

**A NONCONSTRUCTIVE PROOF OF GENTZEN'S
HAUPTSATZ FOR SECOND ORDER
PREDICATE LOGIC**

BY W. W. TAIT

Communicated by D. Scott, April 26, 1966

Takeuti [3] showed that the consistency of analysis (i.e. second order number theory) is finitistically implied by the *Hauptsatz* for second order logic, i.e. by the proposition that every theorem of this system is derivable without cut.¹ We will prove that, conversely, the *Hauptsatz* for this system follows from a certain generalization of the consistency of analysis; namely from:

I. *Every countable set of relations among natural numbers is included in an ω -model.*

An ω -model is a collection of relations among natural numbers which is closed under the second order comprehension axiom. Henkin [1] has shown that a second order formula is derivable with the cut rule if and only if it is valid in all (countable) ω -models.

When the given set of relations consists only of the successor relation, I asserts the consistency of analysis.

The formalism for second order predicate logic which we will use is obtained from the system of predicate logic of finite order given in Schütte [2] by dropping all expressions and bound variables of types other than 0 (individuals), 1 (propositions) and $(0, 0, \dots, 0)$ (relations among individuals). Thus, expressions of type 0 are built up from constants and free variables of type 0 using function constants. The expressions of type $(0, \dots, 0)$ are constants, free variables and expressions $\lambda x_1^0 \dots x_n^0 A(x_1^0, \dots, x_n^0)$, where $A(a_1^0, \dots, a_n^0)$ is a wff (expression of type 1). The logical symbols other than λ are \neg, \vee and \forall .

The notation and terminology of [2] will be assumed. In particular, the notions of *strict derivation* and *partial valuation* will be the same as in [2], except that they refer to the second order logic and not the full system of [2], and that we require of a partial valuation that whenever $\forall x^r A(x^r)$ is true (t), then so is $A(a^r)$ for some free

¹ Actually, Takeuti asserts this result, not for second order logic, but for G^1LC which contains constants and free variables for third order relations. His proof however is valid for second order logic. On the other hand, the proof of the result of this paper cannot be extended to G^1LC , as far as I know.

variable a^τ (and not simply for some expression of type τ , as in [2], 6.1.5). With this modification, the proof of [2], 6.5 still goes through (since the partial valuation actually constructed satisfies our stronger condition); and so from 5.6 and 6.5:

II. *If a wff is not strictly derivable, then there is a partial valuation in which it is false (f).*

Since derivability without cut is equivalent to strict derivability (see [2], p. 306), we can prove the desired result by showing that every wff which is false in some partial valuation is also false in some ω -model (for certain values of its constants and free variables). Of course, if we identify the terms of type 0 with the integers ≥ 0 , then a total valuation is an ω -model. Hence, we need only consider *proper* partial valuations, i.e., ones which are not total.

Let b be a free variable of type $\neq 0$. That is, b is a propositional variable or else it ranges over relations among individuals. To each wff A we assign the expression A^b , which is either a wff or else the symbol ϕ , as follows: If A is prime and contains b , then it is of the form b (if b is of type 1) or $(e_1^0, \dots, e_n^0 \in b)$. In this case $(\pm A)^b = \phi$. \pm refers of course to A and $\neg A$. If A is prime and does not contain b , then $(\pm A)^b = \pm A$. $(\neg\neg A)^b = \neg\neg A^b$ if $A^b \neq \phi$; otherwise, $(\neg\neg A)^b = \phi$. $(A \vee B)^b = A^b \vee B^b$, A^b , B^b or ϕ , according to whether $A^b \neq \phi \neq B^b$, $A^b \neq \phi = B^b$, $A^b = \phi \neq B^b$ or $A^b = \phi = B^b$, resp. $(\neg(A \vee B))^b = \neg(A_1 \vee B_1)$ if $(\neg A)^b = \neg A_1$ and $(\neg B)^b = \neg B_1$. If either $(\neg A)^b$ or $(\neg B)^b = \phi$, then $(\neg(A \vee B))^b = \phi$. $(\forall x B(x))^b = \forall x B(x)^b$ or ϕ according to whether $A(c)^b \neq \phi$ or $A(c)^b = \phi$. (Here c is chosen $\neq b$.) $(\neg \forall x B(x))^b = \neg \forall x B_1(x)$ if $(\neg B(x))^b = \neg B_1(x)$. If $(\neg B(x))^b = \phi$, then $(\neg \forall x B(x))^b = \phi$. $(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n))^b = (e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n)^b)$ or ϕ according to whether $B(x_1, \dots, x_n)^b \neq \phi$ or $= \phi$. Finally, $(\neg(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n)))^b = \neg(e_1, \dots, e_n \in \lambda x_1 \dots x_n B_1(x_1, \dots, x_n))$ or ϕ according to whether $(\neg B(x_1, \dots, x_n))^b = \neg B_1(x_1, \dots, x_n)$ or $= \phi$.

Let $b^\tau = b$ be fixed ($\tau \neq 0$), and let $A(x)$ denote the result of replacing all occurrences of b in the wff A by $x^\tau = x$.

III. *Let A^b be $\neq \phi$ and be satisfied in the ω -model M for certain values assigned to its constants and free variables. Then $\lambda x A(x)$ is satisfied in M by the same values.*

The proof is by induction on A . If A is prime then it does not contain b (since $A^b \neq \phi$), and so $\lambda x A(x)$ is logically implied by $A^b = A$. If $A = A_1 \vee A_2$, then for $i = 1$ or 2 , A_i^b is satisfied, and so by the induction hypothesis, so is $\lambda x A_i(x)$, from which $\lambda x A(x)$ follows. If $A = \neg(A_1 \vee A_2)$, then $\neg A_1^b$ and $\neg A_2^b$ are satisfied, and hence, so are $\lambda x \neg A_1(x)$ and $\lambda x \neg A_2(x)$, from which $\lambda x A(x)$ follows. If $A = \neg\neg B$, then B^b is satis-

fied, and hence so is $\Lambda xB(x)$ from which $\Lambda xA(x)$ follows. If $A = VyB(y)$, then $B(c)_1^b$ is satisfied for some value of c (chosen $\neq b$), and so $\Lambda xB(c, x)$ is satisfied for this value of c . Therefore, $\Lambda xA(x) = \Lambda xVyB(y, x)$ is satisfied. If $A = \neg VyB(y)$, then $\neg B(c)^b$ is satisfied for all values of c , and hence, so is $\Lambda x\neg B(c, x)$. Therefore, $\Lambda xA(x) = \Lambda x\neg VyB(y, x)$ is satisfied. If $A = \pm(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n))$, then $\pm B(e_1, \dots, e_n)^b$ is satisfied, and hence, so is $\Lambda x \pm B(e_1, \dots, e_n, x)$, from which $\Lambda xA(x) = \Lambda x \pm(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n, x))$ follows.

The *depth* dA of the variable b in the wff A is inductively defined as follows: If b does not occur in A , then $dA = 0$. Assume that b does occur in A . If A is prime, $dA = 1$. If $A = BvC$, $dA = dB + dC + 1$. If $A = \neg B$, $dA = dB + 1$. If $A = VyB(y)$, $dA = dB(c) + 1$ (choosing $c \neq b$).

If $A = (e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n))$, then $dA = dB(c_1, \dots, c_n) + 1$ (choosing $c_i \neq b$).

Using the same notation as in III:

IV. If V is a partial valuation which is not total, and $VxA(x)$ is f in V , then so is $A^b (\neq \phi)$.

Since V is not total, there is a wff F to which V assigns no truth value. If τ (the type of b) = 1, set $e = F$, and if $\tau = (\tau_1, \dots, \tau_n)$, where each $\tau_i = 0$, set $e = \lambda x_1^0 \dots x_n^0 F$. Then since $VxA(x)$ is f , so is $A(e)$. By induction on the depth dA of b in $A = A(b)$, we prove that if $A(e)$ is f in V , then so is A^b . We can assume here that no wff to which V assigns a value contains b . If $dA = 0$, then b is not in A , so that $A^b = A = A(e)$. $dA = 1$ is impossible, since otherwise $A(e)$ is either e or else is of the form $(e_1, \dots, e_n \in e)$, and so $A(e)$ has no value for V . Assume $dA > 1$. If $A = A_1vA_2$, then $A_1(e)$ and $A_2(e)$ are f , and since $dA_i < dA$, it follows from the induction hypothesis that A_1^b and A_2^b are f . Hence, A^b is f . If $A = \neg(A_1vA_2)$, then $\neg A_i(e)$ is f for some $i = 1, 2$, and so $\neg A_i^b$ is f (since $dA_i < dA$), and hence, so is A^b . If $A = \neg\neg B$, then $B(e)$ is f , and so since $dB < dA$, B^b is f . Hence, A^b is f . If $A = Vy^\sigma B(y^\sigma)$, then for all e^σ , $B(e, e^\sigma)$ is f . Since e^σ does not contain b , $dB(e^\sigma) < dA$, and so $B(e^\sigma)^b$ is f . Hence, $A = Vy^\sigma B(y^\sigma)$ is f . If $A = \neg Vy^\sigma B(y^\sigma)$, then for some variable c^σ with $c^\sigma \neq b$, $\neg B(c^\sigma)$ is f . Since $d\neg B(c^\sigma) < dA$, $(\neg B(c^\sigma))^b = \neg B_1(c^\sigma)$ is f , and so $A^b = \neg Vy^\sigma B_1(y^\sigma)$ is f . If $A = \pm(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n))$ then $\pm B(e_1, \dots, e_n, e)$ is f , and since $d \pm B(e_1, \dots, e_n) (= d \pm B(e_1, \dots, e_n, b)) < dA$, $(\pm B(e_1, \dots, e_n))^b$ is f . Hence, A^b is f .

Let V be a partial valuation which is not total. If f is a function constant of n arguments, we set $f^* = \{(e_1, \dots, e_n, f(e_1, \dots, e_n)) : \text{all expressions } e_1, \dots, e_n \text{ of type } 0\}$. If α is a constant or variable of type $(0, \dots, 0)$ (n arguments), we set $\alpha^* = \{(e_1, \dots, e_n) :$

$(e_1, \dots, e_n \in \alpha)$ has the value t for V }. If α is a constant or variable of type 1, we set $\alpha^* = t$ if V assign α t , and otherwise $\alpha^* = f$. If we identify the expressions of type 0 with the integers ≥ 0 , then, by **I**, the relations f^* and α^* are included in an ω -model M .

V. *If A is $t(f)$ in V , then it is $t(f)$ in M , with f denoting f^* and α denoting α^* .*

We prove this by induction on the number of occurrences of logical constants in A . If A is prime, the result is clear. If $A = \angle B$ or BvC , the result follows by the induction hypothesis applied to its components. Let $A = Vx^0B(x^0)$. If A is true in V , then so is $B(a^0)$ for some a^0 ; and so $B(a^0)$ is true in M , and hence, so is A . If A is false in V , then $B(e^0)$ is false in V , and so in M , for all e^0 . But since the domain of individuals of M consists of the e^0 , A is false in M . Let $A = Vx^\tau B(x^\tau)$ where $\tau \neq 0$. If A is true in V , then so is $B(a^\tau)$ for some variable a^τ , and so $B(a^\tau)$, and hence A , are true in M . If A is false in V , then by **IV**, so is $B(b^\tau)^{b^\tau}$ (choosing b^τ not in A), and so $B(b^\tau)^{b^\tau}$ is false in M by the induction hypothesis. Therefore, by **III**, A is false in M . Let $A = \pm(e_1, \dots, e_n \in \lambda x_1 \dots x_n B(x_1, \dots, x_n))$. Then if A is t (f) in V , so is $\pm B(e_1, \dots, e_n)$. Hence $\pm B(e_1, \dots, e_n)$ is t (f) in M , and therefore, so is A .

THEOREM. *If A is derivable, then it is derivable without cut.*

If A is not derivable without cut, i.e. strictly derivable, then there is a partial valuation V in which it is false. If V is a total valuation, then A cannot be derivable by [2, 7.4]. If V is not total, then by **V**, there is an ω -model in which A is false, and so, in any case, A is not derivable. Q.E.D.

REFERENCES

1. L. Henkin, *Completeness in the theory of types*, J. Symbolic Logic **15** (1950), 89-91.
2. K. Schütte, *Properties of simple type theory*, J. Symb. Logic **25** (1960), 305-326.
3. G. Takeuti, *On a generalized logical calculus*, Japan. J. Math. **23** (1953), 39-96.

UNIVERSITY OF ILLINOIS AT CHICAGO CIRCLE