A SIMPLE TREATMENT OF CHURCH'S THEOREM ON THE DECISION PROBLEM

Thomas SCHWARTZ

The proof I will give of Church's undecidability theorem for quantification theory is simpler than any other known to me, and it is more direct: the non-effectiveness of quantificational validity is not inferred from some *other* unsolvability result.

Church's Theorem is deduced from his Thesis (CTh) that every computable function in the set Z of natural numbers is recursive.

To define recursiveness we set up a formal system whose vocabulary consists of o, S, $\equiv$, the parentheses and comma, the *variables* x, y, z, w, w_1, w_2, etc. and for all i, k, the k-ary *function letter* ${}^kf_i$.

S and the ${}^kf_i$ are *function symbols*.

Terms are defined thus: variables and o are terms, and if $\tau_1, \ldots, \tau_k$ are terms, so are $S(\tau_1)$ and ${}^kf_i(\tau_1, \ldots, \tau_k)$.

o, $S(o)$, $S(S(o)))$ etc. are *numerals*. For n in Z, the numeral n is defined by setting $\bar{0} = o$ and $\overline{k+1} = \mathrm{S}(\bar{k})$.

An *equation* is $\equiv$ flanked by terms.

If E is a set of equations, an *E-letter* is a function letter in E.

"$E \vdash e$" means e belongs to the set E of equations or is derivable from E by these rules:

R1. If B results from substituting numerals for variables in A, from A to infer B.

R2. If $A(b)$ results from replacing one occurrence of a in $A(a)$ by b, from $a \equiv b$ and $A(a)$ to infer $A(b)$.

A *recursion* is a finite set E of equations such that (1) for all numerals a, b, if $\mathrm{E} \vdash a \equiv b$ then $a = b$, and (2) for all E-letters ${}^kf_i$ and $n_1, \ldots, n_k$ in Z, there is a unique n for which $E \vdash {}^kf_i(\bar{n}_1, \ldots, \bar{n}_k) \equiv \bar{n}$.

If E is a recursion and h a k-ary E-letter, $\bar{h}^E$ is the function defined on Z^k thus: $\bar{h}^E(n_1, \ldots, n_k) = n$ iff $E \vdash h(\bar{n}_1, \ldots, \bar{n}_k) \equiv \bar{n}$.

A function φ is *recursive* iff $\varphi = \bar{h}^E$ for some recursion E and E-letter h.

If $\Pi \subset Z$, the *characteristic function* Γ_{Π} of Π is defined on Z by setting $\Gamma_{\Pi}(n) = 0$ if $n \varepsilon \Pi$ and $\Gamma_{\Pi}(n) = 1$ if not $n \varepsilon \Pi$.

Clearly, if Π is effective, Γ_{Π} is computable, and so, by CTh, Γ is recursive.

I will show there is no effective test of (quantificational) validity for the language L whose well-formed formulas (wffs) are generated in the usual way from the parentheses and comma, the universal quantifier $\forall$ and conditional sign $\rightarrow$, the terms introduced above and the 2-place predicate $\equiv$ and 1-place predicate P.

"$\models A$" means A is a valid wff.

LEMMA. *If Π is an effective subset of Z, there is an A such that for all n in Z,* $\models A \rightarrow P(\bar{n})$ iff $n \varepsilon \Pi$.

Proof. By CTh there is a recursion E with E-letters $h_1, \ldots, h_m$ such that $\bar{h}^E_p = \Gamma_{\Pi}$ for some $p = 1, \ldots, m$.

Let $\bar{L}$ be L shorn of all function letters but E-letters and interpreted in Z by assigning 0 to o, the successor function to S, identity to $\equiv$, Π to P and $\bar{h}^E{}_i$ to h_i ($i = 1, \ldots, m$).

Let A be a conjunction of universal closures of the wffs in E plus these wffs:

(1) $h_p(x) \equiv o \rightarrow P(x)$, (2) $x \equiv y \rightarrow (x \equiv z \rightarrow y \equiv z)$.

(3) $(x \equiv z \rightarrow y \equiv z) \rightarrow (z \equiv x \rightarrow z \equiv y)$,

(4gi) $(x \equiv z \rightarrow y \equiv z) \rightarrow (g(w_1, \ldots, w_{i-1}, x, w_{i+1}, \ldots, w_r) \equiv z \rightarrow g(w_1, \ldots, w_{i-1}, y, w_{i+1}, \ldots, w_r) \equiv z)$, for all function symbols g of $\bar{L}$ and $i = 1, \ldots, r$ (r = the number of arguments of g).

$\bar{L}$ and A have these two properties:

(I) (α) *If* $E \vdash \tau \equiv \sigma$, $\tau \equiv \sigma$ *is true for all values* $q_1, \ldots, q_v$ *of its variables* $t_1, \ldots, t_v$. (β) *A is true.*

Proof. (α) By induction on ν plus the number 1 of occurrences of E-letters in $\tau \equiv \sigma$.

If $\nu = 1 = 0$, τ and σ are numerals; thus, since E is a recursion, $\tau = \sigma$, so $\tau \equiv \sigma$ is true.

If $\nu = 0$ and $1 > 0$, $\tau \equiv \sigma$ has a part $h_i(\bar{r}_1, \ldots, \bar{r}_k)$. Let $\bar{h}^E{}_i(r_1, \ldots, r_k) = r$.

Then $E \vdash h_i(\bar{r}_1, \ldots, \bar{r}_k) \equiv \bar{r}$.

Let $\tau' \equiv \sigma'$ result from replacing one occurrence of $h_i(\bar{r}_1, \ldots, \bar{r}_k)$ in $\tau \equiv \sigma$ by $\bar{r}$.

Since $h_i(\bar{r}_1, \ldots, \bar{r}_k)$ and $\bar{r}$ both name r, $\tau \equiv \sigma$ is true if $\tau' \equiv \sigma'$ is.

But by R2, $E \vdash \tau' \equiv \sigma'$, so by inductive hypothesis, $\tau' \equiv \sigma'$ is true, hence so is $\tau \equiv \sigma$.

If $\nu > 0$ and $\tau^* \equiv \sigma^*$ results from substituting each $\bar{q}_i$ for t_i in $\tau \equiv \sigma$, $E \vdash \tau^* \equiv \sigma^*$ by R_1, so $\tau^* \equiv \sigma^*$ is true by inductive hypothesis, so $\tau \equiv \sigma$ is true for $q_1, \ldots, q_v$.

(β) (1)-(4*gi*) are obviously true for all values of their variables, and by (α), so are the wffs in E. So A is true.

(II) (α) *If* $\models A \rightarrow a \equiv b$ *and* $\tau(b) \equiv \sigma$ *follows from* $a \equiv b$ *and* $\tau(a) \equiv \sigma$ *by R2,* $\models A \rightarrow (\tau(a) \equiv \sigma \rightarrow \tau(b) \equiv \sigma)$.

(β) *If* $\models A \rightarrow a \equiv b$ *and* $\tau \equiv \sigma(b)$ *follows from* $a \equiv b$ *and* $\tau \equiv \sigma(a)$ *by R2,* $\models A \rightarrow (\tau \equiv \sigma(a) \rightarrow \tau \equiv \sigma(\mathrm{b}))$.

Hence, (γ) *if* $\models A \rightarrow a \equiv b$, $\models A \rightarrow \tau \equiv \sigma$ *and* $\tau' \equiv \sigma'$ *follows from* $a \equiv b$ *and* $\tau \equiv \sigma$ *by R2,* $\models A \rightarrow \tau' \equiv \sigma'$.

Proof. (α) By induction on the length of $\tau(a)$.

If $a = \tau(a)$, $b = \tau(b)$, so $\models A \rightarrow \tau(a) \equiv \tau(b)$, whence $\models A \rightarrow (\tau(a) \equiv \sigma \rightarrow \tau(b) \equiv \sigma)$ by (2).

If a is a proper part of $\tau(a)$, $\tau(a)$ has the form $g(d_1, \ldots, d_{i-1}, \varrho(a), d_{i+1}, \ldots, d_r)$ and $\tau(b)$ the form $g(d_1, \ldots, d_{i-1}, \varrho(b), d_{i+1}, \ldots, d_r)$.

By inductive hypothesis, $\models A \rightarrow (\varrho(a) \equiv \sigma \rightarrow \varrho(b) \equiv \sigma)$, so by (4*gi*), $\models A \rightarrow \tau(a) \equiv \sigma \rightarrow \tau(b) \equiv \sigma)$.

(β) $\models A \rightarrow (\sigma(a) \equiv \tau \rightarrow \sigma(b) \equiv \tau)$ by (α), so $\models A \rightarrow (\tau \equiv \sigma(a) \rightarrow \tau \equiv \sigma(b))$ by (3).

The set of consequences of A includes E and is closed under R1 (since R1 is valid) and R2 (by (IIγ)). So if $E \vdash e$, $\models A \rightarrow e$. But if $n \varepsilon \Pi$, $\Gamma_\Pi, \Gamma_\Pi(n) = \bar{h}^{\mathrm{E}} p(\mathrm{n}) = 0$, whence $E \vdash h_{\mathrm{p}}(\bar{n}) \equiv o$, so $\models A \rightarrow h_p(\bar{n}) \equiv o$, and thus $\models A \rightarrow P(\bar{n})$ by (1). Conversely, if $\models A \rightarrow P(\bar{n})$, $P(\bar{n})$ is true by (Iβ), so $n \varepsilon \Pi$.

Theorem. The class of valid wffs of L is not effective.

Proof. Suppose the theorem false. Treating each superscript and subscript digit as a separate symbol, L has a finite vocabulary, so we get an effective enumeration A_0, A_1, A_2 etc. of L's expressions by listing L's symbols in some chosen "alphabetic" order, then the 2-symbol expressions in lexicographic order, next the 3-symbol expressions etc. Thus, since validity is effective, we can effectively tell for each n in Z whether or not $\models A_n \to P(\bar{n})$. Hence $\Pi = \{n \varepsilon Z \mid \text{not} \models A_n \to P(\bar{n})\}$ is effective. So by the Lemma, there is a k such that for all n in Z, $\models A_k \to P(\bar{n})$ iff $n \varepsilon \Pi$. So $\models A_k \to P(\bar{k})$ iff $k \varepsilon \Pi$, i.e. $\models A_k \to P(\bar{k})$ iff not $\models A_k \to P(\bar{k})$, a contradiction !

University of Pittsburgh Thomas SCHWARTZ