Incompleteness theorems and S_2^i versus S_2^{i+1}

Gaisi Takeuti

In [2], S. Buss introduced the systems of Bounded Arithmetic for $S_2^i(i = 0, 1, 2, \ldots)$ which have a close relationship to classes in polynomial hierarchy. As Buss stated in the introduction of his book, one of the most important problems on Bounded Arithmetic is the separation problems on $S_2^i(i = 1, 2, 3, \ldots)$ i.e., the problems to show $S_2^i \neq S_2^{i+1}(i = 1, 2, 3, \ldots)$. We believe that the separation problems of S_2^i and the separation problems of classes of the polynomial hierarchy are the same problem in the sense that the difficulty of these two problems comes from the same source. We also believe that the solution of one of them will lead to the solution of the other problem.

This idea is partially supported by the following theorem in [4].

Theorem. (Krajíček-Pudlák-Takeuti) *If* $S_2^i = S_2^{i+1}$, *then* $\Sigma_{i+2}^P = \Pi_{i+2}^p$.

Very often, a stronger theory is shown to be strictly stronger than a weaker theory by proving that the stronger theory proves the consistency of the weaker system. In [2], S. Buss proved that the second incompleteness theorem also holds for S_2^i. However this method does not work for the separations of S_2^i since the theorem of Wilkie and Paris in [6] immediately implies that $S_2 = \bigcup_i S_2^i$ does not prove the consistency of S_2^0. The reason for this phenomenon is that the consistency here is the consistency of the theory with unbounded quantifiers. The expressing power of unbounded quantifiers is too strong to be handled by Bounded Arithmetic. The ordinary consistency is totally inadequate for Bounded Arithmetic. We need some more delicate consistency.

In [5], we introduced a delicate notion of proof in S_2^i and delicate notions of consistency of S_2^i and Gödel sentences of S_2^i. Using them we proved that a Gödel sentence of S_2^i is provable in $S_{2,n}^{i+1}$ though it is not provable in S_2^i, therefore $S_{2,n}^{i+1} \neq S_2^i$ holds in the language of S_2^i. S_2^{i+1} is a limit of $S_{2,n}^{i+1}$ if n goes to ∞ but this result does not imply $S_2^{i+1} \neq S_2^i$.

[1]Received August 96; revised version December 1996.

In this paper we further develop the idea of [5] and propose many conjectures on delicate consistency and delicate Gödel sentences of S_2^i which imply $S_2^{i+1} \neq S_2^i$. We believe that a little more advance on the knowledge of these consistencies and Gödel sentences would prove $S_2^i \neq S_2^{i+1}$ and $P \neq NP$.

1 The formalized terms

We define $|a|_n$ for $n = 0, 1, 2, \ldots$ by $|a|_0 = a$ and $|a|_{n+1} = |\,|a|_n\,|$. In [5], we expand the language of S_2^i by introducing $a \dot{-} b$, $\max(a, b), \ldots, \beta(i, w)$. In this paper we further expand the language by adding finitely many function symbols whose intended meanings are polynomial time computable functions.

Let $\tilde{n}$ be the Gödel number of a formalized term in the language of S_2^i with only free variables $\ulcorner a_1 \urcorner, \ldots, \ulcorner a_n \urcorner$ where $n = 0, 1, 2, \ldots$. In this case, we often denote $\tilde{n}$ by $\ulcorner t(a_1, \ldots, a_n) \urcorner$ though we cannot find a term $t(a_1, \ldots, a_n)$ in general. Let $v(\ulcorner t(a_1, \ldots, a_n) \urcorner, b_1, \ldots, b_n)$ be the value which $\ulcorner t(a_1, \ldots, a_n) \urcorner$ represents when $\ulcorner a_i \urcorner$ represents b_i respectively. Let $\vec{a}$ and $\vec{b}$ express $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ respectively. In [5], we proved that $\exp(\ulcorner t(\vec{a}) \urcorner |b|^{|\ulcorner t(\vec{a}) \urcorner|})$ is a bound of $v(\ulcorner t(\vec{a}) \urcorner, \vec{b})$ where $\exp(a) = 2^a$ and b is the maximum of $\vec{b}$ and 2 and also that

$$v(\ulcorner t \urcorner) \leq \exp(\ulcorner t \urcorner)$$

if t is the Gödel number of a closed term.

Let f be a new polynomial time computable function whose function symbol is in the language and $f(a) \leq a \# \ldots \# a$, where the number of a is n. Let $t(a) = f(t_0(a))$. We are going to prove

$$v(\ulcorner t(a) \urcorner, b) \leq \exp(\ulcorner t(a) \urcorner |b|^{|\ulcorner t(a) \urcorner|})$$

under the assumption $v(\ulcorner t_0(a) \urcorner, b) \leq \exp(\ulcorner t_0(a) \urcorner |b|^{|\ulcorner t_0(a) \urcorner|})$

$$\begin{aligned}
v(\ulcorner t(a) \urcorner, b) &\leq v(\ulcorner t_0(a) \urcorner, b) \# \ldots \# v(\ulcorner t_0(a) \urcorner, b) \\
&\leq \exp(\ulcorner t_0(a) \urcorner |b|^{|\ulcorner t_0(a) \urcorner|}) \# \ldots \# \exp(\ulcorner t_0(a) \urcorner |b|^{\ulcorner t_0(a) \urcorner}) \\
&\leq \exp((\ulcorner t_0(a) \urcorner |b|^{|\ulcorner t_0(a) \urcorner|} + 1)^n) \\
&\leq \exp(c(\ulcorner t_0(a) \urcorner |b|^{|\ulcorner t_0(a) \urcorner|})^n) \\
&\quad \text{for some constant } c \\
&\leq \exp(\ulcorner t(a) \urcorner \cdot |b|^{|\ulcorner t(a) \urcorner|})
\end{aligned}$$

where we define $\ulcorner f \urcorner$ to be sufficiently large and $f(a)$ to be an abbreviation of $f(a, \ldots, a)$. More precisely we define $\ulcorner f(t) \urcorner$ to be the sequence number of $(\ulcorner f \urcorner, \ulcorner t \urcorner, \ldots, \ulcorner t \urcorner)$ where $\ulcorner t \urcorner$ occurs n times for a fixed number n so that the above calculation goes through. It is very easy to find such $\ulcorner f \urcorner$ and n for every polynomial time computable function f.

2 Truth definition

We say 'a is n-small' if there exists x such that $a \leq |x|_n$. We say 'a is small' if a is 1-small. In this section we always assume $i > 0$. Let u satisfy the following condition. $|u|_2$ is greater than the Gödel number $\ulcorner t(\vec{a}) \urcorner$ of a term $t(\vec{a})$ in the language of S_2^i with only free variables $\vec{a}$. The length of $\vec{a}$ is less than $|\ulcorner t(\vec{a}) \urcorner|$ and it is 3-small. Let $\vec{b}$ be a sequence with the same length as $\vec{a}$. As before we define $b = \max(2, \vec{b})$. As in Section 1, we define $v(\ulcorner t(\vec{a}) \urcorner, \vec{b})$ and the following holds.

$$v(\ulcorner t(\vec{a}) \urcorner, \vec{b}) \leq \exp(|u|_2 |b|^{|u|_3}).$$

Here $\exp(|u|_2 |b|^{|u|_3})$ is a Σ_1^b-definable function in S_2^1 when b is small.

The system S_2^i is defined by the following axioms and inferences.

(a) Basic axioms.

The language of S_2^i consists of $\leq$ and finitely many function symbols which express polynomial time computable functions.

The defining axioms of the functions in the language and the predicate $\leq$. All these axioms are included in the form of initial sequents without logical symbols.

(b) Σ_i^b - PIND

$$\frac{A(\lfloor \frac{1}{2} a \rfloor), \Gamma \rightarrow \Delta, A(a)}{A(0), \Gamma \rightarrow \Delta, A(t)}$$

where $A(0)$ is Σ_i^b and a satisfies an eigenvariable condition. We further extend a) by introducing finitely many forms of initial sequents without logical symbols. E.g.

$$\longrightarrow |s| \leq s.$$

This saves unnecessary use of induction in order to prove some necessary properties. Here the following must be satisfied.

1. The number of the forms of initial sequents must be finite.

2. The initial sequent thus introduced must be valid and has no occurrences of logical symbols. As a stronger case of this type of extension, later we also consider the following S_2^i under the assumption of $P = NP$.

If $P = NP$, there exists an NP-complete predicate $\exists x \leq t(a)A(x,a)$ with a sharply bounded $A(x,a)$ and a polynomial time computable function f satisfying the following condition

$$\exists x \leq t(a)A(x,a) \rightarrow f(a) \leq t(a) \wedge A(f(a),a).$$

If we introduce finitely many polynomial time computable functions, then we can assume that $A(x,a)$ is an atomic formula and $P = NP$ can be expressed by the following forms of initial sequents without logical symbols

$$s' \leq t(s), A(s', t(s)) \rightarrow f(s) \leq t(s)$$
$$s' \leq t(s), A(s', t(s)) \rightarrow A(f(s), s).$$

Our theory developed later also holds for extended S_2^i in this way and will be used to find conjectures to prove $P \neq NP$. The outline of this plan goes as follows.

$$P = NP \rightarrow \Sigma_i^b = \Sigma_{i+1}^b \dashrightarrow S_2^i = S_2^{i+1}.$$

Therefore $S_2^i \neq S_2^{i+1} \dashrightarrow P \neq NP$. Here $\dashrightarrow$ holds if S_2^i and S_2^{i+1} are extended S_2^i and S_2^{i+1} discussed in the above.

In [5], it is proved that $v(\ulcorner t(\vec{a})\urcorner, \vec{b})$ is definable under the assumption that $\exp(|u|_2 |b|^{|u|_3})$ is definable where $\ulcorner t(\vec{a})\urcorner \leq |u|_2$ and $b = \max(2, \vec{b})$. In the same way, we can show that $v(\ulcorner t(\vec{a})\urcorner, \vec{b})$ is Σ_1^b-definable in S_2^1 and satisfies the following conditions if $\ulcorner t(\vec{a})\urcorner \leq |u|_2$ and b is small.

1. If $\ulcorner f(t_1(\vec{a}), \ldots, t_k(\vec{a}))\urcorner$ is a subterm of $\ulcorner t(\vec{a})\urcorner$, then

$$v(\ulcorner f(t_1(\vec{a}), \ldots, t_k(\vec{a})\urcorner, \vec{b}) = f(v(\ulcorner t_1(\vec{a})\urcorner, \vec{b}), \ldots, v(\ulcorner t_n(\vec{a})\urcorner, \vec{b})).$$

2. $v(\ulcorner 0\urcorner, \vec{b}) = 0$ and $v(\ulcorner a_i\urcorner, \vec{b}) = b_i$.

All these properties are provable in S_2^1.

Let $\ulcorner t\urcorner$ be a formalized closed term and small. Then in the same way as above, $v(\ulcorner t\urcorner)$ is Σ_1^b-definable function in S_2^1 and satisfies the following conditions.

1. $v(\ulcorner 0\urcorner) = 0$.
2. $v(\ulcorner f(t_1, \ldots, t_n)\urcorner) = f(v(\ulcorner t_1\urcorner), \ldots, v(\ulcorner t_n\urcorner))$.

These properties are provable in S_2^1.

Now let u satisfy the following condition. $|u|_2$ is greater than the Gödel number $\ulcorner \varphi(\vec{a})\urcorner$ of a quantifier free formula in the language of S_2^i with only free variables $\vec{a}$. Let $\vec{b}$ be a sequence with the same length as $\vec{a}$ and $b = \max(\vec{b}, 2)$ be small. In [5], the truth definition of $T_0(\ulcorner \varphi(\vec{a})\urcorner, \vec{b})$ was defined by using $\exp(|u|_2 |b|^{|u|_3})$ is definable. The following properties were proved.

1. If $\ulcorner t_1 \leq t_2 \urcorner$ is a subformula of $\ulcorner \varphi(\vec{a}) \urcorner$, then
$$T_0(\ulcorner t_1 \leq t_2 \urcorner, \vec{b}) \text{ iff } v(\ulcorner t_1 \urcorner, \vec{b}) \leq v(\ulcorner t_2 \urcorner, \vec{b}).$$
2. If $\ulcorner t_1 = t_2 \urcorner$ is a subformula of $\ulcorner \varphi(\vec{a}) \urcorner$, then
$$T_0(\ulcorner t_1 = t_2 \urcorner, \vec{b}) \text{ iff } v(\ulcorner t_1 \urcorner, \vec{b}) = v(\ulcorner t_2 \urcorner, \vec{b}).$$
3. If $\ulcorner \psi_1 \wedge \psi_2 \urcorner$ is a subformula of $\ulcorner \varphi(\vec{a}) \urcorner$, then
$$T_0(\ulcorner \psi_1 \wedge \psi_2 \urcorner, \vec{b}) \text{ iff } T_0(\ulcorner \psi_1 \urcorner, \vec{b}) \wedge T_0(\ulcorner \psi_2 \urcorner, \vec{b}).$$
4. If $\ulcorner \psi_1 \vee \psi_2 \urcorner$ is a subformula of $\ulcorner \varphi(\vec{a}) \urcorner$, then
$$T_0(\ulcorner \psi_1 \vee \psi_2 \urcorner, \vec{b}) \text{ iff } T_0(\ulcorner \psi_1 \urcorner, \vec{b}) \vee T_0(\ulcorner \psi_2 \urcorner, \vec{b}).$$
5. If $\ulcorner \neg\psi \urcorner$ is a subformula of $\ulcorner \psi(\vec{a}) \urcorner$, then
$$T_0(\ulcorner \neg\psi \urcorner, \vec{b}) \text{ iff } \neg T_0(\ulcorner \psi \urcorner, \vec{b}).$$

$T_0(\ulcorner \varphi(\vec{a}) \urcorner, \vec{b})$ is Σ_1^b-definable in S_2^1 and all these properties are provable in S_2^1.

In the same way, we can Σ_1^b-define a truth definition $T_0(\ulcorner \varphi \urcorner)$ in S_2^1 if $\ulcorner \varphi \urcorner$ is the Gödel number of a quantifier free sentence in S_2^1 and $\ulcorner \varphi \urcorner$ is small and the following properties are provable in S_2^1.

1. If $\ulcorner t_1 \leq t_2 \urcorner$ is a subformula of $\ulcorner \psi \urcorner$, then
$$T_0(\ulcorner t_1 \leq t_2 \urcorner) \text{ iff } v(\ulcorner t_1 \urcorner) \leq v(\ulcorner t_2 \urcorner).$$
2. If $\ulcorner t_1 = t_2 \urcorner$ is a subformula of $\ulcorner \varphi \urcorner$, then
$$T_0(\ulcorner t_1 = t_2 \urcorner) \text{ iff } v(\ulcorner t_1 \urcorner) = v(\ulcorner t_2 \urcorner).$$
3. If $\ulcorner \psi_1 \wedge \psi_2 \urcorner$ is a subformula of $\ulcorner \psi \urcorner$, then
$$T_0(\ulcorner \psi_1 \wedge \psi_2 \urcorner) \text{ iff } T_0(\ulcorner \psi_1 \urcorner) \wedge T_0(\ulcorner \psi_2 \urcorner).$$
4. If $\ulcorner \psi_1 \vee \psi_2 \urcorner$ is a subformula of $\ulcorner \varphi \urcorner$, then
$$T_0(\ulcorner \psi_1 \vee \psi_2 \urcorner) \text{ iff } T_0(\ulcorner \psi_1 \urcorner) \vee T_0(\ulcorner \psi_2 \urcorner).$$
5. If $\ulcorner \neg\psi \urcorner$ is a subformula of $\ulcorner \psi \urcorner$, then
$$T_0(\ulcorner \neg\psi \urcorner) \text{ iff } \neg T_0(\ulcorner \psi \urcorner).$$

A formula $\varphi(\vec{a})$ in the language of S_2^1 is said to be a pure i-form if the following conditions are satisfied.

1. The only free variables in $\varphi(\vec{a})$ are $\vec{a}$.

2. $\varphi(\vec{a})$ is of the form

$$\exists x_1 \leq t_1(\vec{a})\forall x_2 \leq t_2(\vec{a}, x_2)\dots Q_i x_i \leq t_i(\vec{a}, x_1, \dots, x_{i-1})$$
$$Q_{i+1}x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \dots, x_i)| A(\vec{a}, x_1, \dots, x_{i+1})$$

where Q_i is $\forall$ and Q_{i+1} is $\exists$ if i is even and Q_i is $\exists$ and Q_{i+1} is $\forall$ if i is odd and $A(\vec{a}, a_1, \dots, a_{i+1})$ is a quantifier free formula in the language of S_2^1. The formula described in the above is denoted by

$$\exists x_1 \leq t_1 \dots Q_{i+1}x_{i+1} \leq |t_{i+1}| A(\vec{a}, \vec{x}).$$

Since S_2^1 has $\beta(i, a)$, every Σ_i^p-formula with only free variables $\vec{a}$ is equivalent to a pure i-form. The formalized notion of pure i-form, i.e., "x is a Gödel number of pure i-form formula" is Δ_1^b with respect to S_2^1.

A formula $\varphi(\vec{a})$ in the language of S_2^1 is said to be of i-form if the following conditions are satisfied.

1. The only free variables in $\varphi(\vec{a})$ are $\vec{a}$.

2. $\varphi(\vec{a})$ is a subformula of a pure i-form formula, i.e., it is of the form

$$Q_j x_j \leq t_j \dots Q_{i+1}x_{i+1} \leq |t_{i+1}| A(\vec{a}, \vec{x})$$

where $A(\vec{a}, \vec{x})$ is quantifier free and Q_k is $\forall$ if k is even and Q_k is $\exists$ if k is odd and t_k is of the form $t_k(\vec{a}, x_j, \dots, x_{k-1})$. Quantifier free formulas and formulas of the form $Q_{i+1}x_{i+1} \leq |t_{i+1}(\vec{a})| A(\vec{a}, x_{i+1})$ are included as special cases of i-form formulas.

If $\ulcorner\varphi(\vec{a})\urcorner$ is the Gödel number of an i-form formula, then $\varphi(\vec{a})$ is of the form

$$Q_j x_j \leq t_j \dots Q_i x_i \leq t_i Q_{i+1}x_{i+1} \leq |t_{i+1}| A(\vec{a}, \vec{x})$$

and j is calculated from $\ulcorner\varphi(\vec{a})\urcorner$.

We are going to define $\widetilde{T}(\widetilde{u}, \ulcorner\varphi(\vec{a})\urcorner, \vec{a})$, which is a truth definition of $\ulcorner\varphi(\vec{a})\urcorner$ by assigning the value a_i to $\ulcorner a_i \urcorner$.

Later we will give a condition such that all the terms occurring in the computation of $\widetilde{T}(\widetilde{u}, \ulcorner\varphi(\vec{a})\urcorner, \vec{a})$ are bounded by $\widetilde{u}$. Under this condition, $\widetilde{T}(\widetilde{u}, \ulcorner\varphi(\vec{a})\urcorner, \vec{a})$ is defined to be $\bigwedge_j$ (if $\ulcorner\varphi(\vec{a})\urcorner$ is of the form

$$\ulcorner Q_j x_j \leq t_j \dots Q_{i+1}x_{i+1} \leq |t_{i+1}| A(\vec{a}, \vec{x})\urcorner,$$

then

$$\begin{array}{ll} Q_j x_j \leq v(\ulcorner t_j \urcorner, \vec{a}) & \ldots Q_{i+1} x_{i+1} \leq |v(\ulcorner t'_{i+1} \urcorner, \vec{a}, x_j, \ldots, x_i) \\ \quad T_0(\ulcorner A(\vec{a}, b_j, & \quad \ldots, b_i \urcorner), \vec{a}, x_j, \ldots, x_i)), \end{array}$$

where we denote $t_k(\vec{a}, b_j, \ldots, b_{k-1})$ by t'_k. Then $\widetilde{T}(\widetilde{u}, \ulcorner \varphi(\vec{a}) \urcorner, \vec{a})$ is Σ_i^b in S_2^1.

In [5], we defined strictly i-normal proof and i-normal proof. This notion is very useful to evaluate the proof in $S_{2,n}^{i+1}$. Now we need a stronger notion since we would like to replace $S_{2,n}^{i+1}$ by S_2^{i+1}. In this paper we call this stronger notion i-normal proof.

Definition. *A proof P in S_2^i is said to be strictly i-normal if the following conditions are satisfied.*

1. *Every formula in P is i-normal.*

2. *P is in free variable normal form.*

3. *Let $\vec{c}$ be all parameter variables in P and $\vec{b}$ be an enumeration of all other free variables in P satisfying the condition that if the elimination inference for b_i is below the elimination inference for b_j then $i < j$. There exists an assignment $t_i(\vec{c})$ for b_i satisfying the following conditions.*

   (a) $t_i(\vec{c})$ is a term in the language of S_2^1 and all function symbols occurring in $t_i(\vec{c})$ are function symbols of increasing functions.

   (b) If the elimination inference of b_i is

   $$\frac{A(\lfloor \frac{1}{2} b_i \rfloor), \Gamma \rightarrow \Delta, A(b_i)}{A(0), \Gamma \rightarrow \Delta, A(t(b_1, \ldots, b_{i-1}, \vec{c}))}$$

   or

   $$\frac{b_i \leq t(b_1, \ldots, b_{i-1}, \vec{c}), A(b_i), \Gamma \rightarrow \Delta}{\exists x \leq t(b_1, \ldots, b_{i-1}, \vec{c}) A(x), \Gamma \rightarrow \Delta}$$

   or

   $$\frac{b_i \leq t(b_1, \ldots, b_{i-1}, \vec{c}), \Gamma \rightarrow \Delta, A(b_i)}{\Gamma \rightarrow \Delta, \forall x \leq t(b_1, \ldots, b_{i-1}, \vec{c}) A(x)}$$

   then $a_1 \leq t_1(\vec{c}), \ldots, a_{i-1} \leq t_{i-1}(\vec{c}) \rightarrow t(a_1, \ldots, a_{i-1}, \vec{c}) \leq t_i(\vec{c})$ is provable without using logical inference, induction, or any free variables other than $a_1, \ldots, a_{i-1}$ and $\vec{c}$. All the information for condition 3) is called a supplementary proof. The proof P includes all these supplementary proofs.

4. *A sequence $\ldots, t(t_1(\vec{c}), \ldots, t_i(\vec{c}), \vec{c}), \ldots$ is provided where $t(b_1, \ldots, b_i, \vec{c})$ ranges over all terms in the proof.*

Precisely P together with all supplementary proofs and the sequence of terms described in 4) is called strictly i-normal proof. Let $\Gamma \to \Delta$ be provable in S_2^i, where all formulas in Γ and Δ are i-normal. Then we first make a free cut free proof of $\Gamma \to \Delta$ in S_2^i and then we can easily make a strictly i-normal proof of $\Gamma \to \Delta$.

Let $\varphi(a)$ be an i-form formula. Then a proof P of $\forall x \neg\varphi(x)$ is said to be i-normal if P is obtained from a strictly i-normal proof P_0 in the following way.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\ P_0\\ \\ \varphi(a)\ \to\end{array}}{\to \neg\varphi(a)}}{\to \forall x \neg\varphi(x)}$$

We denote the formalized notion "w is a strictly i-normal proof of $\ulcorner\Gamma \to \Delta\urcorner$" and "$w$ is an i-normal proof of $\ulcorner\forall x\neg\varphi(x)\urcorner$ by $Prf^i(w, \ulcorner\Gamma \to \Delta\urcorner)$ and $Prf^i(w, \ulcorner\forall x\neg\varphi(x)\urcorner)$ respectively."

In §2 Lemma 3 in [5], we proved the following theorem.

Theorem. *Let $\varphi(c)$ be an i-normal formula with only free variable c. Then*

$$S_{2,n}^{i+1} \vdash Prf^i(|w|_n, \ulcorner\varphi(c)\urcorner) \to \varphi(c).$$

The key point of the proof of this theorem is that if $\ulcorner t(\vec{a})\urcorner$ is n-small, i.e., of the form $|u|_n$, then $|b|^{|\ulcorner t(\vec{a})\urcorner|}$ is small in $S_{2,n}^1$. Therefore the bound $\exp(\ulcorner t(\vec{a})\urcorner |b|^{|\ulcorner t(\vec{a})\urcorner|})$ can be expressed in $S_{2,n}^1$ and $v(\ulcorner t(\vec{a})\urcorner, \vec{b})$ can be expressed in $S_{2,n}^1$.

We add several remarks on the theorem. In $Prf^i(|w|_n, \ulcorner\varphi(c)\urcorner)$, c in $\ulcorner\varphi(c)\urcorner$ is a variable, therefore we might write the theorem

$$S_{2,n}^{i+1} \vdash Prf^i(|w|_n, \ulcorner\varphi(a)\urcorner) \to \varphi(c).$$

We define I_k by the following as usual

$$\begin{aligned} I_0 &= 0 \\ I_{2k+1} &= I_{2k} + 1 \\ I_{2(k+1)} &= 2 \cdot I_{k+1} \end{aligned}$$

where 2 is $1 + 1$. Then $v(\ulcorner t(a_1, \dots, a_n)\urcorner, b_1, \dots, b_n) = v(\ulcorner t(I_{b_1}, \dots, I_{b_n})\urcorner)$. Therefore we have the following theorem with the same proof.

Theorem. *Let $\varphi(c)$ be an i-normal formula with only free variable c. Then*

$$S_{2,n}^{i+1} \vdash Prf^i(|w|_n, \ulcorner\varphi(I_c)\urcorner) \to \varphi(c).$$

Now we come back to our present case in S_2^{i+1}. Then $v(\ulcorner t(\vec{a})\urcorner, \vec{b})$ is definable if $\ulcorner t(\vec{a})\urcorner$ is 2 small and b is small and $v(\ulcorner t\urcorner)$ is definable if $\ulcorner t\urcorner$ is the Gödel number of a closed term and $\ulcorner t\urcorner$ is small.

These two conditions give the following two theorems which are proved in the same way as in the proof of §2 Lemma 3 in [5].

Theorem 1. *Let $\Gamma(c) \to \Delta(c)$ be a sequent with only free variable c and all formulas in $\Gamma(c)$ and $\Delta(c)$ be i-normal. Then*

$$S_2^{i+1} \vdash c \leq |d|, Prf^i(|w|_2, \ulcorner\Gamma(a) \to \Delta(a)\urcorner), \Gamma(c) \to \Delta(c).$$

Theorem 2. *Let $\Gamma \to \Delta$ be a sequent and all formulas in Γ and Δ be i-normal sentences. Then*

$$S_2^{i+1} \vdash Prf^i(|w|, \ulcorner\Gamma \to \Delta\urcorner), \Gamma \to \Delta.$$

We have the following corollaries.

Corollary 3. *Let $\varphi(a)$ be an i-normal formula in which a is only free variable. Then*

$$S_2^{i+1} \vdash \exists w Prf^i(|w|_2, \ulcorner\forall x \neg\varphi(x)\urcorner) \to \forall x \neg\varphi(|x|).$$

Corollary 4.

$$S_2^{i+1} \vdash \neg Prf^i(|w|, \ulcorner \to \urcorner).$$

Remark. The following theorem is obtained by the same method with the proof of Corollary 3.

Theorem. *Let $\varphi(a)$ be an i-normal formula in which a is only free variable. Then*

$$S_2^1 \vdash \exists w Prf^i(|w|_3, \ulcorner\forall x \neg\varphi(x)\urcorner) \to \forall x \neg\varphi(|x|_2).$$

3 Gödel sentences

See §7.5 in [2] for the general theory of Gödel sentences in Bounded Arithmetic. We define Gödel sentences φ_k^i satisfying

$$S_2^1 \vdash \varphi_k^i \longleftrightarrow \forall x \neg Prf^i(|x|_k, \ulcorner \varphi_k^i \urcorner).$$

From the definition of φ_k^i, φ_k^i is of the form $\forall x \neg G_k^i(|x|_k)$ where G_k^i is an i-form formula and we have

$$S_2^1 \vdash G_k^i(a) \longleftrightarrow Prf^i(a, \ulcorner \varphi_k^i \urcorner).$$

The following properties on Gödel sentences are proved by the standard argument.

Theorem 5. *S_2^i does not prove $\varphi_k^i (k = 0, 1, 2, \ldots)$ and $S_2^1 \vdash \varphi_0^i \longleftrightarrow \forall x \neg\ Prf^i(x, \ulcorner \rightarrow \urcorner)$ therefore $S_2^i \nvdash \forall x \neg Prf^i(x, \ulcorner \rightarrow \urcorner)$.*

The following natural question arises here. Is $\forall x \neg Prf^i(x, \ulcorner \rightarrow \urcorner)$ provable in S_2^{i+1}? If the answer is yes, then we have $S_2^i \neq S_2^{i+1}$ and $P \neq NP$. But we believe the answer is no in the following reason.

Let $S_2^{-\infty}$ be the equational theory involving equations $s = t$, where s and t are closed terms in the Buss' original language of S_2^1 with the natural rules of the function symbols. Therefore $S_2^{-\infty}$ does not have any free variables, any logical symbols or any inductions.

As is stated in [3], we conjecture

$$S_2 \nvdash Con(S_2^{-\infty})$$

where Con$(S_2^{-\infty})$ is the consistency of $S_2^{-\infty}$. $S_2^{-\infty}$ is an extremely weak system and therefore Con$(S_2^{-\infty})$ is much weaker than $\forall x \neg Prf^0(x, \ulcorner \rightarrow \urcorner)$. Therefore our conjecture implies

$$S_2^i \nvdash \forall x \neg Prf^0(x, \ulcorner \rightarrow \urcorner)$$

and we believe that there is no hope to prove

$$S_2 \vdash \forall x \neg Prf^i(x, \ulcorner \rightarrow \urcorner).$$

On the other hand, we have the following comjecture.

Conjecture 6. $S_2^i \nvdash \forall x \neg Prf^i(|x|, \ulcorner \rightarrow \urcorner)$.

This conjecture together with Corollary 4 certainly implies $S_2^i \neq S_2^{i+1}$ and
$P \neq NP$. However it should be noted that $S_2^1 \vdash \forall x \neg Prf^i(|x|_2, \ulcorner \rightarrow \urcorner)$.

Remark. For the feasibility of our conjecture we would like to discuss its relation with Baker-Gill-Solovay's result in [1].

First our system S_2^i must satisfy the following basic conditions as we stated before.
(a) The predicate constants are only $\leq$ and $=$.
(b) The number of function constants are finite and all the function constants express polynomial time computable functions.
(c) Extra true axioms are allowed to be used only when they can be expressed by finitely many initial sequents without logical symbols.
These basic conditions cannot accept Baker-Gill-Soloray type relativizations. Moreover we discuss the comparison of the typical Baker-Gill-Soloray case and our case.

Let $\widetilde{S_2^i}$ be S_2^i+ (Axiom on PSPACE–complete predicate) and $\widetilde{P}$ and $\widetilde{NP}$ be the class of P and NP in the language of $\widetilde{S_2^i}$. Then we have $\widetilde{P} = \widetilde{NP}$. We consider a similar stronger example $\widetilde{S_2^i} = S_2^i+$ (Axiom on exponential function) and let $\overset{\approx}{P}$ and $\overset{\approx}{NP}$ be P and NP formulated in the language of $\overset{\approx}{S_2^i}$. Then $\overset{\approx}{P}=\overset{\approx}{NP}$. In this case of $\overset{\approx}{S_2^i}$, the situation is totally opposite to our case of conjecture i.e.,

$$\overset{\approx}{S_2^{i+1}} \nvdash \forall x \neg \overset{\approx}{Prf^i} (|x|, \ulcorner \rightarrow \urcorner)$$

in the place of $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|, \ulcorner \rightarrow \urcorner)$ and $\overset{\approx}{S_2^i} \nvdash \forall x \neg \overset{\approx}{Prf^i} (|x|, \ulcorner \rightarrow \urcorner)$ is trivial in the place of our conjecture $S_2^i \nvdash \forall x \neg Prf^i(|x|, \ulcorner \rightarrow \urcorner)$. Therefore Baker-Gill-Solovay's result has no relation with the feasibility of our conjecture.

Theorem 7. *We have for* $k > 2$

$$S_2^{i+1} \vdash \forall x \neg G_k^i(|x|_{k+1})$$

especially $S_2^{i+1} \vdash \forall x \neg G^i(|x|_3)$.

Proof. By the definition of φ_k^i, we have

$$S_2^{i+1} \vdash \neg \varphi_k^i \rightarrow \exists x Prf^i(|x|_k, \ulcorner \forall x \neg G_k^i(|x|_k) \urcorner).$$

From this and Corollary 3 follows

$$S_2^{i+1} \vdash \neg \varphi_k^i \rightarrow \forall x \neg G_k^i(|x|_{k+1}).$$

Therefore $S_2^{i+1} \vdash \forall x \neg G_k^i(|x|_k) \vee \forall x \neg G_k^i(|x|_{k+1})$. Since $G_k^i(a)$ is equivalent to a form $Prf^i(a, \ulcorner \varphi_k^i \urcorner)$, we have $S_2^{i+1} \vdash \forall x \neg G_k^i(|x|_{k+1})$. $\square$

Corollary 8. *For* $k \geq 2$ *we have*

$$S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_{k+1}, \ulcorner \varphi_k^i \urcorner)$$

especially $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_3, \ulcorner \varphi_2^i \urcorner)$.

Conjecture 9.

$$S_2^i \nvdash \forall x \neg Prf^i(|x|_3, \ulcorner \varphi_2^i \urcorner).$$

Obviously this conjecture implies $S_2^i \neq S_2^{i+1}$ and $P \neq NP$. It should be noted that for $k \geq 3$, $S_2^1 \vdash \forall x \neg G_k^i(|x|_{k+2})$ i.e., $S_2^1 \vdash \forall x \neg Prf^i(|x|_{k+2}, \ulcorner \varphi_k^i \urcorner)$.

Theorem 10. *We have*

$$S_2^{i+1} \vdash \varphi_2^i \vee \forall x \neg Prf^i(|x|_2, \ulcorner \forall x \neg G_2^i(|x|) \urcorner).$$

Proof. We discuss this inside of S_2^{i+1}. Suppose

$$\neg \forall x \neg Prf^i(|x|_2, \ulcorner \forall x \neg G_2^i(|x|) \urcorner)$$

i.e, $\exists x Prf^i(|x|_2, \ulcorner \forall x \neg G_2^i(|x|) \urcorner)$. Then from Corollary 3 follows

$$\forall x \neg G_2^i(|x|_2) \text{ i.e., } \varphi_2^i.$$

□

Conjecture 11. *The following statement holds.*

$$S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_2, \ulcorner \forall x \neg G_2^i(|x|) \urcorner) \rightarrow \varphi_2^i$$

It is easily seen that this conjecture also implies $S_2^i \neq S_2^{i+1}$ and $P \neq NP$.

Theorem 12. *If* $S_2^{i+1} \vdash \varphi_k^i \rightarrow \forall x \neg A(x)$ *and* $k \geq 3$ *and* $A(x)$ *is* i*-normal, then*

$$S_2^{i+1} \vdash \forall x \neg A(|x|).$$

Proof. $S_2^{i+1} \vdash \varphi_k^i \rightarrow \forall x \neg A(x)$ implies that there exists a constant size proof of the following form

$$\begin{array}{c}
\vdots \\
\cdots \quad \rightarrow \cdots \mathrm{Prf}^i(|t|_k, \ulcorner \psi_k^i \urcorner) \\
\hline
\neg \mathrm{Prf}^i(|t|_k, \ulcorner \psi_k^i \urcorner), \cdots \rightarrow \cdots \\
\hline\hline
\forall x \neg \mathrm{Prf}^i(|x|_k, \ulcorner \psi_k^i \urcorner), \cdots \rightarrow \cdots \\
\vdots \\
\hline
A(a), \forall x \neg \mathrm{Prf}^i(|x|_k, \ulcorner \psi_k^i \urcorner), \cdots \rightarrow \cdots
\end{array}$$

Let us denote the size of this proof by a constant c_0. Now we discuss inside of S_2^{i+1} and suppose $\neg\varphi_k^i$ i.e., $\exists x Prf^i(|x|_k, \ulcorner\forall x \neg Prf^i(|x|_k, \ulcorner\psi_k^i\urcorner)\urcorner)$. Then there exists a proof $|x_0|_k$ of

$$Prf^i(|b|_k, \ulcorner\psi_k^i\urcorner) \rightarrow \quad .$$

Using the previous proof of size c_0 and this proof $|x_0|_k$ and making many cut of the following form, we get a proof of

$$A(a) \rightarrow$$

i.e.,

$$\dfrac{\begin{array}{c}\vdots \\ \cdots \quad \rightarrow \cdots, \mathrm{Pr}f^i(|t|_k, \ulcorner\psi_k^i\urcorner)\mathrm{Pr}f^i(|t|_k, \ulcorner\psi_k^i\urcorner) \rightarrow\end{array}}{\begin{array}{c}\cdots \quad \rightarrow \cdots \\ \vdots\end{array}}$$

$$\dfrac{}{A(a) \rightarrow}$$

Since the whole procedure is polynomial time computable from two proofs, the size of the new proof is not greater than $t(|x_0|_k)$ for some term t.

Since $2 < k$, we have

$$\exists z Prf^i(|z|_2, \ulcorner\forall x \neg A(x)\urcorner)$$

and we proved

$$S_2^{i+1} \vdash \neg\varphi_k^i \rightarrow \exists z Prf^i(|z|_2, \ulcorner\forall x \neg A(x)\urcorner).$$

Then by Corollary 3 we have

$$S_2^{i+1} \vdash \neg\varphi_k^i \rightarrow \forall x \neg A(|x|).$$

Therefore we have

$$S_2^{i+1} \vdash \forall x \neg A(x) \vee \forall x \neg A(|x|)$$

i.e.,

$$S_2^{i+1} \vdash \forall x \neg A(|x|).$$

□

Corollary 13. *If $S_2^{i+1} \vdash \varphi_k^i \to \forall x > Prf^i(|x|_{k-1}, \ulcorner\varphi_k^i\urcorner)$ and $k > 2$, then $S_2^{i+1} \vdash \varphi_k^i$ and therefore $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_{k-1}, \ulcorner\varphi_k^i\urcorner)$ and $S_2^{i+1} \neq S_2^i$ and $P \neq NP$.*

Proof. By Theorem 12, we have $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_k, \ulcorner\varphi_k^i\urcorner)$ i.e., $S_2^{i+1} \vdash \varphi_k^i$ therefore $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_{k-1}, \ulcorner\varphi_k^i\urcorner)$ and $S_2^{i+1} \neq S_2^i$ and $P \neq NP$. □

Theorem 14. *For $k \geq 1$ we have*

$$S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_{k+1}, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner).$$

Proof. We have

$$S_2^{i+1} \vdash \exists x Prf^i(|x|_2, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner) \to \forall x \neg G_k^i(|x|_k).$$

On the other hand, we have

$$\begin{aligned} S_2^1 \quad \vdash & \exists x Prf^i(|x|_{k+1}, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner) \\ & \to \exists x Prf^i(|x|_k, \ulcorner\forall x \neg G_k^i(|x|_k)\urcorner) \end{aligned}$$

since a proof of $\forall x \neg G_k^i(|x|_k)$ is obtained from a proof of $\forall x \neg G_k^i(|x|_{k-1})$ by a polynomial time computable operation. Therefore we have

$$\begin{aligned} S_2^{i+1} \vdash \quad & \neg(\exists x Prf^i(|x|_2, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner) \\ & \wedge \exists x Prf^i(|x|_{k+1}, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner) \end{aligned}$$

since $\exists x Prf^i(|x|_k, \ulcorner\forall x \neg G_k^i(|x|_k)\urcorner)$ is equivalent to $\neg\varphi_k^i$ and $\forall x \neg G_k^i(|x|_k)$ is φ_k^i. Therefore we have

$$S_2^{i+1} \vdash \neg\exists x Prf^i(|x|_{k+1}, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner).$$

□

Conclusion. At this moment, there are many mysteries regarding the nature of Gödel sentences in Bounded Arithmetic. We strongly believe that $S_2^i \neq S_2^{i+1}$ and $P \neq NP$ would be proved if our knowledge on these Gödel sentences could be improved. We list several problems on Gödel sentences.

1. Is $S_2^{i+1} \vdash \varphi_k^i$ true? We conjecture that this is false for $k = 0, 1$.
2. Is $S_2^{i+1} \vdash \forall x \neg G_0^i(|x|)$ true? Is $S_2^{i+1} \vdash \forall x \neg G_1^i(|x|_2)$ true? At this moment we have no idea about them.
3. Is $S_2^{i+1} \vdash \forall x \neg Prf^i(|x|_k, \ulcorner\forall x \neg G_k^i(|x|_{k-1})\urcorner)$ true for some k?

Especially is
$S_2^{i+1} \vdash \forall x \neg Prf^i(|x|, \ulcorner\forall x \neg G_1^i(x)\urcorner)$ true?

4 References

[1] T. Baker, J. Gill & R. Solovay, Relativization of the $P =?NP$ question, SIAM Journal of Computing 4 (1975), 431–442.

[2] S. Buss, Bounded Arithmetic, Bibliopolis, Napoli, (1986)

[3] P. Clote & J. Krajíček, Problem 1.2.9 in p 5 in Open problems, Arithmetic, proof theory and computational complexity, Oxford, (1993)

[4] J. Krajíček, P. Pudlák & G. Takeuti, Bounded Arithmetic and the Polynomial Hierachy, Annals of Pure and Applied Logic, vol. 52 (1991), 143–153

[5] G. Takeuti, Bounded Arithmetic and Truth Definition, Annals of Pure and Applied Logic, vol. 39 (1988), 75–104.

[6] A. Wilkie & J. Paris, On the scheme of induction for bounded arithmetic formulas, Proc. of an ASL Conference in Manchester, (1986) North Holland.

Author address

Department of Mathematics
University of Illinois
1409 W Green Street
Urbana, Il 61801 USA
email: `takeuti@symcom.math.uiuc.edu`