INSTITUTUL
DE
MATEMATICA

INSTITUTUL NATIONAL
PENTRU CREATIE
STIINTIFICA SI TEHNICA

ISSN 0250-3638

HERBRAND'S THEOREM AND ELIMINATION OF QUANTIFIERS

by

Şerban A. BASARAB

PREPRINT SERIES IN MATHEMATICS

No.52/1980

BUCURESTI

HERBRAND'S THEOREM AND ELIMINATION OF QUANTIFIERS

by

Şerban A. BASARAB*)

October 1980

*) Institute of Mathematics, Str. Academiei 14, Bucharest, Romania

HERBRAND'S THEOREM AND ELIMINATION OF QUANTIFIERS

by Serban A. Basarab

The aim of this note is to show that the necessary and sufficient condition for a first-order theory T "to satisfy Herbrand's theorem" is that T admits elimination of quantifiers.

As Herbrand's theorem is the basis for most modern automatic proof procedures it seems that the above characterization could be of some interest for the mechanical theorem proving.

§ 1. Background from Logic

Let τ be a similarity type and L_τ the first-order language with equality attached to τ (see for definitions [3] or [8]). We consider in the present work only the case when τ is at most countable.

Definition. Lt H_0 be the set of constants appearing in τ. If no constant appears in τ, then H_0 is to consist of a single constant, say $H_0 = \{a\}$. For $i \in \mathbb{N}$, let H_{i+1} be the union of H_i and the set of all terms of the form $f(t_1, \ldots, t_n)$ for all n-place function symbols f occuring in τ, where $t_1, \ldots, t_n$ are members of the set H_i. Then the Herbrand universe of τ (or of the languange L_τ) is the union $H = H(\tau) = \bigcup_{i \in \mathbb{N}} H_i$.

Definition. The union of the set of atomic formulas of the form $R(t_1, \ldots, t_n)$ for all n-place relation symbols R occuring in τ, where $t_1, \ldots, t_n$ are elements of $H(\tau)$, and the set of atomic formulas of the form $t_1 = t_2$, where t_1, t_2 are elements of $H(\tau)$, is called the atom set, or the Herbrand base $B=B(\tau)$ of τ.

Definition. A literal is either a member of $B(\tau)$ or the negation $\neg\varphi$ of a member φ of $B(\tau)$. Denote by $B^* = B^*(\tau)$ the set of literals attached to τ.

Now let T be a theory in L_τ, i.e. a consistent set of sentences in L_τ. In particular we can consider the case when T is the empty set ϕ. In this case the deductive closure $\bar{\phi}$ is the set of all sentences in L_τ derived from the logical axioms of L_τ.

Definition. A subset I of B^* is called a T-*interpretation* on the Herbrand universe H if the union $T \cup I$ is consistent. The T-interpretation I is called *total* if I is maximal with respect to inclusion in the set of T-interpretations on H. In other words, the T-interpretation I is total iff for every atomic formula $\varphi \in B$ either $\varphi \in I$ or $\neg\varphi \in I$.

It follows easily that every T-interpretation on H can be extended to a total T-interpretation on H. On the other hand every total T-interpretation on H induces a canonical structure of type τ on the factor set $H_I = H/\sim$, where the equivalence relation $\sim$ is defined as follows: for arbitrary $t_1, t_2 \in H$, let $t_1 \sim t_2$ iff $(t_1 = t_2) \in I$. Equipped with the structure of type τ induced by I, H_I becomes by Gödel's consistency theorem [8] Theorem 7.1 a substructure of some model of T.

§ 2. Semantic trees

The semantic trees are very useful tools in the proof theory (see for instance [1], [4], [7]).

Let τ be a similarity type, and L_τ the frist order language associated to τ .

Definition. Given a theory T in L_τ , a T-semantic tree on the Herbrand universe H of τ is a (downward) tree A, where each link is attached with a non-empty finite set of literals from B^* in such a way that:

i) For each node N, there are only finitely many immediate links $L_1, \ldots L_n$ from N. Let Q_i be the set of literals attached to L_i, $i = 1,\ldots,n$. Denote by $\overline{Q}_i$ the conjuction of all the literals in the set Q_i, $i = 1,\ldots,n$. Then $T \vdash \overline{Q}_1 \vee \overline{Q}_2 \vee \ldots \vee \overline{Q}_n$

ii) For each node N, let I(N) be the union of all the sets attached to the links of the path of A down to and including N. Then I(N) is a T-interpretation on H.

If D is a path of a T-semantic tree A on H we denote by I(D) the union $\bigcup I(N)$ where N ranges over the set of nodes of D. Thanks to the finitary character of the consequence relation $\vdash$ it follows that I(D) is also a T-interpretation on H. The set I(D) is finite iff the path D is finite; if this is the case then I(D) = I(N) where N is the terminal node of D.

Definition. A T-semantic tree A on H is said to be *total* if for each branch (i.e. maximal path) D of A, I(D) is a total T-interpretation on H.

When the Herbrand base B of τ is infinite, any total T-semantic tree on H will be infinite. As is easily seen, if A is a total T-semantic tree on H then for each T-interpretation I on H there is at least one branch D of A such that $I \subset I(D)$. Thus, a total T-semantic tree on H corresponds to an exhaustive survey of all possible total T-interpretations on H.

Now let S be a set of sentences in the language L_τ.

Definition. A node N in a T-semantic tree A on H is a *failure node for* S if there is at least one sentence $\varphi \in S$ such that $T \cup I(N) \vdash \neg\varphi$, but for every ancestor node N' of N and for every sentence $\varphi \in S$, the set $T \cup I(N') \cup \{\varphi\}$ is consistent.

Definition. A T-semantic tree A on H is said to be *closed for* S if every branch of A terminates at a failure node.

By König's lemma [3] § 49 Lemma lo it follows that every closed (for S) T-semantic tree A is a finite tree.

Examples

i) Let $\tau = \{a, f, P, Q\}$ where a is a constant, f is a one place function symbol, and P, Q are one-place relation symbols. Let T be the empty set, and S be the set containing the following sentences:

$$(\forall x)\, P(x)$$

$$(\forall x)\, \neg P(x) \vee Q(f(x))$$

$$\neg Q\,(f(a))$$

Figure 2.1 shows a closed (for S) T-semantic tree on H(τ).

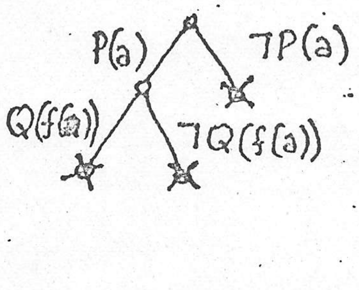

Figure 2.1

ii) Let $\tau = \{f, P\}$, where f is a one-place function symbol and P is a one-place relation symbol. Let T be the empty set, and let $S = \{(\forall x)\, P(x),\ (\exists x) \neg P(f(x))\}$.

Figure 2.2 shows a T-semantic tree on $H(\tau)$ which is not closed for S. In fact it is easy to see that any T-semantic tree on $H(\tau)$ is not closed for S.

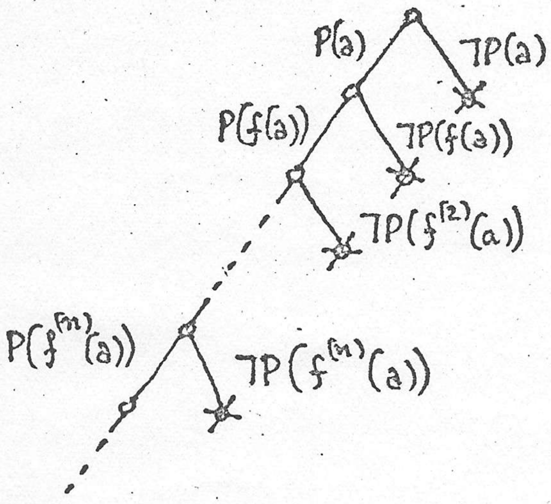

Figure 2.2

iii) Let $\tau = \{0, 1, -, +, \cdot, \leq\}$, where 0 and 1 are constants, − is a one-place function symbol, + and · are two-place function symbols, and $\leq$ is a two-place relation symbol. Let T be the theory of real closed, ordered fields, where the symbols of τ have the usual interpretations. Let $\tau' = \tau \cup \{a, b, c\}$ where a, b, c are new constants and let S be the set containing the following two sentences in $L_{\tau'}$:

$$(\exists x)\ ax^2 + bx + c = 0$$

$$(\forall x)\ ax^2 + bx + c \neq 0$$

Figure 2.3 shows a closed (for S) T-semantic tree on $H(\tau')$.

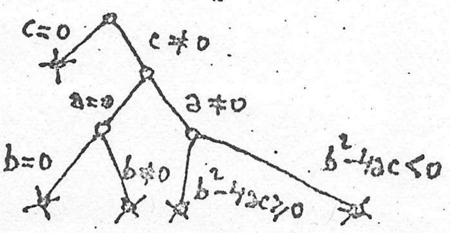

Figure 2.3

The following lemma folows easily from definitions.

(2.1) Lemma. Let A be a finite T-semantic tree on $H(\tau)$ and let $N_1,\dots,N_n$ be the terminal nodes of A. Denote by $\overline{I(N_i)}$ the conjunction of the literals belonging to $I(N_i)$, $i = 1,\dots,n$. Then $T \vdash \overline{I(N_1)} \vee \overline{I(N_2)} \vee \dots \vee \overline{I(N_n)}$.

The following result gives a sufficient condition for the set of sentences $T \cup S$ to be inconsistent.

(2.2) Proposition. If there exists a T-semantic tree on $H(\tau)$ which is closed for S then $T \cup S$ is inconsistent.

Proof. Let A be a T-semantic tree on $H(\tau)$ which in closed for S. By König's lemma, A is a finite tree; let $N_1,\dots N_n$ denote the terminal nodes of A.

Since $N_1,\dots,N_n$ are failure nodes for S there exist some sentences $\varphi_i \in S$, $i = 1,\dots,n$ such that $T \cup I(N_i) \vdash \neg\varphi_i$, $i = 1,\dots,n$. It follows $T \vdash \overline{I(N_1)} \vee \dots \vee \overline{I(N_n)} \rightarrow \neg(\varphi_1 \wedge \dots \wedge \varphi_n)$ On the other hand, by (2.1), $T \vdash \overline{I(N_1)} \vee \dots \vee \overline{I(N_n)}$ and hence $T \cup \{\varphi_1,\dots,\varphi_n\}$ is inconsistent. We conclude that $T \cup S$ is inconsistent too. Q.E.D.

Remark. The converse is not generally true (see the example ii). This motivates the following definition.

Definition. Let τ be a similarity type and T be a theory in L_τ. T is called a Herbrand theory if for each similarity type τ' which extends τ only with constants and for each set of sentences S in $L_{\tau'}$, there exists a closed (for S) T-semantic tree on $H(\tau')$ iff $T \cup S$ is inconsistent.

§ 3. Herbrand theories and elimination of quantifiers

Definition. Let τ be a similarity type and T be a theory in L_τ. T admits elimination of quantifiers if for eac formula $\varphi(x_1,\dots,x_n)$ in L_τ there is a quantifier free formul $\psi(x_1,\dots,x_n)$ in L_τ such that $T \vdash \varphi(x_1,\dots,x_n) \leftrightarrow \psi(x_1,\dots$

As examples of theories which admit elimination of quantifiers let us mention here the following remarkable theories:

i) The theory of algebrically closed fields [8] Corollary 13.3.

ii) The theory of real closed ordered fields [8] Corollary 17.4 (for the converse of Tarski's theorem see [6]).

iii) The theory of p-adic fields [2],[5].

The following result is well known (see for instance [8] Theorem 13.1). For the convenience of the reader we include here a proof.

(3.1) Lemma. Let τ be a similarity type, T a theory in L_τ and I a total T-interpretation on $H(\tau')$, where τ' is an extension of τ with new constants. Then $T \cup I$ is a complete theory in $L_{\tau'}$, if T admits elimination of quantifiers.

Proof. Let φ be a sentence in $L_{\tau'}$. Since by hypothesis T admits elimination of quantifiers it follows that there is a sentence ψ in $L_{\tau'}$, without quantifiers such that $T \vdash \varphi \leftrightarrow \psi$. On the other hand ψ is logically equivalent with a sentence in $L_{\tau'}$, of the form $\overline{\Psi}_1 \vee \ldots \vee \overline{\Psi}_n$, where Ψ_i, $i = 1, \ldots, n$, is a finite non-empty set of literals in $B^*(\tau')$; here $\overline{\Psi}_i$ denotes as above the conjunction of the literals belonging to Ψ_i, $i = 1, \ldots, n$. Since I is a total T-interpretation on $H(\tau')$ the following statements are equivalent:

i) $T \cup I \vdash \psi$

ii) there exist $i \in \{1, \ldots, n\}$ such that $\Psi_i \subset I$.

It follows that either $T \cup I \vdash \psi$ or $T \cup I \vdash \neg\psi$ and hence either $T \cup I \vdash \varphi$ or $T \cup I \vdash \neg\varphi$. We conclude that the theory $T \cup I$ is complete as contended. Q.E.D.

The main result of the present note is the following.

(3.2) Theorem. The necessary and sufficient condition for a theory T to be a Herbrand theory is that T admits elimination of quantifiers.

Proof. Assume that T is a Herbrand theory. We have to show that T admits elimination of quantifiers. Let $\varphi(x_1, \ldots, x_n)$ be a formula in the language of T, say L_τ. We have to show that there is a quantifier free formula $\psi(x_1, \ldots, x_n)$ in L_τ such that $T \vdash \varphi(x_1, \ldots, x_n) \leftrightarrow \psi(x_1, \ldots, x_n)$. Denote by τ' the extension of τ with the new constants $c_1, \ldots, c_n$, and let S be the set whose members are the following two sentences in $L_{\tau'}$: $\varphi(c_1, \ldots, c_n)$ and $\neg\varphi(c_1, \ldots, c_n)$. Since T is by hypothesis a Herbrand theory and S is inconsistent, there is at least one T-semantic tree A on $H(\tau')$ which is closed for S. Let A be such a tree; by König's lemma A is finite. The set of terminal nodes of A is a disjoint union of two subsets $\{N_1,\dots,N_m\}$ and $\{N'_1,\dots,N'_{m'}\}$ subject to: $T \cup I(N_i) \vdash \varphi(c_1,\dots,c_n)$, $i = 1,\dots,m$, and $T \cup I(N'_i) \vdash \neg\varphi(c_1,\dots c_n)$, $i = 1,\dots,m$

Let $\psi(c_1,\dots,c_n) := \overline{I(N_1)} \vee \dots \vee \overline{I(N_m)}$ and $\psi'(c_1,\dots,c_n) := \overline{I(N'_1)} \vee \dots \vee \overline{I(N'_{m'})}$. It follows $T \vdash \psi(c_1,\dots,c_n) \rightarrow \varphi(c_1,\dots,c_n)$ and $T \vdash \psi'(c_1,\dots,c_n) \vdash \neg\varphi(c_1,\dots,c_n)$.

On the other hand, by (2.1), $T \vdash \psi(c_1,\dots,c_n) \vee \psi'(c_1,\dots,c_n)$ and hence $T \vdash \psi(c_1,\dots,c_n) \leftrightarrow \varphi(c_1,\dots,c_n)$. As the constants $c_1,\dots,c_n$ don't appear in T we can replace them by variables. Thus we obtain $T \vdash \varphi(x_1,\dots,x_n) \leftrightarrow \psi(x_1,\dots,x_n)$ where $\psi(x_1,\dots,x_n)$ is a quantifier free formula in L_τ. We conclude that T admits elimination of quantifiers.

Conversely, let us suppose that T admits elimination of quantifiers. We have to show that T is a Herbrand theory. Let τ' be an extension with new constants of the similarity type τ, and S be an arbitrary set of sentences in $L_{\tau'}$, such that $T \cup S$ is inconsistent. Let A be a total T-semantic tree on $H(\tau')$. We show that A contains a subtree closed for S, i.e. every branch D of A contains a failure node for S. Let D be an arbitrary branch of A. Then I(D) is a total T-interpretation on $H(\tau')$ and, by (3.1), $T \cup I(D)$ is complete. Since in addition $T \cup S$ is inconsistent it follows that there is at least one sentence $\varphi \in S$ such that $T \cup I(D) \vdash \neg\varphi$.

Thanks to the finitary character of the consequence relation $\vdash$, there exists a node $N \in D$ subject to $T \cup I(N) \vdash \neg\varphi$ and hence there exists a failure node for S on the branch D. As every branch of A contains a failure node for S we conclude that A contains a closed (for S) subtree, and hence T is a Herbrand theory, as contended. Q.E.D.

As a corollary we derive a version of Herbrand's theorem [1] Theorem 4.3.

Let τ be a similarity type and $L = L_\tau$ be the first order language attached to τ. Denote by $\tilde{\tau}$ the similarity type obtained from τ by Skolemization [8] § 11. Let $\tilde{L}$ be the first order language associated to $\tilde{\tau}$. Denote by Λ the family of sentences in $\tilde{L}$ of the following type:

$(\forall x_1)\ldots(\forall x_n)[(\exists x_{n+1})\varphi(x_1,\ldots,x_{n+1}) \rightarrow \varphi(x_1,\ldots,x_n,f_\varphi(x_1,\ldots,x_n))]$ where $\varphi(x_1,\ldots,x_{n+1})$, $n \geqslant 0$, ranges over the set of formulas of $\tilde{L}$. If T is a theory in L we denoteby $\tilde{T} = T \cup \Lambda$ the Skolemization of T. In particular Λ is the Skolemization of the empty set $\emptyset$. With these preparations we have:

(3.3) Corollary (Herbrand's theorem). Let S be a set of sentences in L. Then the necessary and sufficient condition for S to be inconsistent is that there exists a Λ-semantic tree on $H(\tilde{\tau})$ which is closed for S.

Proof. By [8] Exercise 11.3, S is inconsistent iff $\tilde{S} = \Lambda \cup S$ is inconsistent. On the other hand, by [8] Proposition 11.1, the theory Λ admits elimination of quantifiers, and hence, by (3.2), Λ is a Herbrand theory. Thus the inconsistency of $\tilde{S}$ is equivalent with the existence of a closed (for S) Λ-semantic tree on $H(\tilde{\tau})$.

Q.E.D.

References

[1] Chin-Liang Chang, and Richard Char-Tung Lee, Symbolic Logic and Mechanical Theorem Proving, Academic Press, New York, 1973.

[2] Paul J.Cohen, Decision procedures for real and p-adic fields Communications in Pure and Applied Mathematics, XXII(1969), 131-152.

[3] Stephen Cole Kleene, Mathematical Logic, John Wiley & Sons, Inc., New York, 1967.

[4] R.Kowalski, and P.Hayes, Semantic trees in automatic theorem proving, "Machine Intelligence", Vol.4 (B.Meltzer and D. Michie, eds), American Elsevier, New York, 1969, 87-101.

[5] Angus Mac Intyre, On definable subsets of p-adic fields, Journal of Symbolic Logic 41 (1976), 605-610.

[6] A.Mac Intyre, K.Mc Kenna, and L.van den Dries, Quantifier elimination in algebraic structures, to appear.

[7] J.A.Robinson, The generalized resolution principle, "Machine Intelligence", vol.3 (D.Michie, ed.), American Elsevier, New York, 1968, 77-94.

[8] Gerald E.Sacks, Saturated Model Theory, W.A.Benjamin, Inc., Massachusetts, 1972.