Rosser provability and normal modal logics

Taishi Kurahashi

Abstract

In this paper, we investigate Rosser provability predicates whose provability logics are normal modal logics. First, we prove that there exists a Rosser provability predicate whose provability logic is exactly the normal modal logic $KD$. Secondly, we introduce a new normal modal logic $KDR$ which is a proper extension of $KD$, and prove that there exists a Rosser provability predicate whose provability logic includes $KDR$.

1 Introduction

In the paper [9], we raised the problem of the existence of a $\Sigma_2$ representation of each theory $T$ such that the provability logic of the provability predicate of $T$ constructed from the representation is exactly the modal logic $KD = K + \neg \Box \bot$. This problem has not been settled yet. Here we consider the following more general question: Is there a provability predicate whose provability logic is exactly $KD$? In this paper, we give an affirmative answer to this problem by considering Rosser provability predicates.

Let $T$ be any consistent recursively enumerable extension of Peano Arithmetic $\text{PA}$. We say a formula $Pr_T(x)$ is a provability predicate of $T$ if it weakly represents the set of all theorems of $T$ in $\text{PA}$, that is, for any natural number $n$, $\text{PA} \vdash Pr_T(n)$ if and only if $n$ is the Gödel number of some theorem of $T$. An arithmetical interpretation based on $Pr_T(x)$ is a mapping $f$ from modal formulas to sentences of arithmetic such that $f$ commutes with every propositional connective and $f$ maps $\Box$ to $Pr_T(x)$. Let $\text{PL}(Pr_T)$ be the set of all modal formulas $A$ such that $T \vdash f(A)$ for any arithmetical interpretation $f$ based on $Pr_T(x)$. This set is called the provability logic of $Pr_T(x)$. Solovay [18] proved that for each standard $\Sigma_1$ provability predicate $Pr_T(x)$ of $T$, if $T$ is $\Sigma_1$-sound, then the provability logic of $Pr_T(x)$ is equal to the modal logic $GL$. This is Solovay’s arithmetical completeness theorem.

On the other hand, Feferman [4] found a $\Pi_1$ representation of a theory $T$ such that the consistency statement defined by using the
provability predicate $\text{Pr}_F^F(x)$ constructed from the representation is provable in $\text{PA}$. The provability logic $\text{PL}(\text{Pr}_F^F)$ of Feferman’s predicate includes the modal logic $\text{KD}$, and it is completely different from $\text{GL}$.

The problem of exact axiomatization of $\text{PL}(\text{Pr}_F^F)$ was studied by Montagna [11] and Visser [19], but it has not been settled yet. Shavrukov [10] found a Feferman-like $\Sigma_2$ provability predicate whose provability logic is exactly the modal logic $\text{KD} + \Box p \rightarrow \Box(\Box q \rightarrow q) \lor \Box p$.

Rosser provability predicate $\text{Pr}_R^F(x)$ was essentially introduced by Rosser [13] to improve Gödel’s first incompleteness theorem. It is well-known that the consistency statement defined by $\text{Pr}_R^F(x)$ is provable in $\text{PA}$. Then by the proof of Gödel’s second incompleteness theorem, at least one of the principles $(\text{K})$: $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ and $(\text{4})$: $\Box p \rightarrow \Box \Box p$ is invalid for each Rosser provability predicate (Whether the principle $(\text{K})$ is valid for Rosser provability predicates was asked by Kreisel and Takeuti [7]). Actually, Guaspari and Solovay [5] and Arai [7] showed that whether $(\text{K})$ or $(\text{4})$ is invalid for $\text{Pr}_R^F(x)$ is dependent on the choice of $\text{Pr}_R^F(x)$. More precisely, by using the modal logical result of Guaspari and Solovay, it can be shown that there exists a Rosser provability predicate for which both of these principles are not valid. Also Arai proved the existence of a Rosser provability predicate satisfying $(\text{K})$ and a Rosser provability predicate satisfying $(\text{4})$.

Modal logical investigations of Rosser provability predicates were initiated by Guaspari and Solovay, and continued by Visser [19], Shavrukov [15] and others. In particular, Shavrukov introduced the bimodal logic $\text{GR}$ for usual provability and Rosser provability, and proved the arithmetical completeness theorem for $\text{GR}$. Although Shavrukov’s arithmetically complete logic $\text{GR}$ does not contain $(\text{K})$ for the modality of Rosser provability as an axiom, it is worth considering $(\text{K})$ for Rosser provability from modal logical viewpoint. That is, it is easy to show that the provability logic $\text{PL}(\text{Pr}_R^F)$ is a normal modal logic if and only if $(\text{K})$ is valid for $\text{Pr}_R^F(x)$. If $\text{PL}(\text{Pr}_R^F)$ is normal, then $\text{PL}(\text{Pr}_R^F)$ includes $\text{KD}$.

In this paper, we investigate Rosser provability predicates whose provability logics are normal. In Section 3, we give an affirmative answer to the problem raised in the first paragraph of this section, that is, we prove that there exists a Rosser provability predicate $\text{Pr}_R^T(x)$ of $T$ such that $\text{PL}(\text{Pr}_R^T)$ is exactly $\text{KD}$. In Section 4 we introduce and study a new normal modal logic $\text{KDR} = \text{KD} + \Box \neg p \rightarrow \Box \neg \Box p$. In particular, we prove that there exists a Rosser provability predicate $\text{Pr}_R^T(x)$ of $T$ such that $\text{KDR} \subseteq \text{PL}(\text{Pr}_R^T)$. Thus we obtain a Rosser provability predicate whose provability logic is a proper extension of $\text{KD}$. Whether there exists a Rosser provability predicate whose provability logic is exactly $\text{KDR}$ is still open.
2 Preliminaries

The axioms of the modal logic K are all propositional tautologies in the language of propositional modal logic and the formula \( \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q) \). The inference rules for K are modus ponens, necessitation and substitution. Each modal logic \( L \) is identified with the set of all theorems of \( L \). We say a modal logic \( L \) is normal if \( K \subseteq L \) and \( L \) is closed under modus ponens, necessitation and substitution. For any modal logic \( L \) and modal formula \( A \), let \( L + A \) denote the least normal modal logic whose axioms are those of \( L \) and the formula \( A \). Several normal modal logics are obtained by adding axioms to \( K \) as follows (see [2, 3] for more details):

1. \( KD = K + \neg \Box \bot \).
2. \( T = K + \Box p \rightarrow p \).
3. \( KD4 = KD + \Box p \rightarrow \Box \Box p \).
4. \( KD5 = KD + \neg \Box p \rightarrow \Box \neg \Box p \).
5. \( GL = K + \Box (\Box p \rightarrow p) \rightarrow \Box p \).

A Kripke frame is a tuple \((W, \prec)\) where \( W \) is a nonempty set and \( \prec \) is a binary relation on \( W \). A Kripke model is a tuple \( M = (W, \prec, \models) \) where \((W, \prec)\) is a Kripke frame, and \( \models \) is a binary relation between \( W \) and the set of all modal formulas satisfying the usual conditions for satisfaction and the following condition: \( x \models \Box A \) if and only if for all \( y \in W \), \( y \models A \) if \( x \prec y \). We say a modal formula \( A \) is valid in a Kripke model \( M = (W, \prec, \models) \) if for all \( w \in W \), \( w \models A \). We say that \( M \) is finite if \( W \) is finite. Also we say that \( M \) is serial if for any \( x \in W \), there exists \( y \in W \) such that \( x \prec y \).

It is known that the modal logic \( KD \) is sound and complete with respect to the class of all finite serial Kripke models. Moreover, the following theorem holds.

**Theorem 2.1** (Kripke completeness theorem for \( KD \) (see [12])). *For each modal formula \( A \) which is not provable in \( KD \), we can primitive recursively find a finite serial Kripke model in which \( A \) is not valid.*

Throughout this paper, we assume that \( T \) always denotes a recursively enumerable consistent extension of Peano Arithmetic \( \text{PA} \) in the language \( L_A \) of first-order arithmetic. We also assume that \( L_A \) contains function symbols for all primitive recursive functions. Let \( \omega \) be the set of all natural numbers. For each \( n \in \omega \), the numeral for \( n \) is denoted by \( \ulcorner n \urcorner \). We fix a natural G"{o}del numbering such that 0 is not a G"{o}del number of any object, and that the G"{o}del number of \( \ulcorner \sigma \urcorner \) is larger than \( n \). For each \( L_A \)-formula \( \varphi \), let \( \ulcorner \varphi \urcorner \) be the numeral for the G"{o}del number of \( \varphi \). Let \( \{ \ulcorner \varphi \urcorner \}_{k \in \omega} \) be the repetition-free effective sequence of
all formulas arranged in ascending order of whose Gödel numbers. We
assume that if \( \varphi_k \) is a subformula of \( \varphi_l \), then \( k \leq l \).

We say a formula \( \text{Pr}_T(x) \) is a provability predicate of \( T \) if for any \( n \in \omega \), \( \text{PA} \vdash \text{Pr}_T(\overline{n}) \) if and only if \( n \) is the Gödel number of some \( T \)-provable formula. We fix a \( \Delta_1 \) (PA) formula \( \text{Proof}_T(x, y) \) which is a natural formalization of the relation “\( y \) is a \( T \)-proof of a formula \( x \)” with the usual adequate properties. Let \( \text{Prov}_T(x) \) be the formula \( \exists y \text{Proof}_T(x, y) \).

Then \( \text{Prov}_T(x) \) is a provability predicate of \( T \) satisfying several conditions such as \( \text{PA} \vdash \text{Prov}_T(\gamma \varphi \rightarrow \psi) \rightarrow (\text{Prov}_T(\gamma \varphi) \rightarrow \text{Prov}_T(\gamma \psi)) \) and \( \text{PA} \vdash \text{Prov}_T(\gamma \varphi) \rightarrow \text{Prov}_T(\gamma \text{Prov}_T(\gamma \varphi)) \). Let \( \text{Con}_T \) be the sentence \( \neg\text{Prov}_T(\gamma 0 = 1) \) expressing the consistency of \( T \). We say a formula \( \text{Prf}_T(x, y) \) is a provability predicate of \( T \) if \( \text{Prf}_T(x, y) \) satisfies the following conditions:

1. \( \text{Prf}_T(x, y) \) is \( \Delta_1 \) (PA),
2. \( \text{PA} \vdash \forall x(\text{Prov}_T(x) \leftrightarrow \exists y \text{Prf}_T(x, y)) \),
3. for any \( n \in \omega \) and formula \( \varphi \), \( \text{N} \models \text{Proof}_T(\gamma \varphi, \overline{n}) \leftrightarrow \text{Prf}_T(\gamma \varphi, \overline{n}), \)
4. \( \text{PA} \vdash \forall x \forall x' \forall y (\text{Prf}_T(x, y) \land \text{Prf}_T(x', y) \rightarrow x = x') \).

Here \( \text{N} \) is the standard model of arithmetic. The last clause means that our proof predicates are single conclusion ones. For each proof predicate \( \text{Prf}_T(x, y) \) of \( T \), the \( \Sigma_1 \) formula

\[
\exists y (\text{Prf}_T(x, y) \land \forall z \leq y \neg \text{Prf}_T(\neg(x), z))
\]

is said to be the Rosser provability predicate of \( \text{Prf}_T(x, y) \) or a Rosser provability predicate of \( T \), where \( \neg(x) \) is a term corresponding to a primitive recursive function calculating the Gödel number of \( \neg \varphi \) from the Gödel number of a formula \( \varphi \). Each Rosser provability predicate of \( T \) is a provability predicate of \( T \). Also the following proposition holds.

**Proposition 2.2.** Let \( \text{Prf}^R_T(x) \) be a Rosser provability predicate of \( \text{Prf}_T(x, y) \) or a Rosser provability predicate of \( T \) and \( \varphi \) be any formula. If \( T \vdash \neg \varphi \), then \( T \vdash \neg \text{Prf}^R_T(\gamma \varphi) \).

As a consequence of Proposition \( \text{2.2} \) we have \( T \vdash \neg \text{Prf}^R_T(\gamma 0 = 1) \).

Let \( \text{Prf}_T(x) \) be any provability predicate of \( T \). A mapping \( f \) from the set of all modal formulas to the set of all \( \mathcal{L}_A \)-sentences is said to be an arithmetical interpretation based on \( \text{Prf}_T(x) \) if \( f \) satisfies the following conditions:

1. \( f(\bot) = 0 = 1 \),
2. \( f \) commutes with each propositional connective,
3. \( f(\square A) = \text{Prf}_T(\gamma f(A)) \).

The set \( \text{PL}(\text{Prf}_T) = \{ A : A \text{ is a modal formula and for all arithmetical interpretations } f \text{ based on } \text{Prf}_T(x), T \vdash f(A) \} \) is called the provability logic of \( \text{Prf}_T(x) \). One of the major achievements of the investigation of provability logics is Solovay’s arithmetical completeness theorem (see \( \text{2} \) \( \text{17} \) \( \text{18} \)).
Theorem 2.3 (Solovay’s arithmetical completeness theorem for GL).
If $T$ is $\Sigma_1$-sound, then $\text{PL}(\text{Prov}_T) = \text{GL}$.

Provability logics of nonstandard provability predicates have been also studied by many authors. Feferman [4] found a nonstandard $\Sigma_2$ provability predicate $\text{Pr}_{T}^{F}(x)$ such that $\text{KD} \subseteq \text{PL}(\text{Pr}_{T}^{F})$ (see also [11, 19]). Shavrukov [16] found a Feferman-like $\Sigma_2$ provability predicate whose provability logic is exactly $\text{KD} + □p \rightarrow □(□q \rightarrow q) \lor □p$. Also it was proved in [8, 9] that for each $L \in \{\text{K} \cup \{\text{K} + □(□n^p \rightarrow p) \rightarrow □p : n \geq 2\}\}$, there exists a $\Sigma_2$ provability predicate $\text{Pr}_{T}(x)$ of $T$ such that $\text{PL}(\text{Pr}_{T})$ is precisely $L$ (The modal logic $\text{K} + □(□n^p \rightarrow p) \rightarrow □p$ for $n \geq 2$ was introduced by Sacchetti [14]).

In this paper, we are interested in the provability logics $\text{PL}(\text{Pr}_{T}^{R})$ of Rosser provability predicates $\text{Pr}_{T}^{R}(x)$. In particular, we study the situation where $\text{PL}(\text{Pr}_{T}^{R})$ is a normal modal logic. We introduce the following terminology.

Definition 2.4. A Rosser provability predicate $\text{Pr}_{T}^{R}(x)$ of $T$ is normal if $\text{PL}(\text{Pr}_{T}^{R})$ is a normal modal logic.

It is easy to show the following proposition.

Proposition 2.5. For any Rosser provability predicate $\text{Pr}_{T}^{R}(x)$ of $T$, the following are equivalent:
1. $\text{Pr}_{T}^{R}(x)$ is normal.
2. $\text{KD} \subseteq \text{PL}(\text{Pr}_{T}^{R})$.
3. $T \vdash \text{Pr}_{T}^{R}(\lnot \varphi \rightarrow \psi) \rightarrow (\text{Pr}_{T}^{R}(\lnot \varphi) \rightarrow \text{Pr}_{T}^{R}(\lnot \psi))$ for any sentences $\varphi$ and $\psi$.

We can define a Rosser provability predicate which is not normal by using modal logical results of Guaspari and Solovay [5] or Shavrukov [15]. On the other hand, Arai [1] defined a normal Rosser provability predicate (This was also mentioned by Shavrukov [15]). Thus whether $\text{Pr}_{T}^{R}(x)$ is normal or not is dependent on the choice of $\text{Pr}_{T}^{R}(x)$. Model theoretic properties of normal Rosser provability predicates were investigated by Kikuchi and Kurahashi [6].

We say an $\mathcal{L}_A$-formula $\varphi$ is propositionally atomic if it is not a Boolean combination of proper subformulas of $\varphi$. We prepare a new propositional variable $p_\varphi$ for each propositionally atomic formula $\varphi$. Then there exists a primitive recursive injection $I$ from $\mathcal{L}_A$-formulas to propositional formulas satisfying the following conditions:
1. $I(\varphi) \equiv p_\varphi$ for each propositionally atomic $\varphi$.
2. $I$ commutes with every propositional connective.

Let $X$ be any finite set of $\mathcal{L}_A$-formulas. We say $X$ is propositionally satisfiable if the set $I(X) = \{I(\varphi) : \varphi \in X\}$ of propositional formulas
is satisfiable. An \( \mathcal{L}_A \)-formula \( \psi \) is said to be a **tautological consequence (t.c.)** of \( X \) if \( I(\psi) \) is a tautological consequence of \( I(X) \). The above definitions are formalized in \( PA \), and \( PA \) can prove several facts about them. For instance, \( PA \) proves that “If \( X \cup \{ \varphi \} \) is not propositionally satisfiable for a finite set \( X \) of formulas and a formula \( \varphi \), then \( \neg \varphi \) is a t.c. of \( X \).” Define \( P_{T,n} \) to be the finite set \( \{ \varphi : \mathbb{N} \models \exists y \leq \overline{\text{Proof}}_T(\neg \varphi^+, y) \} \) of formulas. Then \( PA \) proves “If a formula \( \varphi \) is a t.c. of \( P_{T,n} \) for some \( n \), then \( \varphi \) is provable in \( T^n \), and so on.

\[ \]

### 3 The arithmetical completeness theorem for \( KD \)

In this section, we prove that there exists a normal Rosser provability predicate of \( T \) whose provability logic is exactly the modal logic \( KD \).

**Theorem 3.1.** There exists a Rosser provability predicate \( \text{Pr}^R_\varphi(x) \) of \( T \) such that the following conditions hold:

1. (Arithmetical soundness) For any modal formula \( A \), if \( KD \vdash A \), then \( T \vdash f(A) \) for any arithmetical interpretation \( f \) based on \( \text{Pr}^R_\varphi(x) \). In other words, \( \text{Pr}^R_\varphi(x) \) is normal.

2. (Uniform arithmetical completeness) There exists an arithmetical interpretation \( f \) based on \( \text{Pr}^R_\varphi(x) \) such that for any modal formula \( A \), \( KD \vdash A \) if and only if \( T \vdash f(A) \).

Let \( \{A_k\}_{k \in \omega} \) be a primitive recursive enumeration of all modal formulas that are not provable in \( KD \). From each \( A_k \), we can primitive recursively construct a finite serial Kripke model \( M_k = (W_k, \preceq_k, \models_k) \) in which \( A_k \) is not valid by Theorem 2.1. We may assume that \( W_k \) and \( W_l \) for \( k \neq l \) are pairwise disjoint sets of natural numbers and \( \bigcup_{k \in \omega} W_k = \omega \setminus \{0\} \). We define an infinite Kripke model \( M = (W, \preceq, \models) \) which can be primitive recursively represented in \( PA \) as follows:

1. \( W = \omega \setminus \{0\} \),
2. \( x \preceq y \) if and only if for some \( k \in \omega \), \( x, y \in W_k \) and \( x \preceq_k y \),
3. \( x \models p \) if and only if for some \( k \in \omega \), \( x \in W_k \) and \( x \models_k p \).

First, we define a primitive recursive function \( h(x) \) by using the recursion theorem as follows:

- \( h(0) = 0 \),
- \( h(m+1) = \begin{cases} i & \text{if } h(m) = 0 \& i \neq 0 \\ i = \min \{ j \in W : \neg \text{S}(j) \text{ is a t.c. of } P_{T,m} \}, \\ h(m) & \text{if no } i \text{ as above exists.} \end{cases} \)
Here $S(x)$ is the $\Sigma_1$ formula $\exists v(h(v) = x)$. The sentence $S(\overline{j})$ is propositionally atomic because it is an existential sentence. Suppose that $P_{T,m}$ is propositionally satisfiable and $\neg S(\overline{j})$ is a t.c. of $P_{T,m}$. Then there exists a formula $\phi \in P_{T,m}$ containing $S(\overline{j})$ as a subformula. Hence $j \leq m$. Thus $\{j \in W : \neg S(\overline{j})\}$ is a t.c. of $P_{T,m}$ if and only if there exists $j \leq m$ such that $\neg S(\overline{j})$ is a t.c. of $P_{T,m}$. Notice that this equivalence also holds when $P_{T,m}$ is not propositionally satisfiable. From this observation, for each $m$, whether $\{ j \in W : \neg S(\overline{j}) \}$ is empty or not can be primitive recursively determined. This guarantees that $h(x)$ is a primitive recursive function.

By the definition of the function $h(x)$, we obtain the following lemma.

**Lemma 3.2.**

1. PA $\vdash \forall v \exists x (h(v) = x \land x \neq 0 \rightarrow \forall u \geq v h(u) = x)$.
2. PA $\vdash \forall x \forall y (0 < x < y \land S(x) \rightarrow \neg S(y))$.
3. The sentences $\neg \text{Con}_T, \exists x (\text{Prov}_T(\neg \exists S(\overline{x}) \land x \neq 0))$ and $\exists x (S(x) \land x \neq 0)$ are equivalent in PA.
4. For any $i \neq 0$, $T \not\vdash \neg S(\overline{i})$.

**Proof.** 1 is proved by induction in PA. 2 follows from 1 immediately.

3. We prove PA $\vdash \neg \text{Con}_T \rightarrow \exists x (\text{Prov}_T(\neg \exists S(\overline{x}) \land x \neq 0))$ is obvious.

We prove PA $\vdash \exists x (\text{Prov}_T(\neg \exists S(\overline{x}) \land x \neq 0) \rightarrow \exists x (S(x) \land x \neq 0))$. We work in PA: Suppose $\neg S(\overline{i})$ is provable in $T$ for $i \neq 0$. Then $\neg S(\overline{i})$ is contained in $P_{T,m}$ for some $m$. In this case, $\neg S(\overline{i})$ is a t.c. of $P_{T,m}$. Hence $\{ j \in W : \neg S(\overline{j}) \}$ is a t.c. of $P_{T,m}$, is not empty for some $n$. For the least such $n$, $h(n+1) \neq 0$.

We prove PA $\vdash \exists x (S(x) \land x \neq 0) \rightarrow \neg \text{Con}_T$. We reason in PA: Suppose $S(i)$ holds and $i \neq 0$. Then there exists $n$ such that $h(m) = 0$ and $h(m+1) = i$. In this case, $i = \min \{ j \in W : \neg S(\overline{j}) \}$ is a t.c. of $P_{T,m}$, and hence $\neg S(\overline{i})$ is a t.c. of $P_{T,m}$. Then $\neg S(\overline{i})$ is provable in $T$. On the other hand, the sentence $S(\overline{i})$ is provable in $T$ because it is a true $\Sigma_1$ sentence. Therefore $T$ is inconsistent.

4. Suppose $T \vdash \neg \text{Con}_T$ for some $i \neq 0$. Then PA $\vdash \exists x (\text{Prov}_T(\neg \exists S(\overline{x}) \land x \neq 0))$. By 3, PA $\vdash \neg \text{Con}_T$. This contradicts the consistency of $T$. Thus $T \not\vdash \neg S(\overline{i})$.

For each $i \in W$, the set $\{ j \in W : i \prec j \}$ is finite and nonempty because the Kripke model $M$ is a disjoint union of finite serial Kripke models. We can use the sentence $\bigvee_{i \prec j} S(\overline{j})$ which contains at least one disjunct.

We define a PA-provably recursive function $g(x)$ which enumerates all theorems of $T$. The definition of $g(x)$ consists of Procedures 1 and 2. The definition starts with Procedure 1 and the values of $g(0), g(1), \ldots$
are defined by referring to the values of the function $h(x)$ in stages. At the first time of $h(m + 1) \neq 0$, the construction of $g(x)$ is switched to Procedure 2. In Procedure 2, $g(x)$ outputs all formulas in stages.

We start defining the function $g(x)$. In the definition of $g(x)$, we identify each formula with its Gödel number.

Procedure 1.
Stage 1. $m$:
- If $h(m + 1) = 0$, then
  \[
g(m) = \begin{cases} \varphi & \text{if } m \text{ is a proof of } \varphi \text{ in } T, \text{ that is, } \mathsf{Proof}_T(\neg \varphi^\dagger, m) \text{ holds,} \\ 0 & \text{if } m \text{ is not a proof of any formula in } T. \end{cases}
\]
  Go to Stage 1. $(m + 1)$.
- If $h(m + 1) \neq 0$, then go to Procedure 2.

Procedure 2.
Let $m$ be the smallest number such that $h(m + 1) \neq 0$. Let $i = \min\{j \in W : \neg S(j) \text{ is a t.c. of } P_{T,m}\}$ and let $X$ be the finite set $P_{T,m-1} \cup \bigcup_{j<i} S(j)$ of formulas. Then $i = h(m + 1)$. Let \(\{\varphi_k\}_{k \in \omega}\) be the sequence of all formulas introduced in Section 2.

We define the values of $g(m), g(m + 1), \ldots$ and the numbers \(\{t_k\}_{k \in \omega}\) simultaneously in stages. Let $t_0 = 0$.

Stage 2. $k$: We distinguish the following three cases C1, C2 and C3.

**C1** If $\varphi_k$ is a t.c. of $X$, then let $g(m + t_k) = \varphi_k$ and $t_{k+1} = t_k + 1$.

**C2** If $\varphi_k$ is not a t.c. of $X$ and $\neg \varphi_k$ is a t.c. of $X$, then let $g(m + t_k) = \neg \varphi_k$, $g(m + t_k + 1) = \varphi_k$ and $t_{k+1} = t_k + 2$.

**C3** If neither $\varphi_k$ nor $\neg \varphi_k$ is a t.c. of $X$, then for each $0 \leq s \leq m$, let
\[
g(m + t_k + s) = \neg \cdots \neg \varphi_k \text{ and } t_{k+1} = t_k + m + 1.
\]

Go to stage 2. $(k + 1)$.

The definition of $g(x)$ has just been finished. Let $\mathsf{Prf}_g(x, y)$ be the $\Delta_1(\mathsf{PA})$ formula $x = g(y) \land \mathsf{Fml}(x)$, where $\mathsf{Fml}(x)$ is a natural $\Delta_1(\mathsf{PA})$ representation of “$x$ is an $\mathcal{L}_A$-formula”. Also let $\mathsf{Pr}_g(x)$ and $\mathsf{Pr}_g^R(x)$ be the formula $\exists y \mathsf{Prf}_g(x, y)$ and the Rosser provability predicate of $\mathsf{Prf}_g(x, y)$, respectively. Actually, our formula $\mathsf{Prf}_g(x, y)$ is a proof predicate of $T$.

**Lemma 3.3.**

1. $\mathsf{PA} \vdash \forall x(\mathsf{Prov}_T(x) \leftrightarrow \mathsf{Pr}_g(x))$.
2. For any $n \in \omega$ and formula $\varphi$, $\mathbb{N} \models \mathsf{Proof}_T(\neg \varphi^\dagger, n) \leftrightarrow \mathsf{Pr}_g(\neg \varphi^\dagger, n)$.

**Proof.** 1. It is clear that $\neg \exists x(S(x) \land x \neq 0) \rightarrow \forall x(\mathsf{Prov}_T(x) \leftrightarrow \mathsf{Pr}_g(x))$ is proved in $\mathsf{PA}$ by the definition of $g(x)$. Also $\mathsf{PA} \vdash \exists x(S(x) \land x \neq$
0) \rightarrow \forall x(\Pr_g(x) \leftrightarrow \Fml(x)) because each formula \varphi_k is output at Stage 2.k in Procedure 2. Since PA \vdash \neg \text{Con}_P \rightarrow \forall x(\Prov_T(x) \leftrightarrow \Fml(x)) we have PA \vdash \exists x(S(x) \land x \neq 0) \rightarrow \forall x(\Prov_T(x) \leftrightarrow \Pr_g(x)) by Lemma 3.2.3. Thus we obtain PA \vdash \forall x(\Prov_T(x) \leftrightarrow \Pr_g(x))

2. By Lemma 3.2.3, \neg \exists x(S(x) \land x \neq 0) is true in \mathbb{N}. Then \text{Proof}^T(\neg \varphi^n, \pi) and Pr^T_g(\neg \varphi^n, \pi) are equivalent in \mathbb{N} by the definition of g(x).

**Lemma 3.4.** Let \psi be a formula. Then the following statement is provable in PA:

"Let i and m be such that h(m) = 0 and h(m + 1) = i. Let X = Pr_{t,m-1} \cup \{S(j) : j \in \omega, S(j) \not\in X\}.

1. For each j_0 > i, \neg S(j_0) is not a t.c. of X.
2. X is propositionally satisfiable.
3. If a formula \varphi is a t.c. of X, then Pr^R_g(\neg \varphi^n) holds.
4. If \psi is not a t.c. of X, then \neg Pr^R_g(\neg \psi^n) holds.

**Proof.** We reason in PA. Let i and m be such that h(m) = 0 and h(m + 1) = i. In this case, i = \min\{j \in W : \neg S(j) \not\in X\} is a t.c. of Pr_{t,m}.

Let X = Pr_{t,m-1} \cup \{S(j) : j \in \omega, S(j) \not\in X\}.

1. Suppose that \neg S(j_0) is a t.c. of X for some j_0 > i. Then \bigvee_{i < j} S(j) \rightarrow \neg S(j_0) is a t.c. of Pr_{t,m-1}. Since S(j_0) \rightarrow \bigvee_{i < j} S(j) is a tautology, S(j_0) \rightarrow \neg S(j_0) is also a t.c. of Pr_{t,m-1}. This means \neg S(j_0) is a t.c. of Pr_{t,m-1}. Then \{j \in W : \neg S(j) \not\in X\} is a t.c. of Pr_{t,m-1} \not\in \emptyset, and hence h(m) \neq 0. This is a contradiction. Therefore \neg S(j_0) is not a t.c. of X.

2. Since \prec is serial, there exists at least one j_0 such that j_0 > i. Then X \cup \{S(j_0)\} is propositionally satisfiable by 1. It follows that X is propositionally satisfiable.

3. Suppose \varphi is a t.c. of X. For the enumeration \{\varphi_k\}_{k \in \omega} used in the definition of g(x), let \varphi = \varphi_k. Then g(m + t_k) = \varphi by C1. Since X is propositionally satisfiable by 2, \neg \varphi \not\in X. Hence \neg \varphi \not\in Pr_{t,m-1}. Therefore \neg \varphi \not\in \{g(0), \ldots, g(m - 1)\} because the construction of g(x) is switched to Procedure 2 at Stage 1.m. Also \neg \varphi \not\in \varphi_k' for all k' \leq k. Let \varphi_l be a formula obtained by deleting zero or more leading negation symbols \neg from \varphi. Then l \leq k and exactly one of \varphi_l and \neg \varphi_l is a t.c. of X. Thus \neg \varphi is not output at Stage 2.l by the definition of g(x). Therefore \neg \varphi is not output before Stage 2.k. This means that Pr^R_g(\neg \varphi^n) holds.

4. Suppose that \psi is not a t.c. of X. Then \neg \psi \not\in \{g(0), \ldots, g(m - 1)\} since \psi is not contained in X. We distinguish the following two cases.
• Case 1: \(\neg\psi\) is a t.c. of \(X\). Let \(\psi = \varphi_k\). Then \(g(m + t_k) = \neg\psi\) by \textbf{C2}. Also \(\psi \neq \varphi_{k'}\) for all \(k' < k\). If \(\psi\) is of the form \(\neg\varphi_l\) for some \(l < k\), then \(\varphi_l\) is a t.c. of \(X\). Hence \(g(m + t_l) = \varphi_l\) by \textbf{C1}, and \(\psi\) is not output at Stage 2.\(l\). Also let \(\varphi_p\) be a formula obtained by deleting zero or more leading \(\neg\)'s from \(\psi\), then exactly one of \(\varphi_p\) and \(\neg\varphi_p\) is a t.c. of \(X\). Hence \(\varphi_p\) is output by \textbf{C1} or \textbf{C2}, and \(\psi\) is not output at Stage 2.\(p\) if \(p < k\). Therefore \(\psi\) is not output before Stage 2.\(k\). Thus \(\Pr^R_g(\neg\psi^\top)\) does not hold.

• Case 2: \(\neg\psi\) is not a t.c. of \(X\). Let \(\varphi_k\) be a formula obtained by deleting all leading \(\neg\)'s from \(\psi\). Then \(\varphi_k\) does not appear before Stage 2.\(k\). Since neither \(\varphi_k\) nor \(\neg\varphi_k\) is a t.c. of \(X\), \(g(m + t_k + s) = \neg\cdots\neg\varphi_k\) for \(0 \leq s \leq m\) by \textbf{C3}. Let \(n\) be the number of deleted negation symbols from \(\psi\). Notice that \(\neg S(i)\) is not a t.c. of \(P_{T,n}\) by Lemma 3.4.4 (because \(n\) is standard). Hence \(m > n\) holds. Thus for \(s = m - n - 1\), \(g(m + t_k + s) = \neg\psi\) and \(g(m + t_k + s + 1) = \psi\).

Therefore \(\Pr^R_g(\neg\psi^\top)\) does not hold.

\[\square\]

**Lemma 3.5.** Let \(i, k \in W\) and suppose \(i \prec k\).

1. \(\text{PA} \vdash S(i) \rightarrow \Pr^R_g(\neg\bigvee_{i \prec j} S(\overline{j}))\).
2. \(\text{PA} \vdash S(i) \rightarrow \neg\Pr^R_g(\neg S(\overline{k}))\).

**Proof.** Suppose \(i \prec k\). We reason in \(\text{PA} + S(i)\): Let \(m\) be such that \(h(m) = 0\) and \(h(m + 1) = i\). Let \(X = P_{T,m-1} \cup \{\bigvee_{i \prec j} S(\overline{j})\}\).

Since \(\bigvee_{i \prec j} S(\overline{j})\) is a t.c. of \(X\), \(\Pr^R_g(\neg\bigvee_{i \prec j} S(\overline{j}))\) holds by Lemma 3.4.3. Also \(\neg S(\overline{k})\) is not a t.c. of \(X\) by Lemma 3.4.1. Therefore \(\neg\Pr^R_g(\neg S(\overline{k}))\) holds by Lemma 3.4.4 (because \(\neg S(\overline{k})\) is standard).

\[\square\]

**Lemma 3.6.** For any \(\varphi\) and \(\psi\), \(\text{PA} \vdash \Pr^R_g(\neg\varphi \rightarrow \psi^\top) \rightarrow (\Pr^R_g(\neg\varphi^\top) \rightarrow \Pr^R_g(\neg\psi^\top))\).

**Proof.** Since \(\text{PA} + \text{Con}_T \vdash \text{Prov}_T(\neg\varphi^\top) \rightarrow \neg\text{Prov}_T(\neg\neg\varphi^\top)\), we have \(\text{PA} + \text{Con}_T \vdash \Pr^R_g(\neg\varphi^\top) \rightarrow \neg\Pr^R_g(\neg\neg\varphi^\top)\) by Lemma 3.3. Thus \(\text{PA} + \text{Con}_T \vdash \Pr^R_g(\neg\varphi^\top) \leftrightarrow \Pr^R_g(\neg\neg\varphi^\top)\). It follows that \(\text{PA} + \text{Con}_T \vdash \Pr^R_g(\neg\varphi \rightarrow \psi^\top) \rightarrow (\Pr^R_g(\neg\varphi^\top) \rightarrow \Pr^R_g(\neg\psi^\top))\).

Then it suffices to prove \(\text{PA} + \neg\text{Con}_T \vdash \Pr^R_g(\neg\varphi \rightarrow \psi^\top) \rightarrow (\Pr^R_g(\neg\varphi^\top) \rightarrow \Pr^R_g(\neg\psi^\top))\). We work in \(\text{PA} + \neg\text{Con}_T\): By Lemma 3.4.3, there exists \(i \neq 0\) such that \(S(i)\) holds. Then there exists \(m\) such that \(h(m) = 0\) and \(h(m + 1) = i\). Let \(X = P_{T,m-1} \cup \{\bigvee_{i \prec j} S(\overline{j})\}\).
Suppose $Pr^R_g(\psi)$ and $Pr^R_g(\varphi)$. Then $\varphi \to \psi$ and $\varphi$ are t.c.'s of $X$ by Lemma 3.4.3. Hence $\psi$ is also a t.c. of $X$. We conclude $Pr^R_g(\psi)$ by Lemma 3.4.4 (because $\psi$ is standard).

Define $f$ to be the arithmetical interpretation based on $Pr^R_g(x)$ by $f(p) \equiv \exists x(S(x) \land x \neq 0 \land x \vdash \psi)$ for each propositional variable $p$.

**Lemma 3.7.** Let $i \in W$ and $A$ be any modal formula.

1. If $i \vdash A$, then $PA \vdash S(\overline{t}) \to f(A)$.

2. If $i \nvdash A$, then $PA \vdash S(\overline{t}) \to \neg f(A)$.

*Proof.* We prove 1 and 2 simultaneously by induction on the construction of $A$. We give a proof only for the case that $A$ is of the form $\Box B$.

1. Suppose $i \vdash \Box B$. Then for any $j > i$, $j \vdash B$. By induction hypothesis, $PA \vdash \bigvee_{i < j} S(\overline{j}) \to f(B)$. By Lemmas 3.3 and 3.6 $PA \vdash Pr^R_g(\bigvee_{i < j} S(\overline{j})) \to Pr^R_g(f(B))$. Since $PA \vdash S(\overline{t}) \to Pr^R_g(\bigvee_{i < j} S(\overline{j}))$ by Lemma 3.5, we obtain $PA \vdash S(\overline{t}) \to f(\Box B)$.

2. Suppose $i \nvdash \Box B$. Then there exists $j \in W$ such that $i < j$ and $j \nvdash B$. By induction hypothesis, we have $PA \vdash S(\overline{j}) \to \neg f(B)$. By Lemmas 3.3 and 3.6 $PA \vdash \neg Pr^R_g(\neg S(\overline{j})) \to \neg Pr^R_g(f(B))$. Since $i < j$, we obtain $PA \vdash S(\overline{t}) \to \neg Pr^R_g(\neg S(\overline{j}))$ by Lemma 3.5. Therefore $PA \vdash S(\overline{t}) \to \neg f(\Box B)$.

We finish the proof of Theorem 3.1.

*Proof of Theorem 3.1*

1. Arithmetical soundness follows from Lemmas 3.3 and 3.6.

2. Suppose $KD \nvdash A$. Then there exists $i \in W$ such that $i \nvdash A$. By Lemma 3.7, $PA \vdash S(\overline{t}) \to \neg f(A)$. Since $T \nvdash S(\overline{t})$ by Lemma 3.2.4, we obtain $T \nvdash f(A)$.

Notice that our arithmetical interpretation $f$ maps each propositional variable to a $\Sigma_1$ sentence. We say such an arithmetical interpretation a $\Sigma_1$ arithmetical interpretation. Then we obtain the following corollary.

**Corollary 3.8.** Let $A$ be any modal formula. The following are equivalent:

1. $KD \vdash A$.

2. For any normal Rosser provability predicate $Pr^R_T(x)$ of $T$ and any arithmetical interpretation $f$ based on $Pr^R_T(x)$, $T \vdash f(A)$.

3. For any normal Rosser provability predicate $Pr^R_T(x)$ of $T$ and any $\Sigma_1$ arithmetical interpretation $f$ based on $Pr^R_T(x)$, $T \vdash f(A)$.
4 Normal modal logic KDR

It is known that the formalized version of Proposition 2.2 is provable in PA, that is, for any formula \( \varphi \), \( \text{PA} \vdash \text{Prov}_T(\neg \varphi) \rightarrow \text{Prov}_T(\neg \text{Pr}^R_T(\neg \varphi)) \) (see [15]). Relating to this observation, in this section, we consider the following condition for Rosser provability predicates:

For any sentence \( \varphi \), \( T \vdash \text{Pr}^R_T(\neg \varphi) \rightarrow \text{Pr}^R_T(\neg \text{Pr}^R_T(\neg \varphi)) \). (1)

We introduce a new normal modal logic \( KDR \).

**Definition 4.1.** \( KDR = KD + \Box \neg p \rightarrow \Box \neg \Box p \).

Since \( KD \not\vdash \Box \neg p \rightarrow \Box \neg \Box p \), KDR is a proper extension of KD, and hence the condition (1) is not valid for some Rosser provability predicate by Theorem 3.1.

It is easy to show that the validity of the modal formula \( \Box \neg p \rightarrow \Box \neg \Box p \) in a Kripke frame \( F = (W, \prec) \) is characterized by the condition

\[
\forall x \forall y \in W (x \prec y \Rightarrow \exists z \in W (x \prec z \land y \prec z)).
\] (2)

We say a Kripke model \( M = (W, \prec, \vdash) \) is a KDR-model if \( \prec \) is serial and satisfies the condition (2). Then it is proved that KDR is sound and complete with respect to the class of all KDR-models (This follows from Theorem 3 in Boolos [2] p. 89 because KDR is \( K\{\langle 0, 0, 1, 1 \rangle, \langle 0, 1, 1, 1 \rangle \} \) in the terminology of Boolos).

**Proposition 4.2.** For any modal formula \( A \), the following are equivalent:

1. \( KDR \vdash A \).
2. \( A \) is valid in all KDR-models.

Here we give an alternative axiomatization of KDR.

**Definition 4.3.**

1. Let \( KR = K + \Box \neg p \rightarrow \Box \neg \Box p \).
2. Let \( KR^+ \) be the logic obtained by adding the inference rule \( \Box A \rightarrow A \) to \( KR \).

**Proposition 4.4.** For any modal formula \( A \), the following are equivalent:

1. \( KDR \vdash A \).
2. \( KR^+ \vdash A \).

**Proof.** (1 \( \Rightarrow \) 2): It suffices to show \( KR^+ \vdash \neg \Box \bot \). Since \( KR \vdash \Box \neg \bot \rightarrow \Box \neg \Box \bot \) and \( K \vdash \Box \neg \bot \), we have \( KR \vdash \Box \neg \Box \bot \). Then \( KR^+ \vdash \neg \Box \bot \).

(2 \( \Rightarrow \) 1): It suffices to show that the rule \( \Box A \rightarrow A \) is admissible in KDR. Suppose \( KDR \not\vdash A \). Then there exists a KDR-model \( M = (W, \prec, \vdash) \) in
which $A$ is not valid by Proposition 4.2. Let $M' = (W', \prec', \vdash')$ be the Kripke model defined as follows:

- $W' = W \cup \{0\}$, where 0 is an element not contained in $W$,
- $\prec' = \prec \cup \{(0, w) : w \in W\}$,
- $0 \vdash' p$; and $x \vdash' p \iff x \vdash p$ for $x \in W$.

Then $M'$ is also a KDR-model. Since $0 \not\vdash A$, we obtain $\KDR \not\vdash A$ by Proposition 4.2 again.

In Theorem 4.6 below, we prove the existence of a normal Rosser provability predicate $\Pr^R_T(x)$ satisfying $\KDR \subseteq \PL(\Pr^R_T)$.

**Proposition 4.5.** If $\KDR \subseteq \PL(\Pr^R_T)$, then for some formula $\varphi$, $T$ cannot prove $\neg\Con_T \rightarrow (\Pr^R_T(\neg\varphi) \lor \Pr^R_T(\varphi))$.

**Proof.** Suppose $\KDR \subseteq \PL(\Pr^R_T)$. Let $\varphi$ be a sentence satisfying the equivalence $T \vdash \varphi \iff \neg\Pr^R_T(\neg\varphi)$. Since $\KDR \subseteq \PL(\Pr^R_T)$, we have $T \vdash \Pr^R_T(\neg\varphi) \rightarrow \Pr^R_T(\neg\Pr^R_T(\varphi))$. Then $T \vdash \Pr^R_T(\neg\varphi) \rightarrow \Pr^R_T(\varphi)$.

We obtain $T \vdash \neg\Pr^R_T(\neg\varphi)$ because $T \vdash \Pr^R_T(\neg\varphi) \rightarrow \neg\Pr^R_T(\varphi)$. Then $T \vdash \neg\Con_T \rightarrow \Pr^R_T(\neg\varphi)$ by the supposition, and hence $T \vdash \neg\Con_T \rightarrow \neg\varphi$. By the formalized version of Rosser’s first incompleteness theorem, we have $T \vdash \Prov_T(\neg\varphi) \rightarrow \neg\Con_T$ (see [10]). Thus $T \vdash \Prov_T(\neg\varphi) \rightarrow \neg\varphi$. By Löb’s theorem, we obtain $T \vdash \neg\varphi$. Then $T$ is inconsistent by Rosser’s first incompleteness theorem. This is a contradiction.

We prove the main theorem of this section.

**Theorem 4.6.** There exists a normal Rosser provability predicate $\Pr^R_T(x)$ of $T$ such that $\KDR \subseteq \PL(\Pr^R_T)$.

**Proof.** We define a PA-provably recursive function $g'(x)$ in stages. The corresponding formulas $\Pr_{g'}(x, y)$, $\Pr_{g'}(x)$, and $\Pr_{g'}(x, y)$ are defined as in the proof of Theorem 3.1. In the definition of $g'(x)$, we can refer to our defining formula $\Pr^R_{g'}(x)$ with the aid of the recursion theorem.

For each natural number $m$, let $F_m$ be the set of all formulas whose Gödel numbers are less than or equal to $m$. First, we define an increasing sequence $\{Y^i_m\}_{i \in \omega}$ of finite sets of formulas recursively as follows:

- $Y^0_m = \Pr_{m}$,
- $Y^{i+1}_m = Y^i_m \cup \{\neg\Pr^R_{g'}(\varphi') : \neg\Pr^R_{g'}(\varphi') \in F_m \land \neg\varphi' \text{ is a t.c. of } Y^i_m\}$.

13
Also let $Y_m = \bigcup_{i \in \omega} Y^i_m$.

Notice that the definition of $Y_m$ refers to the formula $Pr^R_g(x)$. It is easy to see that for any $i \in \omega$, $Y^i_m \subseteq F_m$. Then $Y_m$ is in fact the finite set $\bigcup_{i \leq m} Y^i_m$ because $F_m$ contains at most $m$ formulas.

We start defining our function $g'(x)$.

Procedure 1.
Stage 1.m:

- If $Y_m$ is propositionally satisfiable, then
  
  $g'(m) = \begin{cases} 
  \varphi & \text{if } m \text{ is a proof of } \varphi \text{ in } T, \text{ that is, } \text{Proof}_T(\neg \varphi, m) \text{ holds}, \\
  0 & m \text{ is not a proof of any formula in } T.
  \end{cases}$

  Go to Stage 1.(m + 1).

- If $Y_m$ is not propositionally satisfiable, then go to Procedure 2.

Procedure 2. Let $m$ be the least number such that $Y_m$ is not propositionally satisfiable. Let $X = Y_m - 1$. Then $X$ is propositionally satisfiable. We define the values of $g'(m + t)$ for $t \geq 0$ by copying the corresponding part of Procedure 2 in the definition of the function $g(x)$ in our proof of Theorem 3.1 with the present definition of $X$. The definition of $g'(x)$ is completed.

**Lemma 4.7.** PA proves the following statement:

“Let $m$ be the least number such that $Y_m$ is not propositionally satisfiable and let $X = Y_{m-1}$. Then every element of $X$ is provable in $T$.

1. For any formula $\varphi \in F_m$, $\varphi$ is a t.c. of $X$ if and only if $Pr^R_g(\neg \varphi, m)$ holds.

2. Every element of $X$ is provable in $T$.

3. Suppose for all $j < i$, there exists no formula $\varphi$ such that $\neg Pr^R_g(\neg \varphi, m)$ is in $F_m$, $\varphi$ is a t.c. of $X$ and $\neg \varphi$ is a t.c. of $Y^i_m$. Then every element of $Y^i_m$ is provable in $T$.”

**Proof.** We proceed in PA. Let $m$ be the least number such that $Y_m$ is not propositionally satisfiable and let $X = Y_{m-1}$. Then the construction of $g'(x)$ goes to Procedure 2 at Stage 1.m.

1. Let $\varphi$ be any formula with $\varphi \in F_m$.

   $(\rightarrow)$: Suppose $\varphi$ is a t.c. of $X$. Then $\neg \varphi \notin P_{T, m-1}$ because $P_{T, m-1} \subseteq X$ and $X$ is propositionally satisfiable. By tracing our proof of Lemma 3.3, we can show that $Pr^R_g(\neg \varphi, m)$ holds.

   $(\leftarrow)$: Suppose $\varphi$ is not a t.c. of $X$. Then $\varphi \notin P_{T, m-1}$. The number of leading $\neg$’s of $\varphi$ is less than $m$ because $\varphi \in F_m$. Then we can show that $Pr^R_g(\neg \varphi, m)$ does not hold by tracing our proof of Lemma 3.4.4 for $\psi = \varphi.$
2. We prove by induction on $j$ that for all $j$, every element of $Y^j_{m-1}$ is provable in $T$. For $j = 0$, the statement trivially holds because $Y^0_{m-1} = P_{T,m-1}$. Suppose that every formula contained in $Y^j_{m-1}$ is $T$-provable. Let $\varphi$ be any formula such that $\neg \Pr^R_{g'}(\varphi) \in F_{m-1}$ and $\neg \varphi$ is a t.c. of $Y^j_{m-1}$. Since $Y^j_{m-1} \subseteq X$ and $X$ is propositionally satisfiable, $\varphi$ is not a t.c. of $X$. Since $\varphi \in F_m$, $\Pr^R_{g'}(\varphi)$ does not hold by 1. Moreover, $g'(x)$ outputs $\neg \varphi$ before any output of $\varphi$, and this fact is provable in $T$. Thus $\neg \Pr^R_{g'}(\varphi)$ is provable in $T$. Therefore every element of $Y^j_{m-1}$ is provable in $T$ because $Y^{j+1}_{m-1} = Y^j_{m-1} \cup \{\neg \Pr^R_{g'}(\varphi)\}$.

3. This is proved in a similar way as in our proof of 2.

Lemma 4.8.

1. $\text{PA} \vdash \text{Con} \iff \forall x \text{ "} Y_x \text{ is propositionally satisfiable".}

2. For any $n \in \omega$, $\text{PA} \vdash \text{ "} Y_n \text{ is propositionally satisfiable".}

Proof. 1. We proceed in $\text{PA}$. 

($\rightarrow$): Let $m$ be the least number such that $Y^m_m$ is not propositional satisfiable, and let $X = Y^m_m$. Also let $i$ be the least number such that $Y^i_m$ is not propositionally satisfiable. We would like to show that $T$ is inconsistent.

If $i = 0$, $Y^0_m = P_{T,m}$ is not propositionally satisfiable. In this case, $g'(m)$ is the formula $\varphi$ with $P_{T,m} = P_{T,m-1} \cup \{\varphi\}$, and $\neg \varphi$ is a t.c. of $P_{T,m-1}$. Thus both $\varphi$ and $\neg \varphi$ are provable in $T$, and hence $T$ is inconsistent. Therefore we may assume $i > 0$.

Suppose that some element of $Y^i_m$ is not provable in $T$. By Lemma 4.7.3, there exists the least $j < i$ such that for some formula $\varphi$, $\neg \Pr^R_{g'}(\varphi) \in F_m$, $\varphi$ is a t.c. of $X$ and $\neg \varphi$ is a t.c. of $Y^j_m$. By Lemma 4.7.1, $\varphi$ is $T$-provable. By the choice of $j$, every element of $Y^j_m$ is provable in $T$ by Lemma 4.7.3. Hence $\neg \varphi$ is also $T$-provable. Therefore $T$ is inconsistent. This contradicts the supposition.

Therefore all elements of $Y^i_m$ are provable in $T$. Let $\psi_0, \ldots, \psi_l$ be formulas such that $Y^i_m = P_{T,m} \cup \{\neg \Pr^R_{g'}(\psi_0), \ldots, \neg \Pr^R_{g'}(\psi_l)\}$. In this case, $\bigwedge_{k \leq l} \neg \Pr^R_{g'}(\psi_k)$ is provable in $T$. On the other hand, $\bigvee_{k \leq l} \Pr^R_{g'}(\psi_k)$ is a t.c. of $P_{T,m}$ because $Y^i_m$ is not propositionally satisfiable. Thus $T$ proves $\bigvee_{k \leq l} \Pr^R_{g'}(\psi_k)$. It follows that $T$ is inconsistent.

($\leftarrow$): Suppose that $T$ is inconsistent. Then $P_{T,m}$ is not propositionally satisfiable for some $m$. Therefore $Y_m$ is not propositionally satisfiable for some $m$.

2. By 1, for any $n \in \omega$, the $\Sigma_1$ sentence “$Y_n$ is propositionally satisfiable” is true, and hence provable in $\text{PA}$.

15
Our formula $\text{Pr}_g(x, y)$ is a proof predicate of $T$.

**Lemma 4.9.**

1. $\text{PA} \vdash \forall x (\text{Prov}_T(x) \leftrightarrow \text{Pr}_g(x))$.
2. For any $n \in \omega$ and formula $\varphi$, $\text{N} \models \text{Proof}_T(\neg \varphi, \pi) \leftrightarrow \text{Pr}_g(\neg \varphi, \pi)$.

**Proof.** From Lemma 4.8 this is proved in a similar way as in our proof of Lemma 3.3. □

**Lemma 4.10.** Let $\varphi$ and $\psi$ be any formulas.

1. $\text{PA} + \text{Con}_T \vdash \text{Pr}_g^R(\neg \varphi \to \psi) \to (\text{Pr}_g^R(\varphi) \to \text{Pr}_g^R(\neg \varphi \to \psi))$.
2. $\text{PA} + \text{Con}_T \vdash \text{Pr}_g^R(\neg \neg \varphi) \to \text{Pr}_g^R(\neg \varphi)$.

**Proof.** Let $U = \text{PA} + \text{Con}_T$. By Lemma 4.8, $U$ proves $\forall x \forall y_x$ “$Y_x$ is propositionally satisfiable”. Then by the definition of $g'(x)$, the formulas $\text{Prov}_T(x)$, $\text{Pr}_g(x)$ and $\text{Pr}_g^R(x)$ are all equivalent in $U$.

2. Since $\text{PA} \vdash \text{Prov}_T(\neg \varphi) \to \text{Pr}_g^R(\neg \varphi)$, we have $U \vdash \text{Pr}_g^R(\neg \varphi \to \psi) \to (\text{Pr}_g^R(\varphi) \to \text{Pr}_g^R(\neg \varphi \to \psi))$.

2. We reason in $U$: Suppose $\text{Pr}_g^R(\neg \neg \varphi)$ holds. Then $\neg \varphi$ is output by $g'(x)$. Moreover, $g'(x)$ outputs $\neg \varphi$ before any output of $\varphi$ since $T$ is consistent, and this fact is provable in $T$. Hence $T$ proves $\neg \text{Pr}_g^R(\varphi)$. Therefore $\text{Pr}_g^R(\neg \text{Pr}_g^R(\varphi))$ holds by the equivalence of $\text{Prov}_T(x)$ and $\text{Pr}_g^R(x)$. □

**Lemma 4.11.** Let $\varphi$ and $\psi$ be any formulas.

1. $\text{PA} + \neg \text{Con}_T \vdash \text{Pr}_g^R(\neg \varphi \to \psi) \to (\text{Pr}_g^R(\varphi) \to \text{Pr}_g^R(\neg \varphi \to \psi))$.
2. $\text{PA} + \neg \text{Con}_T \vdash \text{Pr}_g^R(\neg \neg \varphi) \to \text{Pr}_g^R(\neg \varphi)$.

**Proof.** We reason in $\text{PA} + \neg \text{Con}_T$: By Lemma 4.8, there exists $m$ such that $Y_m$ is not propositionally satisfiable. Let $m$ be the least such number and let $X = Y_{m-1}$. By Lemma 4.8, $m$ is larger than the Gödel numbers of $\varphi$, $\psi$, $\varphi \to \psi$ and $\neg \text{Pr}_g^R(\varphi)$. That is, $\varphi$, $\psi$, $\varphi \to \psi$ and $\neg \text{Pr}_g^R(\varphi)$ are in $F_m$.

1. Suppose $\text{Pr}_g^R(\neg \varphi \to \psi)$ and $\text{Pr}_g^R(\varphi)$ hold. Then $\varphi \to \psi$ and $\varphi$ are t.c.’s of $X$ by Lemma 4.7. Since $\psi$ is also a t.c. of $X$, $\text{Pr}_g^R(\neg \varphi \to \psi)$ holds by Lemma 4.7 again.

2. Suppose $\text{Pr}_g^R(\neg \neg \varphi)$ holds. Then $\neg \varphi$ is a t.c. of $X$ by Lemma 4.7. Since $X = Y_{m-1} = \bigcup_{i \in \omega} Y_{m-1}$, $\neg \varphi$ is a t.c. of $Y_{m-1}$ for some $i \in \omega$. Then $\neg \text{Pr}_g^R(\varphi) \in Y_{m-1} \subseteq X$ because $\neg \text{Pr}_g^R(\varphi)$ is in $F_m$. Thus $\neg \text{Pr}_g^R(\varphi)$ is also a t.c. of $X$. By Lemma 4.7 again, $\text{Pr}_g^R(\neg \text{Pr}_g^R(\varphi))$ holds. □

By Lemmas 4.9, 4.10 and 4.11 we conclude $\text{KDR} \subseteq \text{PL}(\text{Pr}_g^R)$. □
Let $F$ be Shavrukov’s modal logic $\text{KD} + \Box p \rightarrow \Box((\Box q \rightarrow q) \lor \Box p)$. It is easy to see that $F$ is included in $\text{KD}4 \cap T$. Also we obtain the following proposition.

**Proposition 4.12.** $\text{KDR} \subseteq \text{KD}5 \cap F$. Consequently, $\text{KDR} \subseteq \text{KD}5 \cap \text{KD}4 \cap T$.

**Proof.** $\text{KDR} \subseteq \text{KD}5$: This is because $\text{KD} \vdash \Box \neg p \rightarrow \neg \Box p$ and $\text{KD}5 \vdash \neg \Box p \rightarrow \neg \Box p$.

$\text{KDR} \subseteq F$: Since $F \vdash \Box \neg p \rightarrow \Box \neg p \land \Box((\Box p \rightarrow p) \lor \Box \neg p)$, we have $F \vdash \Box \neg p \rightarrow \Box(\neg \Box p \lor \Box \neg p)$. Since $\text{KD} \vdash \Box(\neg \Box p \lor \Box \neg p) \rightarrow \Box \neg \Box p$, we conclude $F \vdash \Box \neg p \rightarrow \Box \neg \Box p$. \hfill $\square$

In Theorem 4.4, we proved that there exists a Rosser provability predicate $\text{Pr}_T^R(x)$ such that $\text{KDR} \subseteq \text{PL}(\text{Pr}_T^R)$. On the other hand, $\text{PL}(\text{Pr}_T^R)$ cannot include $\text{KD}5 \cap \text{KD}4 \cap T$.

**Proposition 4.13.** There exists no Rosser provability predicate $\text{Pr}_T^R(x)$ such that $\text{KD}5 \cap \text{KD}4 \cap T \subseteq \text{PL}(\text{Pr}_T^R)$.

**Proof.** Let $\text{Pr}_T^R(x)$ be any Rosser provability predicate such that $\text{KD} \subseteq \text{PL}(\text{Pr}_T^R)$, and let $\varphi$ be a sentence satisfying $T \vdash \varphi \leftrightarrow \neg \text{Pr}_T^R(\Box \varphi)$. Let $\xi$ be the sentence $\neg \text{Pr}_T^R(\Box \varphi) \rightarrow \text{Pr}_T^R(\neg \neg \text{Pr}_T^R(\neg \Box \varphi))$. Then $T + \xi \vdash \varphi \rightarrow \text{Pr}_T^R(\neg \neg \text{Pr}_T^R(\neg \Box \varphi))$ and $T + \xi \vdash \varphi \rightarrow \neg \varphi$. Hence $T + \xi \vdash \neg \varphi$. Then we have $T \vdash \text{Pr}_T^R(\neg \neg \xi) \rightarrow \text{Pr}_T^R(\neg \varphi)$, and $T \vdash \text{Pr}_T^R(\neg \neg \xi) \rightarrow \neg \text{Pr}_T^R(\neg \varphi)$ because $\text{KD} \subseteq \text{PL}(\text{Pr}_T^R)$. Therefore $T + \text{Pr}_T^R(\neg \neg \xi) \vdash \varphi$. By combining this with $T + \xi \vdash \neg \varphi$, we obtain that $T + \xi \lor \text{Pr}_T^R(\neg \neg \xi)$ is inconsistent.

Also let $\eta$ and $\gamma$ be the sentences $\text{Pr}_T^R(\neg \neg \varphi) \rightarrow \text{Pr}_T^R(\neg \neg \text{Pr}_T^R(\neg \varphi))$ and $\text{Pr}_T^R(\neg \varphi) \rightarrow \varphi$, respectively. We can show that the theories $T + \xi$ and $T + \gamma$ prove $\varphi$. Then $T + \xi \vdash \text{Pr}_T^R(\neg \neg \eta)$ and $T + \gamma \vdash \text{Pr}_T^R(\neg \neg \gamma)$ prove $\neg \varphi$. Therefore the theories $T + \xi \land \text{Pr}_T^R(\neg \neg \eta)$ and $T + \gamma \land \text{Pr}_T^R(\neg \neg \gamma)$ are also inconsistent. Let $\alpha$ be the sentence

$$(\xi \land \text{Pr}_T^R(\neg \neg \eta)) \lor (\eta \land \text{Pr}_T^R(\neg \neg \xi)) \lor (\gamma \land \text{Pr}_T^R(\neg \neg \gamma)).$$

Then we have shown $T \vdash \neg \alpha$. On the other hand, the sentence $\alpha$ is an arithmetical instance of a modal formula which is in $\text{KD}5 \cap \text{KD}4 \cap T$. Therefore we conclude $\text{KD}5 \cap \text{KD}4 \cap T \not\subseteq \text{PL}(\text{Pr}_T^R)$. \hfill $\square$

We propose a question concerning the logics $\text{KDR}$ and $F$.

**Problem 4.14.** Are there respective normal Rosser provability predicates $\text{Pr}_T^R(x)$ of $T$ satisfying each of the following conditions?

1. $\text{KD} \subseteq \text{PL}(\text{Pr}_T^R) \subseteq \text{KDR}$.
2. $\text{KD} \subseteq \text{PL}(\text{Pr}_T^R)$ and $\text{PL}(\text{Pr}_T^R)$ is not compatible with $\text{KDR}$.
3. $\text{KDR} = \text{PL}(\text{Pr}_T^R)$.
4. $F \subseteq \text{PL}(\text{Pr}_T^R)$ or $F = \text{PL}(\text{Pr}_T^R)$. 

17
5. KDR ⊆ PL(Pr^R_T) ⊆ KD5 ∩ KD4 ∩ T.
6. KDR ⊆ PL(Pr^R_T) ⊆ KD5 ∩ KD4 ∩ T.
7. KDR ⊆ PL(Pr^R_T) ⊆ F.
8. KDR ⊆ PL(Pr^R_T) ⊆ F.

In Theorem 3.1, we proved that there exists a Rosser provability predicate Pr^R_T(x) such that PL(Pr^R_T) = KD, and our proof does not guarantee that Pr^R_T(x) satisfies the following stronger version of the principle (K):

(K'): T ⊬ ∀x∀y(Pr^R_T(x→y) → (Pr^R_T(x) → Pr^R_T(y))).

Here x→y is a primitive recursive term corresponding to a primitive recursive function calculating the Gödel number of ϕ → ψ from Gödel numbers of ϕ and ψ. We can say the same comment for Theorem 4.6.

We close this paper with the following problem.

**Problem 4.15.**

1. Is there a Rosser provability predicate Pr^R_T(x) such that PL(Pr^R_T) = KD and Pr^R_T(x) satisfies (K')?
2. Is there a Rosser provability predicate Pr^R_T(x) such that KDR ⊆ PL(Pr^R_T) and Pr^R_T(x) satisfies (K')?

**References**

[1] Toshiyasu Arai. Derivability conditions on Rosser’s provability predicates. *Notre Dame Journal of Formal Logic*, 31(4):487–497, 1990.

[2] George Boolos. *The logic of provability*. Cambridge University Press, Cambridge, 1993.

[3] Alexander Chagrov and Michael Zakharyaschev. *Modal Logic*. Oxford Logic Guides. Clarendon Press, 1997.

[4] Solomon Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49:35–92, 1960.

[5] David Guaspari and Robert M. Solovay. Rosser sentences. *Annals of Mathematical Logic*, 16(1):81–99, 1979.

[6] Makoto Kikuchi and Taishi Kurahashi. Universal Rosser predicates. *The Journal of Symbolic Logic*, 82(1):292–302, 2017.

[7] Georg Kreisel and Gaisi Takeuti. Formally self-referential propositions for cut free classical analysis and related systems. *Dissertationes Mathematicae (Rozprawy Matematyczne)*, 118, 1974.

[8] Taishi Kurahashi. Arithmetical completeness theorem for modal logic K. *Studia Logica*, 106(2):219–235, 2018.
[9] Taishi Kurahashi. Arithmetical soundness and completeness for $\Sigma_2$ numerations. *Studia Logica*, 106(6):1181–1196, 2018.

[10] Per Lindström. *Aspects of Incompleteness*. Number 10 in Lecture Notes in Logic. A K Peters, 2nd edition, 2003.

[11] Franco Montagna. On the algebraization of a Feferman’s predicate. *Studia Logica*, 37(3):221–236, 1978.

[12] Wolfgang Rautenberg. Modal tableau calculi and interpolation. *Journal of Philosophical Logic*, 12(4):403–423, 1983.

[13] John Barkley Rosser. Extensions of some theorems of Gödel and Church. *The Journal of Symbolic Logic*, 1(3):87–91, 1936.

[14] Lorenzo Sacchetti. The fixed point property in modal logic. *Notre Dame Journal of Formal Logic*, 42(2):65–86, 2001.

[15] Vladimir Yurievich Shavrukov. On Rosser’s provability predicate. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 37(4):317–330, 1991.

[16] Vladimir Yurievich Shavrukov. A smart child of Peano’s. *Notre Dame Journal of Formal Logic*, 35(2):161–185, 1994.

[17] Craig Smoryński. *Self-reference and modal logic*. Universitext. Springer-Verlag, New York, 1985.

[18] Robert M. Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25(3-4):287–304, 1976.

[19] Albert Visser. Peano’s smart children: A provability logical study of systems with built-in consistency. *Notre Dame Journal of Formal Logic*, 30(2):161–196, 1989.