INCONSISTENCY PROOFS IN NONSTANDARD MODELS

K. EDA AND T. UESU

Abstract. Presuming the consistency of *ZF*, according to the Gödel completeness theorem and second incomplete theorem we have a model of *ZF* in which there exists a proof of the inconsistency of *ZF*. We prove that formulas of hyperfinite length occur in this proof and the length of the proof is also hyperfinite.

Let *ZF* be the Zermelo-Frankel set theory. We presume that *ZF* is consistent. Then, the Gödel second incompleteness theorem [2] implies $Consis(ZF)$ is not provable in *ZF*, where $Consis(ZF)$ is a closed formula coding the consistency of *ZF* in a standard way. Now the Gödel completeness theorem [1] implies that there exists a model M of *ZF* in which the negation of $Consis(ZF)$ holds, i.e. there exists a proof in M from *ZF* to the contradiction $\neg\forall x(x = x)$. Of course this proof should not be a standard proof and the theory *ZF* in M is not a standard one. To express these more exactly let $\mathbb{N}^M$ be the set of the natural numbers in a model M of *ZF*. Then the initial part of $\mathbb{N}^M$ corresponds to the set of the standard natural numbers $\mathbb{N}$ and we identify this part with $\mathbb{N}$. Then, the set $\mathbb{N}^M \setminus \mathbb{N}$ is non-empty. We call a natural number in $\mathbb{N}^M \setminus \mathbb{N}$ *hyperfinite*.

In this paper we show that in the above proof a formula of hyperfinite length actually occurs and the height of the proof is also hyperfinite. To state our theorem precisely we define some notions and fix our proof system. As we remark in Remark 2.1, our results are depending on axiom systems. We use a Frege-Hilbert style proof system, which we explain more in the next section. Since we only deal with systems without function symbols or individual constant symbols, the usual Peano Arithmetic is not in our scope. When we

---

1991 *Mathematics Subject Classification.* Primary 03F40, 03F20 Secondary 03H05.

Key words and phrases. Incompleteness theorem, Completeness theorem, Inconsistency, Nonstandard models, hyperfinite, length, height.

We thank to Yukako Kezuka for giving us the opportunity to investigate this subject. The first author was supported in part by the Grant-in-Aid for Scientific research (C) of Japan No. 20540110.

mention the Peano Arithmetic (abbreviated by PA), it is the one obtained by adding new predicates for function symbols with additional axioms in a standard way.

As we shall show, the fact that a formula of hyperfinite length actually occurs in the inconsistency proof is proved straightforwardly. But, to show the fact that the height of the proof is hyperfinite, we need some device.

§1. Formal system. For the predicate calculus we adopt a system with the only logical symbols $\supset$, $\neg$ and $\forall$. We use $\wedge$, $\vee$ and $\exists$ as abbreviations as usual, i.e. $\exists x$ is the abbreviation of $\neg\forall x\neg$ and so on. The variables are given as $\boldsymbol{v}_n$ for $n \in \mathbb{N}$ and we usually use x_n and y_n auxiliarily. Formulas are defined as usual. In particular, if P is an n-ary predicate or an n-ary predicate variable, the latter of which we shall introduce later, the prime formula is expressed as $Px_1 \cdots x_n$ for variables $x_1, \cdots, x_n$ to display the real sequence of letters (for $=$ and $\in$ we express as $x_1 = x_2$ and $x_1 \in x_2$ as usual). The *length* of a formula is the number of the occurrences of symbols, where we adopt the Poland expression for formulas though we never explicitly use it. We use the symbol $\equiv$, when the two sequences of letters are the same. We explicitly state our system below.

Logical Axioms

$A \supset (B \supset A)$ $\qquad$ $(A \supset (B \supset C) \supset (A \supset B)) \supset (A \supset C)$

$\neg\neg A \supset A$ $\qquad$ $(A \supset B) \supset ((A \supset \neg B) \supset \neg A)$

$\forall x A(x) \supset A(y)$ $\qquad$ $\forall x(C \supset A(x)) \supset (C \supset \forall x A(x))$

x does not occur in C freely.

$\forall x(x = x)$ $\qquad$ $\forall x \forall y(x = y \ \supset (A(x) \supset A(y)))$

Inference Rules

(I1) $\dfrac{A \quad A \supset B}{B}$ $\qquad$ (I2) $\dfrac{A(x)}{\forall x A(x)}$

As logical axioms, some axiom in *ZF* or in *PA* is given as an axiom schema. Except when the emphasis of a schema is necessary, we simply call it an axiom. Instead of deductions we define the notion of proofs to make a distinction between axioms and conclusions. To describe proofs and formulas, we use finite plane trees with labeled formulas. A finite *tree* T is a finite partially ordered set with the minimal element r, called the root, A *branch* in a tree is a maximal linearly ordered subset of the tree. such that $\{s \in T \,|\, s \preceq t\}$ is linearly ordered for each $t \in T$. For a formula F, let T_F be a labeled plane tree whose labels are formulas appearing in the formation of F. For instance, when $F \equiv (A \supset B) \supset C$, the root of T_F is F and the immediate successors of the root are $A \supset B$ and C etc. On the plane the position of $A \supset B$ is left to C, and $A \supset B$ and C are upper to F. For each $t \in T_F$ we denote the formula labeled to t by A_t. A node is not just a formula, but is with its position. But sometimes we regard a node as a formula for simplicity. The *height* of a formula F is the maximal length of branches in T_F.

A *proof* $\mathcal{P}$ is also a labeled plane tree $T_{\mathcal{P}}$ where the formula labeled to a node t is denoted by A_t and each node of $T_{\mathcal{P}}$ has at most two successors, and which satisfies the following:

(0) if there exists no successor of t, i.e. t is a leaf, then A_t is an axiom;
(1) if s is the unique successor of t, then A_t is a conclusion of A_s by inference rule (I2); and
(2) if s_0 and s_1 are distinct successors of t and s_0 is left to s_1, then A_t is conclusion of A_{s_0} and A_{s_1} by inference rule (I1) where A_{s_0} is left to A_{s_1}.

For the root r of $T_{\mathcal{P}}$ we call A_r the conclusion of $\mathcal{P}$. The height of a proof $\mathcal{P}$ is the maximal length of of branches in $T_{\mathcal{P}}$. A *branch* in a proof $\mathcal{P}$ is the sequence of formulas $(A_t : t \in \boldsymbol{b})$ where the ordering is induced from a branch $\boldsymbol{b}$ in $T_{\mathcal{P}}$.

Definition 1.1. Let $x_1, \cdots, x_n$ be the enumeration of variables appearing in a formula A without repitition. A formula B is *similar* to A, if B is the result of replacing all the occurrences of variables $x_1, \cdots, x_n$ in A by variables $y_1, \cdots, y_n$ respectively where $y_1, \cdots, y_n$ are mutually distinct.

We note that B is similar to A if and only if A is similar to B.

§2. Main theorem, definitions and preliminary lemmas. In this section we prove the first half of the following theorem, i.e. Lemma 2.3.

Theorem 2.1. *Let K be ZF or PA and M be a model of K in which $\neg Consis(K)$ holds.*

Then, a proof of the contradiction $\neg \forall x(x = x)$ from K in M contains a formula of hyperfinite length and the height of the proof is also hyperfinite.

Remark 2.1. Here we remark that our results and proofs depend on the systems. For example, if we adopt the axioms of *ZF* closed under inference rules, then every provable statement is proved in one line and this also holds in M. We adopt the standard axioms of *ZF*, for instance, [3] or [4]. If we adopt a proof system LK which admits only prime formulas in the beginning sequences and as the addition-rules, the appearance of a formula of hyperfinite length in a proof automatically implies that the height of the proof is also hyperfinite.

In the following a theory K stands for *ZF* or *PA*.

Definition 2.2. A proof $\mathcal{P}$ is *without detour*, if no branch of $T_{\mathcal{P}}$ contains distinct nodes s and t such that A_s and A_t are similar.

Lemma 2.2. *If C is a conclusion of a proof $\mathcal{P}$ in K, then there exists a proof $\mathcal{Q}$ of C in K without detour such that the height of $\mathcal{Q}$ is equal to or less than that of $\mathcal{P}$ and the maximal length of formulas occurring in $\mathcal{Q}$ is equal to or less than that of formulas occurring in $\mathcal{P}$.*

Proof. Suppose that there exist distinct $s, t \in T_{\mathcal{P}}$ such that $s \preceq t$, A_s and A_t are similar, and A_t is the result of replacing variables $x_1, \cdots x_n$ by $y_1, \cdots, y_n$ from A_s. For $w \in T_{\mathcal{P}}$, let $\mathcal{P}_w$ be the proof of A_w such that $T_{\mathcal{P}_w} = \{u \in T_{\mathcal{P}} \mid w \preceq u\}$.

Since a similar formula to an axiom of K is also an axiom of K, replacing all the occurrences of x_i in $\mathcal{P}_t$ by y_i and all the occurrences of y_i in $\mathcal{P}_t$ by x_i we have a proof of A_s whose height is the same as that of $T_{\mathcal{P}_t}$. Replace the proof of A_s in $\mathcal{P}$ by this proof. Iterating this replacement we have a proof of C with the requiring properties. $\dashv$

Lemma 2.3. *Every proof of the contradiction $\neg\forall x(x = x)$ from K in M contains a formula of hyperfinite length.*

Proof. To the contradiction, suppose that the length of every formula occurring in a proof $\mathcal{P}$ of $\neg\forall x(x = x)$ is standard. Using Lemma 2.2 in M, we have a proof $\mathcal{Q}$ in K satisfying the properties there. Let m_0 be the the maximal length of formulas appearing in $\mathcal{Q}$. Then, m_0 should be standard by the assumption. Since the number of the symbols in K other than variables are finite in the standard sense and the maximal length of formulas appearing in $\mathcal{Q}$ is bounded by m_0, the number of similarities of formulas appearing in $\mathcal{Q}$ is bounded by some standard natural number m_1. Since $\mathcal{Q}$ is a proof without detour, the lengths of branches in $\mathcal{Q}$ are bounded by m_1 and so is the height of $\mathcal{Q}$.

Now, the total number of formulas appearing in $\mathcal{Q}$ is bounded by 2^{m_1} and consequently the total number of variables occurring in $\mathcal{Q}$ is bounded by $2^{m_1}m_0$. Therefore, replacing variables indexed by hyperfinite natural numbers by variables indexed by standard natural numbers, we have a standard proof of $\neg\forall x(x = x)$, which is a contradiction. ⊣

§3. Core proofs. To prove the remaining statement of Theorem 2.1 we introduce a notion "core proofs".

Let L^+ be the extended language with additional *predicate variables*. According to this extension, we extend axioms in K to K^+ naturally. For a finite set Γ of predicate variables $L \cup \Gamma$ is a sublanguage of L^+ restricted predicates to Γ.

DEFINITION 3.1. Let $*$ be a distinguished symbol which is a variable not in L^+. A *holed formula* is a formula in $L^+ \cup \{*\}$ without any free occurrences of variables other than $*$ and without any bound occurrences of $*$. For each formula F in L^+, let $\overline{F}$ be the holed formula obtained by replacing all the free occurrences of variables by $*$. The arity of a holed formula H is the number of occurrences of $*$. If n is the arity of H, $H(x_1, \cdots, x_n)$ is the formula obtained by replacing the i-th occurrence of $*$ from the left in H by x_i.

DEFINITION 3.2. An *assignment* σ has a finite support Γ, which is a finite set of predicate variables, and assigns a holed formula P^σ to each predicate variable P, where the arities of P and P^σ are the same, and $P^\sigma \equiv P * \cdots *$ for $P \notin \Gamma$, where the number of $*$ is the arity of P. For a predicate letter P in L, define $P^\sigma \equiv P * \cdots *$, for instance, $\in^\sigma \equiv * \in *$.

For every formula in L^+ and holed formula F we inductively define F^σ as follows: For a predicate variable or predicate letter P in L, define $(Px_1 \cdots x_m)^\sigma$ as the formula $P^\sigma(x_1, \cdots, x_m)$, where x_i maybe $*$. Then, define inductively as: $(\neg A)^\sigma \equiv \neg A^\sigma$, $(A \supset B)^\sigma \equiv A^\sigma \supset B^\sigma$ and $(\forall x A)^\sigma \equiv \forall x A^\sigma$.

An assignment σ is *admissible* for $Px_1 \cdots x_m$, if the substitutions of x_i to $*$ do not violate the scopes of quantifiers in P^σ, i.e. the i-th occurence of $*$ is not in a scope of a quantificatio $\forall x_i$. An assignment σ is *admissible* for a formula or a proof, if σ is admissible for every $Px_1 \cdots x_m$ occurring in the formula or the proof respectively.

We remark that sometimes an assignment assigns a predicate variable to a formula in L^+ but not in L.

LEMMA 3.1. *Let σ be an assignment such that σ is admissible for a pime formula $Px_1 \cdots x_m$. If $(Px_1 \cdots x_m)^\sigma \equiv F^\sigma$ for a formula F, then $P^\sigma \equiv \overline{F^\sigma}$ holds.*

In particular if $F \equiv Qy_1 \cdots y_n$ for a predicate variable $Q$and σ is also admissible for $Qy_1 \cdots y_n$, then $P^\sigma \equiv Q^\sigma$, $m = n$ and $x_i \equiv y_i$ for $1 \leq i \leq m = n$.

PROOF. Since σ is admissible, the i-th $*$ in P^σ is not in the scope of bound variable x_i in P^σ and hence the number of free occurrences is m which is the same for F^σ. Therefore we have $P^\sigma \equiv \overline{F^\sigma}$. Now the second statement is clear. ⊣

Since an assignment never increase free occurrences of variables, the assignments follow the regulations for the subset axiom, the replacement axiom in *ZF* and the induction axiom in *PA*. Therefore it is rather straightforward to prove the following lemma and we omit its proof.

LEMMA 3.2. *Let $\mathcal{P}$ be a proof in K^+ and σ is an assignment. Then $\mathcal{P}^\sigma$ is also a proof.*

DEFINITION 3.3. If $\mathcal{P}$ is a proof in K^+ and σ is an assignment which assigns a formula in L to each predicate variable, $\mathcal{P}$ is called a *core proof* of $\mathcal{P}^\sigma$ by σ.

LEMMA 3.3. *Let F be a formula and $\overline{F}$ its holed formula, and σ be an assignment. Then, $\overline{F}^\sigma$ is the holed formula of F^σ.*

PROOF. Let P be a predicate letter in L or a predicate variable. When $F \equiv Py_1 \cdots y_m$, $F^\sigma \equiv P^\sigma(y_1, \cdots, y_m)$ and $\overline{F}^\sigma \equiv P^\sigma(y_1, \cdots, y_m)$ by definition and hence the conclusion holds. Since the induction steps for $\neg$ and $\supset$ are obvious, we only deal with $F \equiv \forall x G$.

By induction hypothesis $\overline{G^\sigma} \equiv G^\sigma$. Let $\overline{F} \equiv \forall x G_0$. Then, $\overline{G}$ is obtained from G_0 by replacing free occurrences of x by $*$ and hence $\forall x G_0^\sigma \equiv \overline{\forall x G^\sigma}$. Since $\overline{F}^\sigma \equiv \forall x G_0^\sigma$, we have the conclusion. ⊣

We have defined the height of a formula in Section 1. An axiom is not always given by a formula and so we need to explain how to define the height of an axiom. We consider explicit forms of axioms and define the heights. For instance, the heights of the logical axioms $A \supset (B \supset A)$ and $\forall x(x = x)$ are 3 and 2, respectively, and the height of the subset axiom in set theory

$$\forall x \exists y \forall z((z \in y \supset z \in x \land A(z)) \land (z \in x \land A(z) \supset z \in y))$$

is 11, considering the abbreviations $\exists$ and $\wedge$.

Let m_0 be the maximal height of axioms, for instance the maximal height of logical axioms is 6. Hence, m_0 becomes that of an axiom of K.

Define function g by:

$$\begin{aligned} g(1, m_0) &= m_0 \\ g(n+1, m_0) &= g(n, m_0)(2^{g(n,m_0)+1} + 1). \end{aligned}$$

THEOREM 3.4. *Let m_0 be the maximal height of axioms in K. Let $\mathcal{P}$ be a proof of its height n. Then, there exist a proof $\mathcal{K}$ in K^+ and an assignment σ such that $\mathcal{K}$ is a core proof of $\mathcal{P}$ by σ, σ is admissible for $\mathcal{K}$ and the heights of formulas occurring in $\mathcal{K}$ are bounded by $g(n, m_0)$.*

§4. Proofs of Theorems 2.1 and 3.4. Before proceeding to the proof of Theorem 3.4 we prove some lemmas. The first lemma is straight forward and we omit its proof.

LEMMA 4.1. *Let φ be an assignment with support Γ and A a formula. Suppose that the height of a formula A is equal to or less than h and the height of P^φ for $P \in \Gamma$ is bounded by h_0. Then the height of A^φ is less than or equal to $h + h_0$.*

LEMMA 4.2. *Let P, Q be predicate variables, A be a formula such that $A \equiv \overline{A}(x_1, \cdots, x_m)$, $\mathcal{K}_0$ be a proof of A, $\mathcal{K}_1$ be a proof of $Px_1 \cdots x_m x_{m+1} \cdots x_n$ for $m \leq n$ and φ be an assignment such that $P^\varphi \equiv \overline{A} \supset Q * \cdots *$, where the arity of Q is $n - m$. If the heights of formulas occurring in $\mathcal{K}_0$ or $\mathcal{K}_1$ are bounded by h, then the heights of formulas in proof* $\dfrac{\mathcal{K}_0 \quad \mathcal{K}_1^\varphi}{Qx_{m+1} \cdots x_n}$ *are bounded by $2h + 1$.*

PROOF. We have

$$\begin{aligned} (Px_1 \cdots x_m x_{m+1} \cdots x_n)^\sigma &\equiv \overline{A}(x_1, \cdots, x_m) \supset Qx_{m+1} \cdots x_n \\ &\equiv A \supset Qx_{m+1} \cdots x_n \end{aligned}$$

and hence $\dfrac{\mathcal{K}_0 \quad \mathcal{K}_1^\varphi}{Qx_{m+1} \cdots x_n}$ is a proof. Since the height of $\overline{A} \supset Q * \cdots *$ is less than or equal to $h + 1$, the conclusion follows from Lemma 4.1. ⊣

Next we prove the main lemma for Teorem 3.4.

LEMMA 4.3. *Let k be the number of predicate variables occurring in F or G, and h_0 the maximal height of F and G. Suppose that*

$F^\sigma \equiv G^\sigma$ for an assignment σ. Then there exists an assignment φ such that

(1) *$(P^\varphi)^\sigma \equiv P^\sigma$ for each predicate variable P; and*
(2) *$F^\varphi \equiv G^\varphi$ and the height of F^φ (equal to that of G^φ) is less than or equal to $(k+1)h_0$;*

Proof. We inductively define assignments $\psi_1, \cdots, \psi_i$ for some $i \leq k$ so that $F^{\psi_1\cdots\psi_i} \equiv G^{\psi_1\cdots\psi_i}$ and the following are satisfied. Let Γ be the set of predicate variables of occurring in F or G and $\{P_j\}$ be the support of ψ_j for $1 \leq j \leq i$.

(1_i) $(Q^{\psi_1\cdots\psi_i})^\sigma \equiv Q^\sigma$ for every predicate variable $Q \in \Gamma$;
(2_i) the height of $(P_j)^{\psi_j}$ is less than or equal to h_0; and
(3_i) the heights of $F^{\psi_1\cdots\psi_i}$ and $G^{\psi_1\cdots\psi_i}$ are bounded by $h + ih_0$.

We consider the trees $T_{F^{\psi_1\cdots\psi_{j-1}}}$ and $T_{G^{\psi_1\cdots\psi_{j-1}}}$ and call a P-leaf for a leaf labeled by $Px_1\cdots x_m$ for some $x_1, \cdots, x_m$. For a node $t \in T_{F^{\psi_1\cdots\psi_{j-1}}}$ or $t \in T_{G^{\psi_1\cdots\psi_{j-1}}}$, $t^* \in T_{G^{\psi_1\cdots\psi_j}}$ or $t^* \in T_{F^{\psi_1\cdots\psi_j}}$ is defined respectively, if t^σ and $(t^*)^\sigma$ are the same node in $T_{F^\sigma}(= T_{G^\sigma})$, and t^* is undefined, otherwise.

Unless $F^{\psi_1\cdots\psi_{j-1}} \equiv G^{\psi_1\cdots\psi_{j-1}}$, there exists a P-leaf t such that t^* is undefined or t^* is not a P-leaf. We choose P_j so that the height of P_j^σ is maximal among such P. Then we have a P_j-leaf t such that t^* exists but is not a P_j-leaf, since the undefinedness of t^* implies the existence of a Q-leaf s such that the height Q^σ is strictly greater than that of P_j^σ, s^* exists and s^* is not a Q-leaf.

We have $T_F \subseteq \cdots \subseteq T_{F^{\psi_{j-1}}} \subseteq T_{F^\sigma}$ and $T_G \subseteq \cdots \subseteq T_{G^{\psi_{j-1}}} \subseteq T_{G^\sigma}$. We prove the following by induction on j; If a P-leaf t has t^* for $P \in \Gamma \setminus \{P_1, \cdots, P_{j-1}\}$, then the tree of the formula A_{t^*} has a copy in T_F or T_G. We remark that for $t \in T_{F^{\psi_{j-1}}} \subseteq T_{F^{\psi_j}}$ the formula A_t is related to $T_{F^{\psi_{j-1}}}$ and $T_{F^{\psi_j}}$ and generally they are distinct.

We have chosen P_j from $\Gamma \setminus \{P_1, \cdots, P_{j-1}\}$ so that the height of P_j^σ is maximal among such P. By induction hypothesis $T_{F^{\psi_j}}$ and $T_{G^{\psi_j}}$ are defined by attaching trees corresponding to A_{t^*}. Hence their heights are less than or equal to h_0. Now let $P \in \Gamma \setminus \{P_1, \cdots, P_j\}$ and suppose that a P-leaf t has t^*. If t^* does not belong to $T_{F^{\psi_{j-1}}} \cup T_{G^{\psi_{j-1}}}$, then t^* belongs to a part added according to P_j and the tree of A_{t^*} has a copy in T_F or T_G. Suppose that t^* belongs to $T_{F^{\psi_{j-1}}} \cup T_{G^{\psi_{j-1}}}$. Since the height of P_j^σ is not less than that of P^σ, no trees are added above t^* in $T_{F^{\psi_{j-1}}} \cup T_{G^{\psi_{j-1}}}$ according to P_j. Therefore, the tree of the formula A_{t^*} is a copy in T_F or T_G by induction hypothesis.

Since the maximal height of T_F and T_G is h_0, in each step the height possibly increases at most h_0 and we have the property (3_i). The other properties are clear now. Since the cardinality of Γ is k, we have the conclusion. $\dashv$

Proof of Theorem 3.4. We prove this theorem by induction on the height of a proof. When the height of $\mathcal{P}$ is 1, the proof consists of an axiom. For instance, we deal with $\forall x A(x) \supset A(y)$.

Let $\overline{A}$ be the holed formulas of $A(x)$ and the numbers of $*$ be m. Then, we have $x_1, \cdots, x_m$ such that $A(x) \equiv \overline{A}(x_1, \cdots, x_m)$. Let the arities of predicate variables P be m and $P^\sigma \equiv \overline{A}$. Now $\forall x Px \supset Py.$ is an axiom and its height is 3, i.e. equal to or less than m_0. Let $y_1, \cdots, y_m$ be the sequence obtained from replacing x_i such that $x_i \equiv x$ in the sequence $x_1, \cdots, x_m$ by y.

We have

$$\begin{aligned} &(\forall x P x_1 \cdots x_m \supset P y_1 \cdots y_m)^\sigma \\ \equiv\ &\forall x (P^\sigma(x_1, \cdots, x_m) \supset P^\sigma(y_1, \cdots, y_m) \\ \equiv\ &\forall x (\overline{A}(x_1, \cdots, x_m) \supset \overline{A}(y_1, \cdots, y_m) \\ \equiv\ &\forall x A(x) \supset A(y). \end{aligned}$$

Let $\mathcal{K}$ be the proof consisting of one formula $\forall x P x_1 \cdots x_m \supset P y_1 \cdots y_m$ then σ is an admissible assignment and we get the conclusion. For other axioms we can prove similarly and omit their proofs.

Suppose that the last inference rule of $\mathcal{P}$ is

$$\text{(I2)} \qquad \frac{A(x)}{\forall x A(x)}$$

and $\mathcal{K}_0$ is a proof of X in L^+ such that $X^\sigma \equiv A(x)$, where σ is admissible for $\mathcal{K}_0$.

Let $\mathcal{K}$ be $\dfrac{\mathcal{K}_0}{\forall x X}$. Then, $\mathcal{K}$ is a proof. Now the admissibility of σ in $\mathcal{K}$ follows from that in $\mathcal{K}_0$. Since the maximal height of formulas possibly increases at most 1, we have the conclusion.

Let the last inference rule of $\mathcal{P}$ be the modus ponens

$$\text{(I1)} \qquad \frac{A \quad A \supset B}{B}$$

and the height of $\mathcal{P}$ is $n+1$. By induction hypothesis we have two proofs $\mathcal{K}_0$ and $\mathcal{K}_1$ in L^+, where $\mathcal{K}_0$ and $\mathcal{K}_1$ contain distinct predicate variables, and an assignment σ such that

(1) $\mathcal{K}_0^\sigma$ is a proof of A;
(2) $\mathcal{K}_1^\sigma$ is a proof of $A \supset B$;

(3) the heights of formulas in $\mathcal{K}_0$ or $\mathcal{K}_1$ are bounded by $g(n, m_0)$.

Let X_0 be the conclusion of $\mathcal{K}_0$ and X be that of $\mathcal{K}_1$. If X is a prime formula $Px_1 \cdots x_n$. Since $X^\sigma \equiv A \supset B$, we introduce a new predicate variable Q and define an assignment φ with its support $\{Q\}$ such that $P^\varphi \equiv \overline{A} \supset Q * \cdots *$ where the arity of $\overline{A}$ is m and that of Q is $n - m$. We extend σ so that $Q^\sigma \equiv \overline{B}$. Now Lemma 4.2 implies the conclusion.

Otherwise, i.e. X is not a prime formula, then we have X_1 and Y such that $X \equiv X_1 \supset Y$, $X_1^\sigma \equiv A$ and $Y^\sigma \equiv B$. If $X_0 \equiv X_1$, then we can apply the modus penons and see that the other requirements are satisfied. Otherwise we apply Lemma 4.3. Since the heights are bounded by $g(n, m_0)$, the number of predicates variables occurring in X_0 or X_1 is less than or equal to $2 \cdot 2^{g(n,m_0)}$. By Lemma 4.1 the heights of formulas are bounded by $g(n, m_0)(2^{g(n,m_0+1} + 1)$. ⊣

Proof of Theorem 2.1. Since the first statement is proved in Lemma 2.3, it suffices to show the second one. To show by contradiction, suppose that $\mathcal{P}$ is a proof of $\neg\forall x(x = x)$ from K in M and the height of $\mathcal{P}$ is a standard number n. We apply Theorem 3.4 in M. Then we have a core proof $\mathcal{K}$ of $\mathcal{P}$ such that the heights of formulas occurring in $\mathcal{K}$ are bounded by $g(n, m_0)$. Replace all prime formulas $Px_1 \cdots x_n$ by $\forall x(x = x)$ for predicate variables P in $\mathcal{K}$. Then we have a proof $\mathcal{Q}$ of $\neg\forall x(x = x)$ from K such that the heights of formulas occurring in $\mathcal{Q}$ are bounded by $g(n, m_0)$. Let c be the maximal length of a prime formula in K, which is a standard number. Then, the lengths of formulas occurring in $\mathcal{Q}$ are bounded by a standard number $(c + 1)g(n, m_0) + c$, which contradicts Lemma 2.3. ⊣

REMARK 4.1. We did not restrict numbers of variables in our proof of getting a bound for heights. Such restrictions are possibly meaningful, when we estimate the length of formulas more precisely. As we have mentioned earlier, our result concerns systems. Lemma 2.3 for *PA* with a standard formulation can be proved similarly, but we have not succeeded to prove Theorem 2.1 for it.

REFERENCES

[1] K. GÖDEL, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls*, ***Monatsh. Math. Phys.***, vol. 37 (1930), pp. 349–360.

[2] ———, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme i*, ***Monatsh. Math. Phys.***, vol. 38 (1931), pp. 173–198.

[3] T. JECH, ***Set theory***, Academic Press, 2003.

[4] K. Kunen, ***Set theory***, North Holland, 1980.

DEPARTMENT OF MATHEMATICS, WASEDA UNIVERSITY, TOKYO 169-8555, JAPAN
E-mail: eda@logic.info.waseda.ac.jp

E-mail: uesu@leaf.ocn.edu