that every simple model satisfies the requirement (†) in the definition of Smullyan models. So, we have the following proposition.

Proposition 4.2. *For every simple model* $M = (\Sigma, \Phi)$*, the triple* $(\Sigma, \mathtt{Pred}_M, \Phi)$ *is a Smullyan model.*

Thus, in the following, we will deal with simple models as Smullyan models. The following proposition says that for every simple model, it is very easy to determine whether a given finite string is a sentence or not. In this sense, simple models are easy to handle.

Proposition 4.3. *For every simple model $M = (\Sigma, \Phi)$, we have*

$$\texttt{Sent}_M = \{X \in \Sigma^* \mid X \textit{ contains at least one } \sharp\}.$$

We are ready to give our counter-models of several implications.

Proposition 4.4. *There is an* r*-simple model which satisfies* F-Tarski$^+$ *but does not satisfy* T-Tarski*.*

Proof. Let M be the r-simple model defined as follows:

- $\Sigma_M := \{\mathsf{r}, \sharp\}$,
- $\Phi_M(\sharp) := \emptyset$.

It is easy to show that $\texttt{Pred}_M = \{\mathsf{r}^i\sharp \mid i \in \mathbb{N}\}$. Then, by the definition of r-Smullyan models, it can be shown that $\Phi_M(H) = \emptyset$ for all $H \in \texttt{Pred}_M$. Therefore, we obtain $\texttt{True}_M = \emptyset = \Phi_M(\sharp)$. This means that M does not satisfy T-Tarski. By Theorem 2.7 and Proposition 3.2, M satisfies F-Tarski$^+$. □

We also give a proof of Proposition 4.4 with a counter-model having a slightly non-trivial function Φ.

Alternative proof of Proposition 4.4. Let M be the r-simple model defined as follows:

- $\Sigma_M := \{\mathsf{r}, \sharp\}$,
- $\Phi_M(\sharp) := \{\sharp^i \mathsf{r} \sharp \mathsf{r} \sharp \mid i \in \mathbb{N}\}$.

As above, we have that M satisfies F-Tarski$^+$ and $\texttt{Pred}_M = \{\mathsf{r}^i\sharp \mid i \in \mathbb{N}\}$. Since the only element of $\Phi_M(\sharp)$ which is of the form KK for some M-predicate K is $\mathsf{r}\sharp\mathsf{r}\sharp$, we have $\Phi_M(\mathsf{r}\sharp) = \{\mathsf{r}\sharp\}$. Then, it is easy to see that $\Phi_M(\mathsf{r}^i\sharp) = \emptyset$ for all $i \geq 2$. Therefore, we obtain

$$\texttt{True}_M = \{\sharp\sharp^i \mathsf{r}\sharp\mathsf{r}\sharp \mid i \in \mathbb{N}\} \cup \{\mathsf{r}\sharp\mathsf{r}\sharp\} = \{\sharp^i \mathsf{r}\sharp\mathsf{r}\sharp \mid i \in \mathbb{N}\} = \Phi_M(\sharp).$$

This implies that M does not satisfy T-Tarski. □

Proposition 4.5. *There is a simple model which satisfies* T-Tarski$^+$ *but does not satisfy* F-Tarski*.*

Proof. Let M be the simple model defined as follows:

- $\Sigma_M := \{\sharp\}$,
- $\Phi_M(\sharp) := \{\sharp^{2i+1} \mid i \in \mathbb{N}\}$.

We have $\texttt{Pred}_M = \{\sharp\}$, $\texttt{Sent}_M = \{\sharp^i \mid i \geq 1\}$, $\texttt{Sent}^+_M = \{\sharp^i \mid i \geq 2\}$, $\texttt{True}_M = \texttt{True}^+_M = \{\sharp^{2i+2} \mid i \in \mathbb{N}\}$, and $\texttt{False}_M = \{\sharp^{2i+1} \mid i \in \mathbb{N}\}$. Since $\texttt{True}^+_M \neq \Phi_M(\sharp) \cap \texttt{Sent}^+_M$, we have that M satisfies T-Tarski$^+$. On the other hand, since $\texttt{False}_M = \Phi_M(\sharp)$, we get that M does not satisfy F-Tarski. □

Proposition 4.6. *There is an* n*-simple model which satisfies both* F-Tarski *and* T-Tarski *but neither* F-Tarski$^+$ *nor* T-Tarski$^+$*.*

Proof. Let M be the n-simple model defined as follows:

- $\Sigma_M := \{\mathsf{n}, \sharp\}$,
- For each $X \in \Sigma^*_M$, let $n(X)$ be the number of occurrences of n in X,
- $\Phi_M(\sharp) := \{X \in \Sigma^*_M \mid n(X) \text{ is even}\}$.

We have $\texttt{Pred}_M = \{\mathsf{n}^i\sharp \mid i \in \mathbb{N}\}$. For each $i \in \mathbb{N}$, it is shown that

1. $\Phi_M(\mathsf{n}^{2i}\sharp) = \{X \in \Sigma^*_M \mid n(X) \text{ is even}\}$,
2. $\Phi_M(\mathsf{n}^{2i+1}\sharp) = \{X \in \Sigma^*_M \mid n(X) \text{ is odd}\}$.

Since the strings ϵ and n are not M-sentences, $\epsilon \in \Phi_M(\mathsf{n}^{2i}\sharp)$, and $\mathsf{n} \in \Phi_M(\mathsf{n}^{2i+1}\sharp)$, we obtain that $\Phi_M(H)$ coincides with neither $\texttt{False}_M$ nor $\texttt{True}_M$ for every $H \in \texttt{Pred}_M$. This means that M satisfies both F-Tarski and T-Tarski.

On the other hand, we have

$$\begin{aligned}
\texttt{True}^+_M &= (\{\mathsf{n}^{2i}\sharp X \mid i \in \mathbb{N} \ \&\ n(X) \text{ is even}\} \\
&\qquad \cup \{\mathsf{n}^{2i+1}\sharp X \mid i \in \mathbb{N} \ \&\ n(X) \text{ is odd}\}) \setminus \texttt{Pred}_M \\
&= \{X \in \texttt{Sent}_M \mid n(X) \text{ is even}\} \setminus \texttt{Pred}_M \\
&= \{X \in \texttt{Sent}^+_M \mid n(X) \text{ is even}\} \\
&= \Phi_M(\sharp) \cap \texttt{Sent}^+_M.
\end{aligned}$$

Therefore, M does not satisfy T-Tarski$^+$. Since M is an n-Smullyan model, by Proposition 3.4, M also does not satisfy F-Tarski$^+$. □

Proposition 4.7. *There is an* r*-simple model which satisfies* F-Tarski$^+$ *and* T-Tarski *but does not satisfy* T-Tarski$^+$*.*

Proof. Let M be the r-simple model defined as follows:

- $\Sigma_M := \{\mathsf{r}, \sharp\}$,
- $\Phi_M(\sharp) := \{\sharp^i, \sharp^i\mathsf{r}\sharp\sharp \mid i \in \mathbb{N}\}$.

We have that M satisfies F-Tarski$^+$ and $\texttt{Pred}_M = \{\mathsf{r}^j\sharp \mid j \in \mathbb{N}\}$. Since the only element of $\Phi_M(\sharp)$ which is of the form KK for some $K \in \texttt{Pred}_M$ is $\sharp\sharp$, we have $\Phi_M(\mathsf{r}\sharp) = \{\sharp\}$. Then, it is easy to see that $\Phi_M(\mathsf{r}^j\sharp) = \emptyset$ for all $j \geq 2$. Therefore, we get

$$\texttt{True}_M = \{\sharp^{i+1}, \sharp^{i+1}\mathsf{r}\sharp\sharp \mid i \in \mathbb{N}\} \cup \{\mathsf{r}\sharp\sharp\} = \{\sharp^{i+1}, \sharp^i\mathsf{r}\sharp\sharp \mid i \in \mathbb{N}\}.$$

Since $\texttt{True}_M \neq \Phi_M(H)$ for all $H \in \texttt{Pred}_M$, we have that M satisfies T-Tarski. On the other hand, since $\epsilon, \sharp \notin \texttt{Sent}^+_M$, we get $\texttt{True}^+_M = \Phi_M(\sharp) \cap \texttt{Sent}^+_M$. Hence, M does not satisfy T-Tarski$^+$. $\square$

Proposition 4.8. *There is a simple model which satisfies* T-Tarski$^+$ *and* F-Tarski *but does not satisfy* F-Tarski$^+$.

Proof. Let M be the simple model defined as follows:

- $\Sigma_M := \{\sharp\}$,
- $\Phi_M(\sharp) := \{\sharp^{2i} \mid i \in \mathbb{N}\}$.

We have that $\texttt{Pred}_M = \{\sharp\}$, $\texttt{Sent}_M = \{\sharp^i \mid i \geq 1\}$, $\texttt{Sent}^+_M = \{\sharp^i \mid i \geq 2\}$, $\texttt{True}^+_M = \{\sharp^{2i+1} \mid i \geq 1\}$, and $\texttt{False}_M = \texttt{False}^+_M = \{\sharp^{2i+2} \mid i \in \mathbb{N}\}$. Since $\texttt{True}^+_M \neq \Phi_M(\sharp) \cap \texttt{Sent}^+_M$, we have that M satisfies T-Tarski$^+$. Since $\epsilon \in \Phi_M(\sharp)$ and $\epsilon \notin \texttt{Sent}_M$, we have $\texttt{False}_M \neq \Phi_M(\sharp)$, that is, M satisfies F-Tarski. On the other hand, we get $\texttt{False}^+_M = \Phi_M(\sharp) \cap \texttt{Sent}^+_M$. Hence, M does not satisfy F-Tarski$^+$. $\square$

Proposition 4.9. *There is an* n*-simple model which satisfies* F-Tarski *but does not satisfy* T-Tarski.

Proof. Let M be the n-simple model defined as follows:

- $\Sigma_M := \{\mathsf{n}, \sharp\}$,
- $\Phi_M(\sharp) := \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\}$, where $n(X)$ is defined as in the proof of Proposition 4.6.

We have $\texttt{Pred}_M = \{\mathsf{n}^i\sharp \mid i \in \mathbb{N}\}$, $\texttt{Sent}_M = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } X \in \Sigma^*_M\}$, and $\Sigma^*_M \setminus \texttt{Sent}_M = \{\mathsf{n}^i \mid i \in \mathbb{N}\}$. It is shown that for each $j \in \mathbb{N}$,

1. $\Phi_M(\mathsf{n}^{2j}\sharp) = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\}$,
2. $\Phi_M(\mathsf{n}^{2j+1}\sharp) = \{\mathsf{n}^i, X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is even}\}$.

We have

$$\begin{aligned}\texttt{True}_M &= \{\mathsf{n}^{2j}\sharp X \sharp \mathsf{n}^i \mid i,j \in \mathbb{N} \,\&\, n(X) \text{ is odd}\} \\ &\quad \cup \{\mathsf{n}^{2j+1}\sharp \mathsf{n}^i, \mathsf{n}^{2j+1}\sharp X \sharp \mathsf{n}^i \mid i,j \in \mathbb{N} \,\&\, n(X) \text{ is even}\} \\ &= \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\} \\ &= \Phi_M(\sharp).\end{aligned}$$

Therefore, M does not satisfy T-Tarski.

Also, since

$$\texttt{False}_M = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is even}\},$$

it is easy to see that for every $H \in \texttt{Pred}_M$, we have $\texttt{False}_M \neq \Phi_M(H)$. This means that M satisfies F-Tarski. $\square$

Proposition 4.10. *There is an* n*-simple model which satisfies* T-Tarski *but does not satisfy* F-Tarski.

Proof. Let M be the n-simple model defined as follows:

- $\Sigma_M := \{\mathsf{n}, \sharp\}$,
- $\Phi_M(\sharp) := \{\mathsf{n}^i, X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\}$.

We have $\mathtt{Pred}_M = \{\mathsf{n}^i\sharp \mid i \in \mathbb{N}\}$, $\mathtt{Sent}_M = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } X \in \Sigma_M^*\}$, and $\Sigma_M^* \setminus \mathtt{Sent}_M = \{\mathsf{n}^i \mid i \in \mathbb{N}\}$. It is shown that for each $j \in \mathbb{N}$,

1. $\Phi_M(\mathsf{n}^{2j}\sharp) = \{\mathsf{n}^i, X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\}$,
2. $\Phi_M(\mathsf{n}^{2j+1}\sharp) = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is even}\}$.

We have

$$\begin{aligned}\mathtt{True}_M &= \{\mathsf{n}^{2j}\sharp \mathsf{n}^i, \mathsf{n}^{2j}\sharp X \sharp \mathsf{n}^i \mid i, j \in \mathbb{N} \;\&\; n(X) \text{ is odd}\} \\ &\quad \cup \{\mathsf{n}^{2j+1}\sharp X \sharp \mathsf{n}^i \mid i, j \in \mathbb{N} \;\&\; n(X) \text{ is even}\} \\ &= \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is odd}\}.\end{aligned}$$

Since it is easy to see that $\mathtt{True}_M \neq \Phi_M(H)$ for every $H \in \mathtt{Pred}_M$, we have that M satisfy T-Tarski.

On the other hand, since

$$\mathtt{False}_M = \{X \sharp \mathsf{n}^i \mid i \in \mathbb{N} \text{ and } n(X) \text{ is even}\} = \Phi_M(\mathsf{n}\,\sharp),$$

we obtain that M does not satisfies F-Tarski. □

5 Smullyan models based on arithmetic

In this section, we discuss the property of first-order arithmetic through the analysis of Smullyan models. We construct two specific Smullyan models $M_{\mathbb{N}}$ and M_{PA} based on arithmetic and we then discuss the correspondence between the properties of these models and those concerning truth and provability in arithmetic.

Let $\mathcal{L}_A$ be the language of first-order arithmetic. We may assume that $\mathcal{L}_A$ contains the corresponding function symbol for each primitive recursive function. Let $\mathbb{N}$ denote the $\mathcal{L}_A$-structure that is the standard model of first-order arithmetic. Peano arithmetic PA is the $\mathcal{L}_A$-theory consisting of basic axioms of arithmetic and the axiom scheme of induction. For each $n \in \mathbb{N}$, let $\overline{n}$ denote the canonical closed $\mathcal{L}_A$-term whose value is n. We fix a natural Gödel numbering of symbols and formulas in $\mathcal{L}_A$ and let $\ulcorner \varphi \urcorner$ denote the Gödel number of an $\mathcal{L}_A$-formula φ. For a variable v, we say that an $\mathcal{L}_A$-formula φ is a *v-formula* if φ contains no free variables other than v. A v-formula having the Gödel number i is denoted by $\varphi_i(v)$. Also, let $\Gamma := \{i \in \mathbb{N} \mid i$ is the Gödel number of some

v-formula$\}$. Let d be the primitive recursive function satisfying the following equality:

$$d(i) = \begin{cases} \ulcorner \varphi_i(\bar{i}) \urcorner & \text{if } i \in \Gamma, \\ 0 & \text{if } i \notin \Gamma. \end{cases}$$

It is allowed that a v-formula φ_i may contain no free variables. In this case, we have $d(i) = \ulcorner \varphi_i \urcorner$. As noted above, we may assume that PA has the function symbol d corresponding to this primitive recursive function. For the sake of simplicity, for each $\mathcal{L}_A$-formula ψ, we will identify $\overline{\ulcorner \psi \urcorner}$ and $\ulcorner \psi \urcorner$ in this section.

5.1 Arithmetic-based Smullyan frames

We prepare an infinite sequence $\{a_i \mid i \in \mathbb{N}\}$ of fresh symbols. We are ready to bring arithmetic into Smullyan models. In the next two subsections, we will define two Smullyan models $M_{\mathbb{N}}$ and M_{PA} based on the standard model $\mathbb{N}$ of arithmetic and Peano arithmetic PA, respectively. For this purpose, in this subsection, we introduce two frames F_{r} and F_{nr} based on arithmetic as a small step.

Definition 5.1 (Arithmetic-based Smullyan frames)**.**

- $\Sigma_{\mathsf{r}} := \{a_i \mid i \in \Gamma\} \cup \{\mathsf{r}\}$.
- $\mathtt{Pred}_{\mathsf{r}} = \{\mathsf{r}^j a_i \mid j \in \mathbb{N} \text{ and } i \in \Gamma\}$.
- $\Sigma_{\mathsf{nr}} := \{a_i \mid i \in \Gamma\} \cup \{\mathsf{n}, \mathsf{r}\}$.
- $\mathtt{Pred}_{\mathsf{nr}} := \{X a_i \mid X \in \{\mathsf{n}, \mathsf{r}\}^* \text{ and } i \in \Gamma\}$.

1. The tuple $F_{\mathsf{r}} := (\Sigma_{\mathsf{r}}, \mathtt{Pred}_{\mathsf{r}})$ is called the *arithmetic-based* r-*Smullyan frame.*
2. The tuple $F_{\mathsf{nr}} := (\Sigma_{\mathsf{nr}}, \mathtt{Pred}_{\mathsf{nr}})$ is called the *arithmetic-based* nr-*Smullyan frame.*

In the following, F_A is assumed to denote F_{r} or F_{nr}. Here, the subscript A stands for 'arithmetic'. If F_A is F_{r}, then Σ_{F_A} and $\mathtt{Pred}_{F_A}$ denote Σ_{r} and $\mathtt{Pred}_{\mathsf{r}}$, respectively. Similarly, if F_A is F_{nr}, then Σ_{F_A} and $\mathtt{Pred}_{F_A}$ denote Σ_{nr} and $\mathtt{Pred}_{\mathsf{nr}}$, respectively. The following proposition easily follows from the definitions of $\mathtt{Pred}_{\mathsf{r}}$ and $\mathtt{Pred}_{\mathsf{nr}}$.

Proposition 5.2. *For any function $\Phi : \mathtt{Pred}_{F_A} \to \mathcal{P}(\Sigma^*_{F_A})$, we have that $(F_A, \Phi) = (\Sigma_{F_A}, \mathtt{Pred}_{F_A}, \Phi)$ is a Smullyan model.*

Notice that for each Smullyan model M, the set of all M-sentences is determined only by its frame $(\Sigma_M, \mathtt{Pred}_M)$. So, we write the set of all sentences determined by the frame F_A as $\mathtt{Sent}_{F_A}$.

To associate our arithmetic-based Smullyan frames with arithmetic, we introduce two mappings I_{F_A} and J_{F_A}. First, we introduce the mapping I_{F_A} which assigns a v-formula to every F_A-predicate.

Definition 5.3 (The mapping I_{F_A})**.** The mapping I_{F_A} from $\texttt{Pred}_{F_A}$ to the set of all v-formulas is defined as follows:

$$I_{F_A}(Xa_i) :\equiv \begin{cases} \neg^k \varphi_i(v) & \text{if } {}^\exists k \geq 0\,(X \equiv \mathsf{n}^k), \\ \neg^{k+l} \varphi_i(d(v)) & \text{if } {}^\exists k, l \geq 0\,(X \equiv \mathsf{n}^k \mathsf{r} \mathsf{n}^l), \\ \neg^{k+l} \bot & \text{if } {}^\exists k, l \geq 0\, {}^\exists Y\,(X \equiv \mathsf{n}^k \mathsf{r} \mathsf{n}^l \mathsf{r} Y). \end{cases}$$

Here $\neg^k$ is the abbreviation for $\underbrace{\neg \cdots \neg}_{k}$. In particular, the definition of the mapping I_{F_r} is simply rewritten as follows:

$$I_{F_\mathsf{r}}(Xa_i) :\equiv \begin{cases} \varphi_i(v) & \text{if } X \equiv \epsilon, \\ \varphi_i(d(v)) & \text{if } X \equiv \mathsf{r}, \\ \bot & \text{if } {}^\exists j \geq 2\,(X \equiv \mathsf{r}^j). \end{cases}$$

Next, we define the mapping J_{F_A} which assigns an $\mathcal{L}_A$-sentence to each F_A-sentence.

Definition 5.4 (The mapping J_{F_A})**.** The mapping J_{F_A} from $\texttt{Sent}_{F_A}$ to a set of $\mathcal{L}_A$-sentences is inductively defined as follows: Let $H \in \texttt{Pred}_{F_A}$ and $X \in \Sigma^*_{F_A}$.

1. For $X \in \texttt{Pred}_{F_A}$, $J_{F_A}(HX) :\equiv I_{F_A}(H)(\ulcorner I_{F_A}(X)(v) \urcorner)$.
2. For $X \notin \texttt{Sent}_{F_A}$,
   - $J_{F_A}(a_i X) :\equiv I_{F_A}(a_i)(0)$,
   - $J_{F_A}(\mathsf{r}HX) :\equiv \bot$,
   - $J_{F_A}(\mathsf{n}HX) :\equiv \neg J_{F_A}(HX)$.
3. For $X \in \texttt{Sent}^+_{F_A}$,
   - $J_{F_A}(a_i X) :\equiv I_{F_A}(a_i)(\ulcorner J_{F_A}(X) \urcorner)$,
   - $J_{F_A}(\mathsf{r}HX) :\equiv \bot$,
   - $J_{F_A}(\mathsf{n}HX) :\equiv \neg J_{F_A}(HX)$.

   In each of the cases of $X \notin \texttt{Sent}_{F_A}$ and $X \in \texttt{Sent}^+_{F_A}$, the third bullet point is applied only when F_A is F_{nr}.

For the arithmetic-based nr-Smullyan frame F_{nr}, we show that the symbol n plays the role of $\neg$ in arithmetic through the mappings $I_{F_{\mathsf{nr}}}$ and $J_{F_{\mathsf{nr}}}$.

Proposition 5.5. *Suppose $F_A = F_{\mathsf{nr}}$.*

1. *For any $H \in \texttt{Pred}_{F_A}$, we have $I_{F_A}(\mathsf{n}H) \equiv \neg I_{F_A}(H)$.*
2. *For any $S \in \texttt{Sent}_{F_A}$, we have $J_{F_A}(\mathsf{n}S) \equiv \neg J_{F_A}(S)$.*

Proof. 1. Since H and $\mathsf{n}H$ contain the same number of r's, we have that $I_{F_A}(\mathsf{n}H)$ is exactly $\neg I_{F_A}(H)$ by the definition of I_{F_A}.

2. Let $S \equiv HX$ for $H \in \mathtt{Pred}_{F_A}$ and $X \in \Sigma^*_{F_A}$. If $X \notin \mathtt{Pred}_{F_A}$, then $X \notin \mathtt{Sent}_{F_A}$ or $X \in \mathtt{Sent}^+_{F_A}$, and hence this proposition is trivial by the definition of J_{F_A}. If $X \in \mathtt{Pred}_{F_A}$, then we have

$$\begin{aligned} J_{F_A}(\mathsf{n}HX) &\overset{\text{Def. 5.4}}{\equiv} I_{F_A}(\mathsf{n}H)(\ulcorner I_{F_A}(X)(v)\urcorner) \\ &\overset{\text{Clause 1}}{\equiv} \neg I_{F_A}(H)(\ulcorner I_{F_A}(X)(v)\urcorner) \\ &\overset{\text{Def. 5.4}}{\equiv} \neg J_{F_A}(HX). \end{aligned}$$

□

Next, we investigate the effect of the symbol r in arithmetic through the mapping $J_{F_{\mathsf{nr}}}$. For this purpose, we prepare the following lemma.

Lemma 5.6. *Let $H \in \mathtt{Pred}_{F_A}$ and $X \in \mathtt{Sent}^+_{F_A}$. If H contains no r, then $J_{F_A}(HX) \equiv I_{F_A}(H)(\ulcorner J_{F_A}(X)\urcorner)$.*

Proof. It suffices to show that for every $k \geq 0$ and $i \in \Gamma$, we have $J_{F_A}(\mathsf{n}^k a_i X) \equiv I_{F_A}(\mathsf{n}^k a_i)(\ulcorner J_{F_A}(X)\urcorner)$. We prove this by induction on k.

For $k = 0$, $J_{F_A}(a_i X) \equiv I_{F_A}(a_i)(\ulcorner J_{F_A}(X)\urcorner)$ directly follows from Definition 5.4.

Suppose that the lemma holds for k. We prove the lemma holds for $k+1$. In this case, we have $F_A = F_{\mathsf{nr}}$.

$$\begin{aligned} J_{F_A}(\mathsf{n}^{k+1} a_i X) &\overset{\text{Prop. 5.5.(2)}}{\equiv} \neg J_{F_A}(\mathsf{n}^k a_i X) \\ &\overset{\text{I.H.}}{\equiv} \neg I_{F_A}(\mathsf{n}^k a_i)(\ulcorner J_{F_A}(X)\urcorner) \\ &\overset{\text{Prop. 5.5.(1)}}{\equiv} I_{F_A}(\mathsf{n}^{k+1} a_i)(\ulcorner J_{F_A}(X)\urcorner). \end{aligned}$$

□

Proposition 5.7. *Let $H \in \mathtt{Pred}_{F_A}$ and $K \in \Sigma^*_{F_A}$.*

1. *If $K \notin \mathtt{Pred}_{F_A}$, then $J_{F_A}(\mathsf{r}HK) \equiv \bot$.*
2. *If $K \in \mathtt{Pred}_{F_A}$, then $\mathsf{PA} \vdash J_{F_A}(\mathsf{r}HK) \leftrightarrow J_{F_A}(HKK)$.*

Proof. 1. Let $K \notin \mathtt{Pred}_{F_A}$. By Definition 5.4, we have $J_{F_A}(\mathsf{r}HK) \equiv \bot$.

2. Let $K \in \mathtt{Pred}_{F_A}$. We distinguish the following two cases:

Case 1: H contains at least one r.
In this case, we find some $l \geq 0$ and $H' \in \mathtt{Pred}_{F_A}$ such that $H \equiv \mathsf{n}^l \mathsf{r} H'$.

$$\begin{aligned} J_{F_A}(\mathsf{r}\mathsf{n}^l\mathsf{r}H'K) &\overset{\text{Def. 5.4}}{\equiv} I_{F_A}(\mathsf{r}\mathsf{n}^l\mathsf{r}H')(\ulcorner J_{F_A}(K)(v)\urcorner) \\ &\overset{\text{Def. 5.3}}{\equiv} \neg^l \bot \\ &\overset{\text{Def. 5.4}}{\equiv} \neg^l J_{F_A}(\mathsf{r}H'KK) && (KK \in \mathtt{Sent}^+_{F_A}) \\ &\overset{\text{Prop. 5.5.(2)}}{\equiv} J_{F_A}(\mathsf{n}^l\mathsf{r}H'KK). \end{aligned}$$

Case 2: H contains no r.

$$\begin{aligned} J_{F_A}(\mathsf{r}HK) &\overset{\text{Def. 5.4}}{\equiv} I_{F_A}(\mathsf{r}H)(\ulcorner I_{F_A}(K)(v)\urcorner) \\ &\overset{\text{Def. 5.3}}{\equiv} I_{F_A}(H)(d(\ulcorner I_{F_A}(K)(v)\urcorner)) \\ &\overset{\text{Def. of } d}{\leftrightarrow_{\mathsf{PA}}} I_{F_A}(H)(\ulcorner I_{F_A}(K)(\ulcorner I_{F_A}(K)(v)\urcorner)\urcorner) \\ &\overset{\text{Def. 5.4}}{\equiv} I_{F_A}(H)(\ulcorner J_{F_A}(KK)\urcorner) \\ &\overset{\text{Lem. 5.6}}{\equiv} J_{F_A}(HKK). \qquad (KK \in \mathtt{Sent}^{+}_{F_A}) \end{aligned}$$

□

5.2 The $\mathbb{N}$-based Smullyan model $M_{\mathbb{N}}$

In this subsection, we introduce the Smullyan model $M_{\mathbb{N}}$ which is defined by referring to the standard model $\mathbb{N}$ of arithmetic. We prove that $\mathbb{N}$ is actually an nr-Smullyan model. Then, we prove that FPT, T-Tarski^{+}, and G1^{+} for $M_{\mathbb{N}}$ correspond to the Fixed Point Theorem over $\mathbb{N}$, the original Tarski's Undefinability Theorem, and a version of Gödel's First Incompleteness Theorem, respectively.

Definition 5.8 (The Smullyan model $M_{\mathbb{N}}$)**.**

1. We define the function $\Phi_{\mathbb{N}} : \mathtt{Pred}_{F_{\mathsf{nr}}} \to \mathcal{P}(\Sigma^*_{F_{\mathsf{nr}}})$ as follows:
$$\Phi_{\mathbb{N}}(H) := \{X \in \Sigma^*_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(HX)\}.$$
2. The triple $M_{\mathbb{N}} := (F_{\mathsf{nr}}, \Phi_{\mathbb{N}}) = (\Sigma_{\mathsf{nr}}, \mathtt{Pred}_{\mathsf{nr}}, \Phi_{\mathbb{N}})$ is called the $\mathbb{N}$-*based Smullyan model*.

From the definition of $\Phi_{\mathbb{N}}$, we obtain

$$\mathtt{True}_{M_{\mathbb{N}}} = \{S \in \mathtt{Sent}_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(S)\}. \tag{1}$$

Theorem 5.9. *The* $\mathbb{N}$*-based Smullyan model* $M_{\mathbb{N}}$ *is an* nr*-Smullyan model.*

Proof. We mentioned in Proposition 5.2 that $M_{\mathbb{N}}$ is a Smullyan model. So, it suffices to show that $\Phi_{\mathbb{N}}$ satisfies the requirements concerning n and r. Let $H \in F_{\mathsf{nr}}$.

$$\begin{aligned} \Phi_{\mathbb{N}}(\mathsf{n}H) &= \{X \in \Sigma^*_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(\mathsf{n}HX)\} \\ &= \{X \in \Sigma^*_{F_{\mathsf{nr}}} \mid \mathbb{N} \models \neg J_{F_{\mathsf{nr}}}(HX)\} \qquad \text{(by Proposition 5.5.(2))} \\ &= \Sigma^*_{F_{\mathsf{nr}}} \setminus \{X \in \Sigma^*_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(HX)\} \\ &= \Sigma^*_{F_{\mathsf{nr}}} \setminus \Phi_{\mathbb{N}}(H). \end{aligned}$$

$$\begin{aligned} \Phi_{\mathbb{N}}(\mathsf{r}H) &= \{X \in \Sigma^*_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(\mathsf{r}HX)\} \\ &= \{K \in \mathtt{Pred}_{F_{\mathsf{nr}}} \mid \mathbb{N} \models J_{F_{\mathsf{nr}}}(HKK)\} \qquad \text{(by Proposition 5.7)} \\ &= \{K \in \mathtt{Pred}_{F_{\mathsf{nr}}} \mid KK \in \Phi_{\mathbb{N}}(H)\}. \qquad \Box \end{aligned}$$

Therefore, by the results we have obtained so far, $M_{\mathbb{N}}$ satisfies FPT, T-Tarski$^+$, and G1$^+$. We show that each of these facts yields a meaningful property in arithmetic. First, we show that the Fixed Point Theorem over $\mathbb{N}$ follows from FPT for $M_{\mathbb{N}}$.

Theorem 5.10 (Fixed Point Theorem over $\mathbb{N}$)**.** *The following statement follows from* FPT *for $M_{\mathbb{N}}$: "For any v-formula $\varphi_i(v)$, there exists an $\mathcal{L}_A$-sentence θ such that $\mathbb{N} \models \theta \leftrightarrow \varphi_i(\ulcorner\theta\urcorner)$".*

Proof. Let $\varphi_i(v)$ be any v-formula. By FPT for $M_{\mathbb{N}}$, we find an $M_{\mathbb{N}}$-fixed point $S \in \mathtt{Sent}^+_{F_{\mathsf{nr}}}$ of $a_i \in \mathtt{Pred}_{F_{\mathsf{nr}}}$. The following equivalences show that the $\mathcal{L}_A$-sentence $J_{F_{\mathsf{nr}}}(S)$ is a fixed point of $\varphi_i(v)$ over $\mathbb{N}$.

$$\begin{aligned}
\mathbb{N} \models J_{F_{\mathsf{nr}}}(S) &\overset{(1)}{\iff} S \in \mathtt{True}_{M_{\mathbb{N}}} \iff M_{\mathbb{N}} \models S \\
&\overset{\text{FPT}}{\iff} M_{\mathbb{N}} \models a_i S \iff a_i S \in \mathtt{True}_{M_{\mathbb{N}}} \\
&\overset{(1)}{\iff} \mathbb{N} \models J_{F_{\mathsf{nr}}}(a_i S) \\
&\overset{\text{Lem. 5.6}}{\iff} \mathbb{N} \models I_{F_{\mathsf{nr}}}(a_i)(\ulcorner J_{F_{\mathsf{nr}}}(S)\urcorner) \qquad (S \in \mathtt{Sent}^+_{F_{\mathsf{nr}}}) \\
&\overset{\text{Def. 5.3}}{\iff} \mathbb{N} \models \varphi_i(\ulcorner J_{F_{\mathsf{nr}}}(S)\urcorner).
\end{aligned}$$

□

Second, we show that T-Tarski$^+$ for $M_{\mathbb{N}}$ corresponds to the original version of Tarski's Undefinability Theorem.

Theorem 5.11 (Tarski's Undefinability Theorem)**.** *The following statement follows from* T-Tarski$^+$ *for $M_{\mathbb{N}}$: "The set TA of all $\mathcal{L}_A$-sentences true in $\mathbb{N}$ is not definable in $\mathbb{N}$".*

Proof. Let $\varphi_i(v)$ be any v-formula. By T-Tarski$^+$ for $M_{\mathbb{N}}$, we have that $\mathtt{True}^+_{M_{\mathbb{N}}} \neq \Phi_{\mathbb{N}}(a_i) \cap \mathtt{Sent}^+_{F_{\mathsf{nr}}}$. That is, there exists $S \in \mathtt{Sent}^+_{F_{\mathsf{nr}}}$ such that $S \in \mathtt{True}_{M_{\mathbb{N}}} \iff S \notin \Phi_{\mathbb{N}}(a_i)$. Then,

$$\begin{aligned}
\mathbb{N} \models J_{F_{\mathsf{nr}}}(S) &\overset{(1)}{\iff} S \in \mathtt{True}_{M_{\mathbb{N}}} \iff S \notin \Phi_{\mathbb{N}}(a_i) \iff a_i S \notin \mathtt{True}_{M_{\mathbb{N}}} \\
&\overset{(1)}{\iff} \mathbb{N} \not\models J_{F_{\mathsf{nr}}}(a_i S) \overset{\text{Lem. 5.6}}{\iff} \mathbb{N} \not\models I_{F_{\mathsf{nr}}}(a_i)(\ulcorner J_{F_{\mathsf{nr}}}(S)\urcorner) \\
&\qquad\qquad (S \in \mathtt{Sent}^+_{F_{\mathsf{nr}}}) \\
&\overset{\text{Def. 5.3}}{\iff} \mathbb{N} \not\models \varphi_i(\ulcorner J_{F_{\mathsf{nr}}}(S)\urcorner).
\end{aligned}$$

These equivalences show that $\varphi_i(v)$ does not define TA. □

Finally, we show that G1$^+$ for $M_{\mathbb{N}}$ corresponds to a version of Gödel's First Incompleteness Theorem with respect to arithmetically definable sound theories (cf. Kikuchi and Kurahashi [1] and Salehi and Seraji [5]).

Theorem 5.12 (A version of Gödel's First Incompleteness Theorem)**.** *The following statement follows from* G1$^+$ *for $M_{\mathbb{N}}$: "Every arithmetically definable sound $\mathcal{L}_A$-theory is incomplete".*

Proof. Let T be any arithmetically definable sound $\mathcal{L}_A$-theory. We find an $\mathcal{L}_A$-formula $\tau(x)$ defining (the set of all Gödel numbers of) T in $\mathbb{N}$. As in the usual proof of Gödel's incompleteness theorems, we can construct a provability predicate $\mathrm{Pr}_T(v)$ of T by using $\tau(v)$, that is, for any $\mathcal{L}_A$-formula ψ, we have $T \vdash \psi \iff \mathbb{N} \models \mathrm{Pr}_T(\ulcorner \psi \urcorner)$. Let $i \in \Gamma$ be such that $\varphi_i(v) \equiv \mathrm{Pr}_T(v)$.

We shall show $\Phi_{\mathbb{N}}(a_i) \cap \mathtt{Sent}^+_{F_{\mathsf{nr}}} \subseteq \mathtt{True}^+_{M_{\mathbb{N}}}$. Let $S_0 \in \Phi_{\mathbb{N}}(a_i) \cap \mathtt{Sent}^+_{F_{\mathsf{nr}}}$. Then, $a_i S_0 \in \mathtt{True}_{M_{\mathbb{N}}}$. As in the proof of Theorem 5.10, we have $\mathbb{N} \models \varphi_i(\ulcorner J_{F_{\mathsf{nr}}}(S_0) \urcorner)$. Hence, $T \vdash J_{F_{\mathsf{nr}}}(S_0)$ because $\varphi_i(v)$ is identical to $\mathrm{Pr}_T(v)$. Since T is sound, we have $\mathbb{N} \models J_{F_{\mathsf{nr}}}(S_0)$. Thus, by (1), we conclude $S_0 \in \mathtt{True}^+_{M_{\mathbb{N}}}$.

By G1$^+$ for $M_{\mathbb{N}}$, we get $S \in \mathtt{Sent}^+_{F_{\mathsf{nr}}}$ such that $S \notin \Phi_{\mathbb{N}}(a_i)$ and $\mathsf{n}S \notin \Phi_{\mathbb{N}}(a_i)$. Since $a_i S$ and $a_i \mathsf{n} S$ are not in $\mathtt{True}_{M_{\mathbb{N}}}$, we have $\mathbb{N} \not\models \varphi_i(\ulcorner J_{F_{\mathsf{nr}}}(S_0) \urcorner)$ and $\mathbb{N} \not\models \varphi_i(\ulcorner J_{F_{\mathsf{nr}}}(\mathsf{n}S_0) \urcorner)$. By Proposition 5.5, we also have $\mathbb{N} \not\models \varphi_i(\ulcorner \neg J_{F_{\mathsf{nr}}}(S_0) \urcorner)$. We conclude that $T \nvdash J_{F_{\mathsf{nr}}}(S_0)$ and $T \nvdash \neg J_{F_{\mathsf{nr}}}(S_0)$. ☐

5.3 PA-based Smullyan model M_{PA}

In this subsection, we introduce the Smullyan model M_{PA} which is defined by referring to Peano arithmetic PA. We show that M_{PA} is a witness of the non-implication F-Tarski$^+$ & T-Tarski $\overset{\mathsf{r}}{\nRightarrow}$ T-Tarski$^+$ (cf. Proposition 4.7).

Definition 5.13 (The Smullyan model M_{PA})**.**

1. We define the function $\Phi_{\mathsf{PA}} : \mathtt{Pred}_{F_{\mathsf{r}}} \to \mathcal{P}(\Sigma^*_{F_{\mathsf{r}}})$ as follows:

$$\Phi_{\mathsf{PA}}(H) := \{X \in \Sigma^*_{F_{\mathsf{r}}} \mid \mathsf{PA} \vdash J_{F_{\mathsf{r}}}(HX)\}$$

2. The triple $M_{\mathsf{PA}} := (F_{\mathsf{r}}, \Phi_{\mathsf{PA}}) = (\Sigma_{F_{\mathsf{r}}}, \mathtt{Pred}_{F_{\mathsf{r}}}, \Phi_{\mathsf{PA}})$ is called the PA-*based Smullyan model.*

From the definition of Φ_{PA}, we obtain

$$\mathtt{True}_{M_{\mathsf{PA}}} = \{S \in \mathtt{Sent}_{F_{\mathsf{r}}} \mid \mathsf{PA} \vdash J_{F_{\mathsf{r}}}(S)\}. \tag{2}$$

The following theorem is proved as in the r-part of the proof of Theorem 5.9.

Theorem 5.14. *The* PA*-based Smullyan model* M_{PA} *is an* r*-Smullyan model.*

So, by Theorem 2.7, M_{PA} satisfies FPT. Also by Proposition 3.2, M_{PA} satisfies F-Tarski$^+$. The following theorem stating that FPT for M_{PA} corresponds to a weak version of the Fixed Point Theorem over PA is also proved as in the proof of Theorem 5.10.

Theorem 5.15 (A weak version of the Fixed Point Theorem over PA)**.** *The following statement follows from* FPT *for* M_{PA}*: "For every v-formula* $\varphi_i(v)$*, there exists an* $\mathcal{L}_A$*-sentence* θ *such that* $\mathsf{PA} \vdash \theta \iff \mathsf{PA} \vdash \varphi_i(\ulcorner \theta \urcorner)$*".*

The Fixed Point Theorem over PA used in the classic proof of the incompleteness theorems is of the form $\mathsf{PA} \vdash \theta \leftrightarrow \varphi_i(\ulcorner \theta \urcorner)$, and the statement above is indeed weak. However, this weak version of the Fixed Point Theorem is sufficient for a proof of the incompleteness of PA as follows: Let $\varphi_i(v)$ be a Σ_1 provability predicate of PA, that is, for any $\mathcal{L}_A$-formula ψ, we have $\mathsf{PA} \vdash \psi \iff \mathsf{PA} \vdash \varphi_i(\ulcorner \psi \urcorner)$. By Theorem 5.15, there exists an $\mathcal{L}_A$-sentence θ such that $\mathsf{PA} \vdash \theta \iff \mathsf{PA} \vdash \neg\varphi_i(\ulcorner \theta \urcorner)$. If $\mathsf{PA} \vdash \theta$, then $\mathsf{PA} \vdash \neg\varphi_i(\ulcorner \theta \urcorner)$ and $\mathsf{PA} \vdash \varphi_i(\ulcorner \theta \urcorner)$. This contradicts the consistency of PA. Hence, we have $\mathsf{PA} \nvdash \theta$. It follows that $\mathsf{PA} \nvdash \varphi_i(\ulcorner \theta \urcorner)$ and $\mathsf{PA} \nvdash \neg\varphi_i(\ulcorner \theta \urcorner)$. Therefore, PA is incomplete. This weak version of the Fixed Point Theorem was discussed in Moschovakis [2, Theorem 5.1] (see also Salehi [3, Remark 3.5]).

Finally, we prove that the PA-based Smullyan model M_{PA} is also a witness of the non-implication F-Tarski$^+$ & T-Tarski $\overset{\mathsf{r}}{\not\Rightarrow}$ T-Tarski$^+$. We already mentioned that M_{PA} satisfies F-Tarski$^+$.

Theorem 5.16.

1. *M_{PA} does not satisfy* T-Tarski$^+$.

2. *M_{PA} satisfies* T-Tarski.

Proof. 1. Let $\varphi_i(v)$ be a provability predicate of PA. We shall prove $\texttt{True}^+_{M_{\mathsf{PA}}} = \Phi_{\mathsf{PA}}(a_i) \cap \texttt{Sent}^+_{F_{\mathsf{r}}}$, which shows that M_{PA} does not satisfy T-Tarski$^+$. Let $S \in \texttt{Sent}^+_{F_{\mathsf{r}}}$.

$$\begin{aligned} S \in \texttt{True}_{M_{\mathsf{PA}}} &\iff \mathsf{PA} \vdash J_{F_{\mathsf{r}}}(S) \iff \mathsf{PA} \vdash \varphi_i(\ulcorner J_{F_{\mathsf{r}}}(S) \urcorner) \\ &\iff \mathsf{PA} \vdash J_{F_{\mathsf{r}}}(a_i S) \iff S \in \Phi_{\mathsf{PA}}(a_i). \end{aligned}$$

2. Suppose, towards a contradiction, that $H \in \texttt{Pred}_{F_{\mathsf{r}}}$ names $\texttt{True}_{M_{\mathsf{PA}}}$. Let φ_i be an $\mathcal{L}_A$-sentence such that $\mathsf{PA} \vdash \varphi_i$. Then, we have $I_{F_{\mathsf{r}}}(a_i) \equiv J_{F_{\mathsf{r}}}(a_i) \equiv I_{F_{\mathsf{r}}}(ra_i) \equiv \varphi_i$ because φ_i is a sentence. On the other hand, $J_{F_{\mathsf{r}}}(ra_i) \equiv J_{F_{\mathsf{r}}}(ra_i\epsilon) \equiv \bot$.

Since $\mathsf{PA} \vdash \varphi_i$, we get $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(a_i)$, and hence $a_i \in \texttt{True}_{M_{\mathsf{PA}}}$. By the supposition, we have $a_i \in \Phi_{\mathsf{PA}}(H)$, and hence $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(Ha_i)$. Since $a_i \in \texttt{Pred}_{F_{\mathsf{r}}}$, we have $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(H)(\ulcorner I_{F_{\mathsf{r}}}(a_i) \urcorner)$ and so $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(H)(\ulcorner \varphi_i \urcorner)$. It follows that $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(H)(\ulcorner I_{F_{\mathsf{r}}}(ra_i) \urcorner)$. Then, $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(Hra_i)$ and thus $ra_i \in \Phi_{\mathsf{PA}}(H)$. By the supposition, we have $ra_i \in \texttt{True}_{M_{\mathsf{PA}}}$. Therefore, $\mathsf{PA} \vdash J_{F_{\mathsf{r}}}(ra_i)$ and this means $\mathsf{PA} \vdash \bot$. This is a contradiction. □

6 Concluding remarks

In this paper, we focused on Smullyan's paper "Truth and Provability" (2013, *The Mathematical Intelligencer*). We introduced the notion of Smullyan models in Section 2. In order to understand deeply the content of Smullyan's paper, we have mainly studied this notion throughout the present paper. In Section 3, we introduced several properties of Smullyan models and studied the equivalences between some of these properties. These equivalences are summarized in Figure 1. In particular, these equivalences show that it is sufficient to consider the four properties T-Tarski, F-Tarski, T-Tarski$^+$, and F-Tarski$^+$ when dealing with the properties we introduced. In Section 4, we provided several Smullyan models which are witnesses of non-implications between some combinations of these four properties. As a consequence of these non-implications, no more arrows can be added into Figure 1. Finally in Section 5, we introduced the nr-Smullyan model $M_{\mathbb{N}}$ and the r-Smullyan model M_{PA} based on $\mathbb{N}$ and PA, respectively. We proved that T-Tarski$^+$ for $M_{\mathbb{N}}$ corresponds to the original Tarski's Undefinability Theorem. Also, we showed that M_{PA} is a witness of the non-implication F-Tarski$^+$ & T-Tarski $\overset{\mathsf{r}}{\not\Rightarrow}$ T-Tarski$^+$.

Through these studies, we have found that Smullyan's framework is simple but powerful, and has a fertile structure. In particular, it may be said that the essence of the proof of Tarski's Undefinability Theorem is reasonably explained within this framework as indicated in Theorem 5.11. However, this framework seems too simple for a more intricate discussion of the First Incompleteness Theorem. In his 2013 book [10], Smullyan introduced an extended framework that enables the analysis of Rosser's First Incompleteness Theorem. It will be interesting to explore what extensions, including this one, can make sense for the analysis of the First Incompleteness Theorem in a future work.

Acknowledgements

We would like to thank Naoyuki Hatanaka and Fumihisa Ikema for their comments during the early stages of the study. We would also like to thank the anonymous referee for suggesting some improvements. The first author was supported by JSPS KAKENHI Grant Number JP23K03200.

References

[1] Makoto Kikuchi and Taishi Kurahashi. Generalizations of Gödel's incompleteness theorems for Σ_n-definable theories of arithmetic. *The Review of Symbolic Logic*, 10(4):603–616, 2017.

[2] Yiannis N. Moschovakis. Kleene's amazing second recursion theorem. *The Bulletin of Symbolic Logic*, 16(2):189–239, 2010.

[3] Saeed Salehi. Tarski's undefinability theorem and the diagonal lemma. *Logic Journal of the IGPL. Interest Group in Pure and Applied Logics*, 30(3):489–498, 2022.

[4] Saeed Salehi. A reunion of Gödel, Tarski, Carnap and Rosser. *Journal of Logic and Computation*, 34(6):1172–1178, 2024.

[5] Saeed Salehi and Payam Seraji. Gödel-Rosser's incompleteness theorem, generalized and optimized for definable theories. *Journal of Logic and Computation*, 27(5):1391–1397, 2017.

[6] Raymond M. Smullyan. *What is the name of this book? The riddle of Dracula and other logical puzzles.* Prentice-Hall, 1978.

[7] Raymond M. Smullyan. *The Lady or the Tiger?: and Other Logic Puzzles.* Knopf, 1982.

[8] Raymond M. Smullyan. *Forever undecided. A puzzle guide to Gödel.* Oxford University Press, 1988.

[9] Raymond M. Smullyan. *Logical labyrinths.* Wellesley, MA: A K Peters, 2009.

[10] Raymond M. Smullyan. *Gödelian Puzzle Book: Puzzles, Paradoxes and Proofs.* Dover Publications Inc., 2013.

[11] Raymond M. Smullyan. Truth and provability. *The Mathematical Intelligencer*, 35(1):21–24, 2013.