TSUKUBA J. MATH.
Vol.6 No. 2 (1982). 201—213

ON REFLECTION PRINCIPLES

By

Akito TSUBOI

Introduction.

In this paper, we shall consider various forms of reflection principles for 1-st order theories containing arithmetic. If a 1-st order theory T contains arithmetic, we can express various notions concerning T within T itself, by using the coding method developed by K. Gödel. K. Gödel assigned each formula ϕ in the language of T a number $\ulcorner\phi\urcorner$ (the Gödel number of ϕ), but our method is slightly different.

We assume that variables, individual constants, relation symbols and function symbols are numbers, and logical symbols ($*\wedge*$, $*\vee*$, $\neg*$, $*\rightarrow*$, $\forall**$, $\exists**$) are operations on numbers. Under these assumptions, a formula ϕ itself is a number and a theory T, which is a set of sentences, can be conceived as a subset of ω (the set of natural numbers). Let S be a theory and A a subset of ω. We say a formula $\alpha(x)$ in the language of S numerates A in S if, for any $n\in\omega$,

$$n\in A \text{ iff } S \text{ proves } \alpha(\bar{n}),$$

where $\bar{n}$ denotes the n-th numeral, i.e., the term of S which expresses the number n. In this case, we call this α a numeration of A in S. If α numerates A in S and $\neg\alpha$ numerates $\omega\backslash A$ in S, we say α binumerates A in S, and α is called a binumeration of A in S.

Let $A=\{n_1, \cdots, n_m\}$ be a subset of ω. Then $[A]$ denotes the formula $x=\bar{n}_1\vee\cdots\vee x=\bar{n}_m$. Clearly $[A]$ binumerates A in any theory S which contains arithmetic.

If a binumeration τ of a theory T in a theory S is given, we can construct a provability formula $Pr_\tau(x)$ whose intuitive meaning is that a formula x is provable in T. The reader should note that this Pr_τ cannot be uniquely determined by T, but is determined by τ. (The explicit definition of Pr_τ can be found in p. 59 of [1].)

Using this Pr_τ, we define the τ-reflection principle $Rfn(\tau)$ and the τ-reflection principle $Rfn_A(\tau)$ based on A:

Received October 9, 1981. Revised March 26, 1982.

$$Rfn(\tau)=\{Pr_\tau(\sigma)\rightarrow\sigma \mid \sigma\in \mathcal{S}ent_T\},$$

$$Rfn_A(\tau)=\{Pr_\tau(\sigma)\rightarrow\sigma \mid \sigma\in A\cap \mathcal{S}ent_T\},$$

where $\mathcal{S}ent_T$ is the set of all sentences in the language of T.

A formula ϕ is said to be a Σ_n-formula (Π_n-formula) if ϕ has the form $Q_1x_1\cdots Q_nx_n\psi(x_1,\cdots,x_n)$ for some quantifier bounded formula ψ, where $Q_1=\exists(\forall)$ and the quantifiers alternate in type. The set of all Σ_n-formulas (Π_n-formulas) is denoted by Σ_n (Π_n).

Let T be a 1-st order theory containing arithmetic. We say T is n-consistent if the following two conditions are not simultaneously satisfied for any Π_{n-1}-formula ϕ:

i) T proves $\exists x\phi(x)$,
ii) T proves $\neg\phi(\bar{m})$ for all $m\in\omega$.

If T is n-consistent for all $n\in\omega$, we say T is ω-consistent. If T proves $Con_{[T_0]}$ $(=\neg Pr_{[T_0]}(\overline{\bar{1}=\bar{0}}))$ for all finite subtheories T_0 of T, T is said to be reflexive. If each extension T^* of T with the same language is reflexive, we say T is essentially reflexive. We next define a more complicated notion A-reflexiveness. Let A be a set of sentences. We say T is A-reflexive, if there exist a truth definition $Tr_A(x)$ for A in T and a numeration $\alpha(x)$ of A in T for which T proves $\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_{[T_0]}(x)\rightarrow Tr_A(x))$ for all finite subtheories T_0 of T, where $Sent(x)$ is a formula which expresses that x is a sentence. (See Definition 1.2 and 1.3 for reference.)

For three sets A, B and C of sentences, we put:

$A\subseteqq_B C$ iff each sentence in A is provable in $B\cup C$,
$A=_B C$ iff $A\subseteqq_B C$ and $C\subseteqq_B A$,
$A\subsetneqq_B C$ iff $A\subseteqq_B C$ and $A\neq_B C$.

In case B is the empty set, we usually omit B in the above definitions. In what follows, we say S is a subtheory of T if $S\subseteqq T$ holds in this sense. If T is a theory and A is a set of sentences in the language of T, then we put $T-A=\{\phi\mid\phi$ is equivalent to some $\psi\in A$ in $T\}$.

It is now possible to state the main theorems of this paper.

THEOREM 1. *Suppose that A is a set of sentences. If T is an A-reflexive theory with a binumeration τ of T in T, then we can effectively construct a binumeration τ' of T in T for which T proves each member of $Rfn_A(\tau')$.*

THEOREM 2. *Suppose that T is a recursively enumerable theory* (*r.e. theory*) *and S is a subtheory of T. If* τ *binumerates T in S, then* $Rfn(\tau)\backslash Rfn_{T-(\Sigma_n\cup\Pi_n)}(\tau)=_T Rfn(\tau)$.

THEOREM 3. *Suppose that T is an* ω*-consistent and essentially reflexive theory and S is a subtheory of T. If* τ *binumerates T in S, we can effectively construct binumerations* τ_1 *and* τ_2 *of T in S for which* $Rfn(\tau_1)\subsetneqq_T Rfn(\tau)\subsetneqq_T Rfn(\tau_2)$.

Theorem 1, which appears in § 2, is closely related to Theorem 5.9 of [1]. Theorem 2 shows that the strength of $Rfn(\tau)$ does not change even if the lower part of it is taken away from it. Theorem 2 also appears in § 2. Theorem 3, which appears in § 3, is an analogy of Theorem 7.4 and 7.5 of [1] and shows that the choice of numerations must be done very carefully.

The reader who is accustomed to the coding method can skip § 1 and may refer to it as occasion demands.

The author wishes to express his heartfelt thanks to Prof. N. Motohashi for a number of helpful suggestions.

§ 1. Preliminaries.

Notations, definitions and conventions in this paper largely correspond with those of [1]. Especially, we assume that a formula ϕ itself is a number and we do not use a notation $\ulcorner\phi\urcorner$ (the Gödel number of ϕ).

For simplicity, we say T is a theory when T is a consistent 1-st order theory containing PA (Peano arithmetic). We use T, S and T_i $(i=1, 2, \cdots)$ as syntactic variables ranging over theories, and usually assume that S is a subtheory of T.

It will be convenient to assume that for every theory T, L_T (the language of T) has all the symbols for p. r. functions (primitive recursive functions) and T contains all the defining axioms for p. r. functions. The function symbol associated with a p. r. function f is denoted by $\dot{f}$ or $\underset{\cdot}{f}$. But symbols which are used very often will be used without dots. For example, we write vr_x and nm_x (the x-th variable and the x-th numeral, respectively) instead of writing $v\underset{\cdot}{r}_x$ and $n\underset{\cdot}{m}_x$. $Tm(x)$, $Fm(x)$, $Sent(x)$, $Prf_\tau(x, y)$, $Pr_\tau(x)$ and Con_τ are formulas whose intuitive meanings are "*x is a term*", "*x is a formula*", "*x is a sentence*", "*y is a proof of x in a theory with a numeration* τ" and "*a theory with a numeration* τ *is consistent*", respectively.

CONVENTION. *Let* α *be a formula with a free variable x and A be a set of formulas. We say* α *numerates A in a theory T, if* α *numerates A in T in the* *usual sense and T proves $\forall x(\alpha(x) \rightarrow Fm(x))$. (See Introduction for reference.)*

The above convention Is trivial: If α numerates A in T in the usual sense and T does not prove $\forall x(\alpha(x) \rightarrow Fm(x))$, then we can define α' as $\alpha(x) \wedge Fm(x)$, and α' will numerate A in T in the above sense.

In the following, we state some definitions which do not appear in [1].

1.1. DEFINITION. Let T be a theory and S a subtheory of T. Then we put:

$$Bin(T, S) = \{\tau \mid \tau \text{ binumerates } T \text{ in } S\}.$$

1.2. DEFINITION. Let T be a theory and A a set of formulas in the language of T. We say a formula $Tr_A(x)$ is a truth definition for A in T, if the following is satisfied:

$$T \vdash Tr_A(\bar{\phi}(x_1, \cdots, x_n)) \leftrightarrow \phi(x_1, \cdots, x_n) \quad \text{for all} \quad \phi(x_1, \cdots, x_n) \in A,$$

where $\bar{\phi}(x_1, \cdots, x_n) = \mathrm{sub}(\bar{\phi}\,;\, \bar{x}_1, \cdots, \bar{x}_n / nm_{x_1}, \cdots, nm_{x_n})$, i.e., the *sentence* obtained from ϕ by substituting $nm_{x_1}, \cdots, nm_{x_n}$ for its *free variables* $x_1, \cdots, x_n$.

1.3. DEFINITION. Let T be a theory and A a set of sentences in the language of T. Then we say

i) T is reflexive if $T \vdash Con_{[T_0]}$ for every finite subtheory T_0 of T,

ii) T is essentially reflexive if every extension T^* of T in the same language as T is reflexive,

iii) T is A-reflexive if there exist a numeration α of A in T and a truth definition Tr_A for A in T for which

$$T \vdash \forall x(\alpha(x) \wedge Sent(x) \wedge Pr_{[T_0]}(x) \rightarrow Tr_A(x))$$

holds for every finite subtheory T_0 of T.

1.4. COROLLARY. *The following* i), ii) *and* iii) *are equivalent:*

i) *T is essentially reflexive,*

ii) *$T \vdash Pr_{\tau \restriction \bar{n}}(\bar{\phi}) \rightarrow \phi$ for every $\tau \in Bin(T, T)$, $\phi \in \mathcal{S}ent_T$, $n \in \omega$, (where, of course, $\tau \restriction \bar{n}$ is an abbreviation for $\tau(x) \wedge x \leqq \bar{n}$.)*

iii) *T is $\{\phi\}$-reflexive for every $\phi \in \mathcal{S}ent_T$.*

Since there is no truth definition for all sentences, a reflection principle cannot be formulated in a single sentence. Although there are many versions of a reflection principle, we restrict our attention to the following two types.

1.5. DEFINITION. Let T and S be theories with $S\subseteq T$ and let τ be a formula which binumerates T in S. Then:

i) Local Reflection Principle;

$$Rfn(\tau)=\{Pr_\tau(\bar{\phi})\to\phi \mid \phi\in \mathcal{S}ent_T\},\quad Rfn_A(\tau)=\{Pr_\tau(\bar{\phi})\to\phi \mid \phi\in A\cap \mathcal{S}ent_T\},$$

ii) Uniform Reflection Principle;

$$\mathrm{RFN}(\tau)=\{\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_\tau(x^*)\to Tr_n(x^*)) \mid n\in\omega\},$$

where Tr_n is the standard truth definition for $\Sigma_n\cup\Pi_n$ and x^* denotes the *sentence* obtained from x by substituting $nm_{(y)_0}$, $nm_{(y)_1}$, $\cdots$ for its *free variables.*

Does T remain consistent when a reflection principle is added to it? The following theorem gives us a partial solution.

1.6. TNEOREM. (Refinement of Theorems 20 and 24 of [4]) *Let T, S and τ be as above. Then:*

i) *If $\tau\in\Sigma_1$ and T is 1-consistent, then $T\cup Rfn(\tau)$ is 1-consistent,*

ii) *If $\tau\in\Sigma_n$ and T is n-consistent, then $T\cup Rfn(\tau)$ is 2-consistent.* ($n=2, 3, \cdots$),

iii) *If T is ω-consistent, then $T\cup RFN(\tau)$ is 2-consistent.*

REMARK. T. Miyatake showed that if $\tau\in\Sigma_1$, then the converse of iii) also holds. If $\tau\notin\Sigma_1$, the 2-consistency is not enough, but the weak converse of iii) holds, and it can be stated as follows: if $\tau\in\Sigma_n$ and $T\cup RFN(\tau)$ is $n+1$-consistent, then T is ω-consistent. It is not hard to give an example of T for which $T\cup Rfn(\tau)$ is inconsistent. The reader may refer to [4] for this purpose.

§ 2. Hierarchy Considerations.

By Gödel's Second Incompleteness Theorem, if $\tau\in Bin(T, T)$ is a Σ_1-formula, Con_τ cannot be proved in T. S. Feferman, however, in [1] showed that in case T is reflexive, we can choose $\tau\in Bin(T, T)$ for which PA proves Con_τ. Since Con_τ and $Rfn_{\Pi_1}(\tau)$ are equivalent over T, we can also prove all elements of $Rfn_{\Pi_1}(\tau)$ in T for the above τ. The following theorem is a generalization of this fact.

2.1. THEOREM. *Suppose that A is a set of sentences. If T is an A-reflexive theory with a binumeration $\tau\in Bin(T, T)$, then we can effectively construct from τ a binumeration $\tau'\in Bin(T, T)$ for which T prove each element of $Rfn_A(\tau')$.*

PROOF. Since T is A-reflexive, we can choose a numeration $\alpha(x)$ of A in T and a truth definition $Tr_A(x)$ for A in T such that

$$T\vdash\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_{\tau\restriction\bar{n}}(x)\rightarrow Tr_A(x)) \quad \text{for all} \quad n\in\omega.$$

Set $\beta(x)=\forall y(\alpha(y)\wedge Sent(y)\wedge Pr_{\tau\restriction x}(y)\rightarrow Tr_A(y))$, $\tau'(x)=\tau(x)\wedge\forall y\leqq\beta(y)$. We prove that this τ' has the desired properties. Since $\tau'\in Bin(T, T)$ is easily obtained from the assumptions, we have only to show that T proves each element of $Rfn_A(\tau')$. First note that $T\vdash Pr_\tau(x)\leftrightarrow\exists yPr_{\tau\restriction y}(x)$, then

$$T\vdash\neg\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_\tau(x)\rightarrow Tr_A(x))\rightarrow\exists y(\neg\beta(y)).$$

Now, by the assumption, $T\vdash\beta(\bar{0})$, hence

$$\begin{aligned}T\vdash\exists y(\neg\beta(y))&\rightarrow\exists y(\neg\beta(y')\wedge\forall z\leqq y\beta(z))\\&\rightarrow\exists y(\beta(y)\wedge\forall x(\tau(x)\wedge x\leqq y\leftrightarrow\tau'(x)))\\&\rightarrow\exists y(\forall z(\alpha(z)\wedge Sent(z)\wedge Pr_{\tau'}(z)\rightarrow Tr_A(z)).\end{aligned}$$

Thus we have

$$\begin{aligned}T\vdash\neg\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_\tau(x)\rightarrow Tr_A(x))&\\ \rightarrow\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_{\tau'}(x)\rightarrow Tr_A(x)).&\end{aligned} \tag{1}$$

On the other hand, by the definition of τ',

$$\begin{aligned}T\vdash\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_\tau(x)\rightarrow Tr_A(x))&\\ \rightarrow\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_{\tau'}(x)\rightarrow Tr_A(x)).&\end{aligned} \tag{2}$$

Combining (1) and (2), we have

$$T\vdash\forall x(\alpha(x)\wedge Sent(x)\wedge Pr_{\tau'}(x)\rightarrow Tr_A(x)).$$

So

$$T\vdash Pr_{\tau'}(\bar{\phi})\rightarrow\phi \quad \text{for all} \quad \phi\in A\cap Sent_T,$$

as desired. □

2.2. COROLLARY. *Suppose that T is an r.e. theory with the same language as PA. Then there is a theory T^* with $T^*=T$, and for each $n\in\omega$, there is a $\tau_n\in Bin(T^*, T^*)$ such that T^* prove each element of $Rfn_{\Sigma_n\cup\Pi_n}(\tau_n)$.*

PROOF. By Theorem 4.13 of [1], there is a theory T^* with $T^*=T$, and there is a $\tau\in Bin(T^*, T^*)$. So it is sufficient to prove that

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_{\tau\restriction\bar{m}}(x^*)\rightarrow Tr_n(x^*)) \quad \text{for all} \quad m, n\in\omega.$$

By formalizing a proof of the soundness of a 1-st order logic, we have

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_{\tau\restriction\bar{0}}(x^*)\to Tr_n(x^*))\quad\text{for all}\quad n\in\omega.$$

Let $m, n\in\omega$ be given. If $\tau\restriction\bar{m}$ is equivalent to $\bigvee_{i\leq l} x=\bar{\chi}_i$, then, for sufficiently large $j\geqq n$,

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_{\tau\restriction\bar{0}}(\overline{\bigwedge_i\chi_i\to x^*})\to Tr_j(\overline{\bigwedge_i\chi_i\underset{\cdot}{\to} x^*}))\,,$$

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_{\tau\restriction\bar{m}}(x^*)\to(\bigwedge_i\chi_i\to Tr_j(x^*)))\,,$$

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Pr_{\tau\restriction\bar{m}}(x^*)\to Tr_j(x^*))\,.$$

Since Tr_n and Tr_j are standard ones, we have

$$T\vdash\forall x\in\Sigma_n\cup\Pi_n\forall y(Tr_j(x^*)\to Tr_n(x^*))\,,$$

which completes our proof. □

If $T=PA$, we don't have to choose T^* as in the above corollary. So the following holds:

2.3. COROLLARY. *For each* $n\in\omega$, *there is a* $\pi_n\in Bin(PA, PA)$ *for which PA proves each element of* $Rfn_{\Sigma_n\cap\Pi_n}(\pi_n)$.

Now, we take another side view of the lower part of a reflection principle w. r. to the formula hierarchy.

2.4. THEOREM. *Suspose that* T *is an r. e. theory and* S *is a subtheory of* T. *Then, for each* $\tau\in Bin(T, S)$ *and* $n\in\omega$,

$$Rfn(\tau)\backslash Rfn_{T-(\Sigma_n\cup\Pi_n)}(\tau)=_T Rfn(\tau)\,.$$

To prove Theorem 2.4 we need some lemmas.

2.5. LEMMA (KENT). *If* T^* *is a consistent extension of an r. e. theory* T, *obtained by the addition of axioms in* $\Sigma_n\cup\Pi_n$, *in which each sentence of* $\Sigma_n\cup\Pi_n$ *is decidable, then* T^* *is incomplete.*

PROOF. See Theorem 3 of [3].

2.6. LEMMA. *Suppose that* T *is an r. e. theory and* $\phi_0, \phi_1\in \mathcal{S}ent_T$. *If* ϕ_0 *and* ϕ_1 *satisfy*

$$T\vdash\phi_0\to\phi_1\quad\&\quad T\nvdash\phi_1\to\phi_0\,,$$

then, for each $n\in\omega$, *there is a* $\chi_n\in\mathcal{S}ent_T$ *for which*

$$T\vdash\phi_0\to\phi_n\quad\&\quad T\vdash\chi_n\to\phi_1\quad\&\quad\chi_n\notin T-(\Sigma_n\cup\Pi_n)\,.$$

PROOF. By way of a contradiction, suppose that for an arbitrary $\phi \in \mathcal{S}ent_T$,

$$T \vdash \phi_0 \to \phi \quad \& \quad T \vdash \phi \to \phi_1 \quad \text{implies} \quad \phi \in T-(\Sigma_n \cup \Pi_n). \tag{1}$$

For each $\phi \in \mathcal{S}ent_T$, $(\phi_0 \vee \phi) \wedge \phi_1$ satisfies the left side of (1). Thus, for any ϕ, there is a $\sigma \in \Sigma_n \cup \Pi_n$ such that T proves $(\phi_0 \vee \phi) \wedge \phi_1 \leftrightarrow \sigma$, i.e.,

$$T \cup \{\neg \phi_0, \phi_1\} \vdash \phi \leftrightarrow \sigma. \tag{2}$$

Adding $\Sigma_n \cup \Pi_n$-sentences to consistent $T \cup \{\neg \phi_0, \phi_1\}$, we can construct a consistent theory T^* which is complete for $\Sigma_n \cup \Pi_n$-sentences. But (2) holds, therefore T^* must be complete. This contradicts the assertion of Lemma 2.5. □

Now we can prove Theorem 2.4.

PROOF OF THEOREM 2.4. It is sufficient to show that

$$T \cup Rfn(\tau) \backslash Rfn_{T-(\Sigma_n \cup \Pi_n)}(\tau) \vdash Pr_\tau(\bar{\phi}) \to \phi,$$

for $\phi \in \Sigma_n \cup \Pi_n$ such that $T \nvdash \phi$. Fix such a sentence $\phi \in \Sigma_n \cup \Pi_n$. By Lemma 2.6, there is a ϕ_1 for which

$$T \vdash \phi \to \phi_1 \quad \& \quad \phi_1 \notin T-(\Sigma_n \cup \Pi_n). \tag{3}$$

This ϕ_1 is unprovable in T. Hence, using Lemma 2.6 again, we can find ϕ_2 for which

$$T \vdash \phi \to \phi_2 \quad \& \quad T \vdash \phi_2 \to \phi \vee \neg \phi_1 \quad \& \quad \phi_2 \notin T-(\Sigma_n - \Pi_n). \tag{4}$$

Combining (3) and (4) yields

$$\phi_1, \phi_2 \notin T-(\Sigma_n \cup \Pi_n) \quad \& \quad T \vdash \phi \leftrightarrow \phi_1 \wedge \phi_2.$$

Therefore, we have

$$\begin{aligned} &(Pr_\tau(\overline{\phi_1}) \to \phi_1), \quad (Pr_\tau(\overline{\phi_2}) \to \phi_2) \in Rfn(\tau) \backslash Rfn_{T-(\Sigma_n \cup \Pi_n)}(\tau) \\ &\& \quad T \vdash (Pr_\tau(\overline{\phi_1}) \to \phi_1) \wedge (Pr_\tau(\overline{\phi_2}) \to \phi_2) \to (Pr_\tau(\bar{\phi}) \to \phi), \end{aligned}$$

as desired. □

The following is an easy consequence of the above theorem, and we can safely leave its proof to the reader.

2.7. COROLLARY. *Suppose that* T *is an r.e. theory and* S *is a subtheory of* T. *Suppose that* τ *is a binumeration of* T *in* S. *Then, for each finite subset* A *of* $Rfn(\tau)$,

$$Rfn(\tau) \backslash (T-A) =_T Rfn(\tau).$$

§ 3. The Ordering of The τ-Reflection Principles.

So far, we have investigated the behavior of the lower parts of the τ-reflection principles (w. r. the formula hierarchy). In this section, we compare the strength of the whole $Rfn(\tau)$'s as new axioms of T, and show that for a fixed ω-consistent and essentially reflexive theory T, there are no maximal elements and no minimal elements in $\{Rfn(\tau) \mid \tau \in Bin(T, S)\}$ w. r. to $\subsetneqq_T$.

3.1. THEOREM. *Suppoe that T is an essentially reflexive theory. Then for each $\tau \in Bin(T, S)$, we can effectively construct a $\tau' \in Bin(T, S)$ for which*

$$Rfn(\tau') \subsetneqq_T Rfn(\tau).$$

PROOF. A simple diagonal argument shows that there is a $(Pr_\tau(\bar{\sigma}) \rightarrow \sigma) \in Rfn(\tau)$ which is not provable in T. Using this $Pr_\tau(\bar{\sigma}) \rightarrow \sigma$, we define $\alpha(x)$ by

$$\alpha(x) = \tau(x) \vee x = \underset{\cdot}{\neg}(\overline{Pr_\tau(\bar{\sigma}) \rightarrow \sigma}).$$

For each formula $\gamma(x)$, define $f_\gamma(m)$ by

$$f_\gamma(m) = \bigwedge_{\phi \leq m} (Pr_\gamma(\bar{\phi}) \rightarrow \phi).$$

Using these, we define a diagonal sentence ϕ_n such that

$$PA \vdash \phi_n \leftrightarrow \forall x(Prf_\alpha(\bar{\phi}_n, x) \rightarrow \neg f_{\tau \restriction x}(n)).$$

ϕ_n can be constructed effectively from n (in fact primitive recursively from n). Hence, there is a corresponding p. r. function symbol $\dot{\phi}$ such that $PA \vdash \bar{\phi}_n = \dot{\phi}(\bar{n})$. Now define $\tau'(x)$ by

$$\tau'(x) = \tau(x) \wedge \forall y, z \leqq x \neg Prf_\alpha(\dot{\phi}(y), z).$$

First we prove that

$$T \nvdash \neg(Pr_\tau(\bar{\sigma}) \rightarrow \sigma) \rightarrow \phi_n \quad \text{for all} \quad n \in \omega. \tag{1}$$

Assume that $T \vdash \neg(Pr_\tau(\bar{\sigma}) \rightarrow \sigma) \rightarrow \phi_n$, then $T \vdash Prf_\alpha(\bar{\phi}_n, \bar{m})$ for some $m \in \omega$. Hence, using the definition of ϕ_n, we have $T \vdash \neg(Pr_\tau(\bar{\sigma}) \rightarrow \sigma) \rightarrow \neg f_{\tau \restriction \bar{m}}(n)$. On the other hand, by the essential reflexiveness of T, $T \vdash f_{\tau \restriction \bar{m}}(n)$. So $T \vdash Pr_\tau(\bar{\sigma}) \rightarrow \sigma$. But this is a contradiction, which leads us to conclude (1). Next we prove that

$$T \vdash \neg \phi_n \rightarrow f_{\tau'}(n) \quad \text{for all} \quad n \in \omega.$$

Since we can assume $T \vdash \forall y(\dot{\phi}(y) > y)$, $T \vdash \forall y \geqq x \forall z \leqq x \neg Prf_\alpha(\dot{\phi}(y), z)$. So we have

$$T \vdash \forall x(\tau'(x) \rightarrow \tau(x) \wedge \forall z \leqq x \neg Prf_\alpha(\bar{\phi}_n, z)) \quad \text{for all} \quad n \in \omega.$$

Now, using the definition of ϕ_n, we have

$$\begin{aligned} T\vdash \neg\phi_n &\to \exists x(Prf_\alpha(\overline{\phi_n}, x) \wedge f_{\tau\upharpoonright x}(n)) \\ &\to \exists x(f_{\tau\upharpoonright x}(n) \wedge \forall y(\tau'(y) \to \tau(y) \wedge y \leqq x)) \\ &\to f_{\tau'}(n)\,, \end{aligned}$$

as desired. From (1) and (2), we conclude that

$$T \nvdash f_{\tau'}(n) \to (Pr_\tau(\bar{\sigma}) \to \sigma) \quad \text{for all} \quad n \in \omega.$$

This directly gives $Rfn(\tau') \subsetneqq_T Rfn(\tau)$. What is left to prove is that τ' binumerates T in S. But this is easily proved by (1). □

Theorem 3.1 asserts that $\{Rfn(\tau) \mid \tau \in Bin(T, S)\}$ has no minimal elements w.r. to $\subsetneqq_T$, if T is essentially reflexive. Maximal elements also do not exist, if T is ω-consistent.

3.2. THEOREM. *Suppose that T is an ω-consistent theory. Then, for each $\tau \in Bin(T, S)$, we can effectively construct a $\tau' \in Bin(T, S)$ for which*

$$Rfn(\tau) \subsetneqq_T Rfn(\tau')\,.$$

PROOF. Set $T' = T \cup Rfn(\tau)$. Then Theorem 1.6 guarantees the consistency of T'. Let $\beta'(x)$ be a Σ_0-formula which binumerates $Rfn(\tau)$ in PA. Using this β' define $\beta(x)$ by

$$\beta(x) = \tau(x) \vee \beta'(x)\,.$$

Clearly, β is a binumeration of T' in S. By Gödel's theorem,

$$T' \nvdash \nu_\beta\,, \tag{1}$$

where ν_β is a fixed point of $\neg Pr_\beta(x)$. Next define $\tau'(x)$ by

$$\tau'(x) = \tau(x) \vee Fm(x) \wedge \exists y < x Prf_\beta(\overline{\nu_\beta}, y)\,.$$

Then τ' is a binumeration of T in S. Since

$$\begin{aligned} PA \vdash \neg\nu_\beta &\to \exists x Prf_\beta(\overline{\nu_\beta}, x) \\ &\to \exists y \exists x < \underset{\cdot}{\neg}(vr_y \underset{\cdot}{=} vr_y) Prf_\beta(\overline{\nu_\beta}, x) \\ &\to \exists y \tau'(\underset{\cdot}{\neg}(vr_y \underset{\cdot}{=} vr_y)) \\ &\to \neg Con_{\tau'}\,, \end{aligned}$$

we have $PA \vdash Con_{\tau'} \to \nu_\beta$. This together with (1) implies

$$T' \nvdash Con_{\tau'}\,.$$

Thus we have $Rfn(\tau) \subsetneqq_T Rfn(\tau')$ as desired. □

The following theorem shows that $\subsetneqq_T$ is a dense ordering of

$$\{Rfn(\tau) \mid \tau \in Bin(T, S)\}.$$

3.3 THEOREM. *If* $\tau_1, \tau_2 \in Bin(T, S)$ *and* $Rfn(\tau_1) \subsetneqq_T Rfn(\tau_2)$, *then there is a* $\tau' \in Bin(T, S)$ *for which*

$$Rfn(\tau_1) \subsetneqq_T Rfn(\tau') \subsetneqq_T Rfn(\tau_2).$$

PROOF. Since $Rfn(\tau_1) \subsetneqq_T Rfn(\tau_2)$, there is a sentence $\phi \in \mathrm{Sent}_T$ for which $Pr_{\tau_2}(\bar{\phi}) \to \phi$ is not provable in $T \cup Rfn(\tau_1)$. For this ϕ, set $T' = T \cup Rfn(\tau_1) \cup \{\neg(Pr_{\tau_2}(\bar{\phi}) \to \phi)\}$. Then T' is a consistent theory. Let $\beta'(x)$ be a Σ_0-formula which binumerates $Rfn(\tau_1)$ in PA, and define $\beta(x)$ by

$$\beta(x) = \tau_1(x) \vee \beta'(x) \vee x = \dot{\neg}(\overline{Pr_{\tau_2}(\bar{\phi}) \to \phi}).$$

Clearly, β is a binumeration of T' in S. If we set

i) $\theta(x, y) = Prf_{\beta}(\dot{\neg} y, x) \wedge \forall z \leqq x \neg Prf_{\beta}(y, z)$,

ii) χ: a fixed point of $\exists x \theta(x, y)$,

iii) $\tau'(x) = \tau_1(x) \vee Fm(x) \wedge \exists y_1, y_2 < x(\theta(y_1, \bar{\chi}) \wedge Prf_{\tau_2}(\bar{\phi}, y_2) \wedge \neg\phi)$, then τ' is a binumeration of T in S. By Rosser's theorem,

$$T \nvdash \chi, \tag{1}$$

$$T \nvdash \neg\chi. \tag{2}$$

First we show that

$$PA \vdash (Pr_{\tau'}(\bar{\phi}) \to \phi) \wedge \chi \to (Pr_{\tau_2}(\bar{\phi}) \to \phi). \tag{3}$$

Note that

$$\begin{aligned} PA \vdash \theta(y_1, \bar{\chi}) \wedge Prf_{\tau_2}(\bar{\phi}, y_2) \wedge \neg\phi \to y_1, y_2 < (\dot{\neg}(vr_{y_1} \dot{=} vr_{y_1}) \dot{\wedge} \dot{\neg}(vr_{y_2} \dot{=} vr_{y_2})) \\ \wedge \theta(y_1, \bar{\chi}) \wedge Prf_{\tau_2}(\bar{\phi}, y_2) \wedge \neg\phi \\ \to \tau'(\dot{\neg}(vr_{y_1} \dot{=} vr_{y_1}) \dot{\wedge} \dot{\neg}(vr_{y_2} \dot{=} vr_{y_2})). \end{aligned}$$

Then, clearly, we have

$$PA \vdash \exists y_1 \theta(y_1, \bar{\chi}) \wedge Pr_{\tau_2}(\bar{\phi} \wedge \neg\phi \to Pr_{\tau'}(\bar{\phi}).$$

This directly gives (3). Next we show that

$$PA \vdash (Pr_{\tau_1}(\bar{\phi}) \to \phi) \wedge \neg\chi \to (Pr_{\tau'}(\bar{\phi}) \to \phi) \quad \text{for all} \quad \phi \in \mathcal{S}ent_T. \tag{4}$$

But it is sufficient for this purpose to show that

$$PA \vdash \neg\chi \to \forall x(\tau'(x) \leftrightarrow \tau_1(x)).$$

And this is verified by the fact that $PA \vdash \neg\chi \to \forall x \neg\theta(x, \bar{\chi})$. Now, from (2) and (3), we have

$$T \cup Rfn(\tau_1) \nvdash Pr_{\tau'}(\bar{\phi}) \to \phi \,. \tag{5}$$

On the other hand, from (1) and (4),

$$T \cup Rfn(\tau') \nvdash Pr_{\tau_2}(\bar{\phi}) \to \phi \,. \tag{6}$$

Since $Rfn(\tau_1) \subseteqq_T Rfn(\tau') \subseteqq_T Rfn(\tau_2)$ is clear from the equivalence of τ' and τ_1 over $T \cup Rfn(\tau_2)$, (5) and (6) will complete our proof. □

NOTES TO THEOREMS 3.1, 3.2 AND 3.3. In the first two theorems, if τ is a Σ_0-binumeration of T, we can choose τ' in Σ_0. Hence there are also infinitely many kinds of the τ-reflection principles, even if τ is restricted to Σ_0. The author doesn't know whether τ' of Theorem 3.3 can be chosen in Σ_0 if both τ_1 and τ_2 are given in Σ_0.

The following theorem shows that there are incomparable elements in $\{Rfn(\tau) \mid \tau \in BinT, S)\}$ w. r. to $\subsetneqq_T$. Since its proof is very similar to that of Theorem 2.14 of [2], we shall give only a sketch of the proof.

3.4. THEOREM. *Suppose thrt T is an essentially reflexive and ω-consistent theory. Then, for each $\tau \in Bin(T, S)$, we can effectively construct a $\tau' \in Bin(T, S)$ for which*

$$Rfn(\tau) \nsubseteqq_T Rfn(\tau') \quad \& \quad Rfn(\tau') \nsubseteqq_T Rfn(\tau) \,.$$

SKETCH OF THE PROOF. Using Theorem 3.1, we choose $\sigma \in \mathcal{S}ent_T$ and $\tau^* \in Bin(T, S)$ such that

$$Rfn(\tau^*) \subsetneqq_T Rfn(\tau) \quad \& \quad T \cup Rfn(\tau^*) \vdash Pr_\tau(\bar{\sigma}) \to \sigma \,.$$

Putting $A_1 = T \cup Rfn(\tau)$ and $A_2 = T \cup Rfn(\tau^*) \cup \{\neg(Pr_\tau(\bar{\sigma}) \to \sigma)\}$, we construct $\tau_1 \in Bin(A_1, S)$ and $\tau_2 \in Bin(A_2, S)$. Then we construct τ' as

$$\tau'(x) = \tau^*(x) \vee Fm(x) \wedge \exists y < x \neg M_{\alpha_1 \alpha_2}(\bar{\mu}, y) \,,$$

where $M_{\alpha_1 \alpha_2}(x, y) = (Prf_{\alpha_1}(x, y) \vee Prf_{\alpha_2}(x, y)) \to \exists z < y (Prf_{\alpha_1}(\underset{\cdot}{\neg} x, z) \vee Prf_{\alpha_2}(\underset{\cdot}{\neg} x, z))$ and μ is a fixed point of $\forall y M_{\alpha_1 \alpha_2}(x, y)$. If we note that μ is independent of A_1 and A_2, we can easily show the desired properties of τ'. □

DISCUSSIONS. The well-known theorem of Löb states that $Pr_\tau(\bar{\phi}) \to \phi$ is provable in T iff ϕ is provable in T. This does not contradict our results. The reason is that Löb's theorem holds only for r. e. theories T and their Σ_1-numerations τ. There are many results that hold only for r. e. theories, and it would be interesting to give examples of them. But we do not go further into these matters.

In this paper, we largely dealt with local reflection principles. The analogous results for uniform reflection principles will be contained in our forthcoming paper.

References

[1] Feferman, S., Arithmetization of metamathematics in a general setting. Fund. Math. **49** (1960), 35-92.
[2] Hájková, M., The lattice of binumerations of arithmetic I. Comment. Math. Univ. Carolinae **12** (1971), 81-104.
[3] Kent, C.F., The relation of A to Prov ⌜A⌝ in the Lindenbaum sentence algebra. J. Symbolic Logic **38** (1973), 295-298.
[4] Kreisel, G. and Lévy, A., Reflection principle and their use for establishing the complexity of axiomatic systems. Z. Math. Logik Grundlagen Math. **14** (1968), 97-142.
[5] Smorynski, C., The incompleteness theorems. Handbook of mathematical logic (Barwise, K.J., Editor), North-Holland, Amsterdam (1977), 821-866.

Department of Mathematics
University of Tsukuba
Sakura-mura, Niihari-gun
Ibaraki, 305 Japan