FIRST ORDER PREDICATE LOGIC WITHOUT NEGATION IS NP-COMPLETE*

Dexter Kozen

TR 77-307

April 1977

Department of Computer Science
Cornell University
Ithaca, New York 14853

*This research has been supported in part by National Science Foundation Grant DCR75-09433.

FIRST ORDER PREDICATE LOGIC WITHOUT NEGATION IS NP-COMPLETE*

Dexter Kozen
Department of Computer Science
Cornell University
Ithaca, NY 14853

Abstract

Techniques developed in the study of the complexity of finitely presented algebras are used to show that the problem of deciding validity of positive sentences in the language of first order predicate logic with equality is $\le_{\log}$-complete for NP.

*This research has been supported in part by National Science Foundation Grant DCR75-09433.

0. Introduction

In this paper we use techniques developed in [1,2] to prove a complexity result for first order predicate logic with equality, namely that deciding validity of positive sentences (those without occurrences of $\neg$) is $\le_{\log}$-complete for NP. This result again attests to the power of negation, as did [3,4] previously, since the general validity problem, even without equality, is undecidable (this result is originally due to Church; see [5] for a very elegant proof, due to Floyd).

It is a little surprising that the problem would be complete for some level of the polynomial time hierarchy rather than some "even" class like P or PSPACE, since universal as well as existential quantification is allowed. This is because universal quantifiers are easy to eliminate, and existential ones not so easy, as we will see.

In [2], we approached a similar problem, that of deciding truth of a sentence of the form

$$Q_1 x_1 \ldots Q_k x_k \; s(\bar{x}) = t(\bar{x})$$

interpreted over a finitely presented algebra, and showed that it was complete for PSPACE. Here we reduce the validity problem to the problem of deciding truth of the sentence under a particular interpretation, a term algebra similar to the algebras of [1,2], so many of the ideas carry over.

1. Preliminaries

The definitions and results of this section are standard; see for example [6,7].

We first describe the language L of first order predicate logic with equality, but without negation. Sentences of this language will be the

positive sentences of ordinary first order logic with equality.

Definition 1.0

The language L consists of the following:

Symbols

(i) a countably infinite set of variables $x_0, x_1, \ldots$;

(ii) a countably infinite set of function symbols $f_0^m, f_1^m, \ldots$ for each finite arity $m \geq 0$ (nullary function symbols $f_0^0, f_1^0, \ldots$ will be called constants and denoted $a_0, a_1, \ldots$);

(iii) a countably infinite set of relational symbols $R_0^m, R_1^m, \ldots$ for each finite arity $m \geq 0$;

(iv) an equality symbol $\approx$;

(v) logical symbols $\wedge, \vee, \exists, \forall$.

Terms $t_1, t_2, \ldots$ are defined inductively:

(i) x_j, a_j are terms;

(ii) if $t_1, \ldots, t_m$ are terms then $f_j^m t_1 \ldots t_m$ is.

Formulas ϕ, ψ are defined inductively:

(i) $t_1 \approx t_2$, $R_j^m t_1 \ldots t_m$ are atomic formulas;

(ii) if ϕ, ψ are formulas then

$\phi \wedge \psi$, $\phi \vee \psi$, $\exists x_j \phi$, $\forall x_j \phi$ are.

Sentences are closed formulas, i.e. those with no free occurrences of variables. ∎

Definition 1.1

τ = {closed terms (those not containing occurrences of variables)}

Sym(t) = {symbols appearing in term t}

Sym(ϕ) = {symbols appearing in formula ϕ}

Free(ϕ) = {variables with free (unquantified) occurrences in ϕ}.

We will write $\phi(x_1,\ldots,x_k)$ to indicate that all the free variables of ϕ are among $x_1,\ldots,x_k$, and $\phi(t_1,\ldots,t_k)$ to represent the formula ϕ with all free occurrences of x_i replaced by t_i, $1\le i\le k$.

Definition 1.2

A structure for L is a pair

$$\mathcal{A} = \langle A,I\rangle$$

where A is a set (the domain), and I is a map (the interpretation) taking function symbols f_i^m to functions $A^m\to A$ of the corresponding arity (constants go to elements of A) and relation symbols R_i^m to m-ary relations on A.

We write $f_{i\,\mathcal{A}}^m$ for $I(f_i^m)$ and $R_{i\,\mathcal{A}}^m$ for $I(R_i^m)$. ∎

The interpretation extends naturally to the set of closed terms, by taking

$$(f_i^m t_1\ldots t_m)_{\mathcal{A}} = f_{1\,\mathcal{A}}^m(t_{1_{\mathcal{A}}},\ldots,t_{m_{\mathcal{A}}}).$$

Definition 1.3

A valuation of variables over $\mathcal{A} = \langle A,I\rangle$ is a map v:{variables}→A.

Let $i\ge 0$, $y\in A$. The map $v[i\backslash y]$ is defined by

$$v[i\backslash y](x_j) = v(x_j) \text{ if } i\ne j,$$

$$v[i\backslash y](x_i) = y.$$ ∎

v extends naturally to the set of all terms, by taking

$$v(a_i) = a_{i_{\mathcal{A}}}$$

$$v(f_i^m t_1\ldots t_m) = f_{1\,\mathcal{A}}^m(v(t_1),\ldots,v(t_m)).$$

If t is any term, we denote v(t) by $t_{\mathcal{A},v}$. Note that for $t\in\tau$, $t_{\mathcal{A},v} = t_{\mathcal{A}}$.

Definition 1.4

A formula ϕ is true in $\mathcal{A}$ under valuation v (notation: $\mathcal{A}\models_v\phi$) if either:

(i) ϕ is of the form $s\approx t$, s,t terms, and $s_{\mathcal{A},v} = t_{\mathcal{A},v}$;

(ii) ϕ is of the form $R_i^m t_1 \ldots t_m$ and

$R_{i_A}^m(t_{1_{A,v}}, \ldots, t_{m_{A,v}})$;

(iii) ϕ is of the form $\psi \wedge \chi$ and

$A \models_v \psi$ and $A \models_v \chi$;

(iv) ϕ is of the form $\psi \vee \chi$ and either

$A \models_v \psi$ or $A \models_v \chi$;

(v) ϕ is of the form $\exists x_i \psi$ and for some $y \in A$,

$A \models_{v[i \backslash y]} \psi$;

(vi) ϕ is of the form $\forall x_i \psi$ and for all $y \in A$,

$A \models_{v[i \backslash y]} \chi$. ∎

<u>Theorem 1.5</u>

Let $A = \langle A, I \rangle$, $A' = \langle A, I' \rangle$ be structures and v, v' be valuations such that I and I' agree on $\mathrm{Sym}(\phi)$ and v and v' agree on $\mathrm{Free}(\phi)$. Then

$A \models_v \phi$ iff $A' \models_{v'} \phi$.

<u>Proof</u>

Induction on structure of ϕ. ∎

<u>Corollary 1.6</u>

Let ϕ be closed, v, v' any two valuations. Then $A \models_v \phi$ iff $A \models_{v'} \phi$. ∎

For this reason we may write $A \models \phi$ unambiguously whenever ϕ is closed, and say ϕ is true in A.

<u>Definition 1.7</u>

A sentence ϕ is <u>valid</u> if ϕ is true in all structures.

The <u>validity problem</u> is the set

($\phi|\phi$ is a valid sentence of L). ∎

Theorem 1.8

Let t be any term, and suppose $v(x_i) = t_{A,v}$. Then $A\models_v\phi(x_i)$ iff $A\models_v\phi(t)$, provided no free variables of t become bound as a result of the substitution.

Proof

Induction on structure of ϕ. ∎

Definition 1.9

Let A and B be structures with domains A and B, respectively. A map $h:A\to B$ is a homomorphism $A\to B$ provided for any f_i^m, R_i^m, and $y_1,\ldots,y_m \in A$,

(i) $h(f_{i\,A}^{m}(y_1,\ldots,y_m)) = f_{i\,B}^{m}(h(y_1),\ldots,h(y_m))$, and

(ii) $R_{i\,A}^{m}(y_1,\ldots,y_m) \to R_{i\,B}^{m}(h(y_1),\ldots,h(y_m))$. ∎

If $h:A\to B$ is a homomorphism and v is a valuation over A, then $h\circ v$ is a valuation over B, and for any term t, $h(t_{A,v}) = t_{B,h\circ v}$.

Theorem 1.10

Let B be a homomorphic image of A, let ϕ be any formula, and let v be any valuation over A. Then

$$A\models_v\phi \;\to\; B\models_{h\circ v}\phi.$$

Proof

By assumption there is a surjective homomorphism $h:A\to B$. Proceeding by induction on the structure of ϕ,

$$A\models_v R_i^m t_1\ldots t_m \;\leftrightarrow\; R_{i\,A}^{m}(t_{1_{A,v}},\ldots,t_{m_{A,v}})$$

$$\to\; R_{i\,B}^{m}(h(t_{1_{A,v}}),\ldots,h(t_{m_{A,v}}))$$

$$\leftrightarrow R_{i\,\mathcal{B}}^{m}(t_{1\,\mathcal{B},h\circ v},\ldots,t_{m\,\mathcal{B},h\circ v})$$

$$\leftrightarrow \mathcal{B}\models_{h\circ v}\phi.$$

and

$$\mathcal{A}\models_v s\approx t \leftrightarrow s_{\mathcal{A},v} = t_{\mathcal{A},v}$$

$$\rightarrow h(s_{\mathcal{A},v}) = h(t_{\mathcal{A},v})$$

$$\leftrightarrow s_{\mathcal{B},h\circ v} = t_{\mathcal{B},h\circ v}$$

$$\leftrightarrow \mathcal{B}\models_{h\circ v} s\approx t.$$

The induction step for ϕ of the form $\psi\wedge\chi$ or $\psi\vee\chi$ is trivial. Finally,

$$\mathcal{A}\models_v \forall x_j\psi \leftrightarrow \forall y\in A\ \mathcal{A}\models_{v[j\backslash y]}\psi$$

$$\rightarrow \forall y\in A\ \mathcal{B}\models_{h\circ(v[j\backslash y])}\psi$$

$$\leftrightarrow \forall y\in A\ \mathcal{B}\models_{(h\circ v)[j\backslash h(y)]}\psi$$

and since h is onto,

$$\leftrightarrow \forall y\in B\ \mathcal{B}\models_{(h\circ v)[j\backslash y]}\psi$$

$$\leftrightarrow \mathcal{B}\models_{h\circ v} \forall x_j\psi.$$

The case of $\phi = \exists x_j\psi$ is similar. ∎

<u>Corollary 1.11</u>

If $\mathcal{B}$ is a homomorphic image of $\mathcal{A}$ and ϕ is closed, then

$$\mathcal{A}\models\phi \rightarrow \mathcal{B}\models\phi.$$ ∎

<u>Definition 1.12</u>

The <u>Herbrand</u> (or <u>free</u>) <u>structure</u> is the structure $\mathcal{T}$ with domain τ, the set of closed terms, and interpretation defined by

$$a_{i_T} = a_i$$

$$f_i^m{}_T = \lambda t_1 \ldots t_m[f_i^m t_1 \ldots t_m],\ m \geq 1$$

$$R_i^m{}_T = \lambda t_1 \ldots t_m[\text{false}].$$ ∎

Note that for any $t \in \tau$, $t_T = t$.

2. Main Results

We wish to give a nondeterministic polynomial time algorithm for deciding validity of sentences in L. Our plan will be to reduce the problem of validity of ϕ to truth of ϕ in the Herbrand structure, then use the techniques of [2] to decide truth of ϕ in this structure in nondeterministic polynomial time.

Let ϕ be any sentence of L.

Theorem 2.0

ϕ is valid iff $T \models \phi$.

Proof

($\rightarrow$) By definition of validity.

($\leftarrow$) Suppose ϕ is not valid. Then there is a model of $\neg\phi$. By the Lowenheim-Skolem theorem, there is a countable or finite model of $\neg\phi$, say U. Let U be the domain of U, and let $h:\tau \rightarrow U$ be any map such that

$$h(a_i) = a_{i_U} \quad \text{for } a_i \in \text{Sym}(\phi)$$

and h maps constants not in Sym(ϕ) onto U. This is possible since Sym(ϕ) is finite and U is at most countable. h then extends uniquely to domain τ by taking

$$h(f_i^m t_1 .. t_m) = f_i^m{}_U(h(t_1), \ldots, h(t_m)).$$

Thus if we define a new structure $\mathcal{U}'$ with domain U and interpretation defined by

$$a_{i_{\mathcal{U}'}} = h(a_i)$$

$$f_{i_{\mathcal{U}'}}^{m} = f_{i_{\mathcal{U}}}^{m},\ m\geq 1,$$

$$R_{i_{\mathcal{U}'}}^{m} = R_{i_{\mathcal{U}}}^{m},\ m\geq 0,$$

then $h:\mathcal{T}\to\mathcal{U}'$ is a surjective homomorphism. But since the interpretations of $\mathcal{U}$ and $\mathcal{U}'$ agree on Sym(ϕ) and $\mathcal{U}\not\models\phi$, by Theorem 1.5, $\mathcal{U}'\not\models\phi$. Since $\mathcal{U}'$ is a homomorphic image of $\mathcal{T}$, by Corollary 1.11, $\mathcal{T}\not\models\phi$. ∎

We can also restrict our attention to sentences of a special form.

Lemma 2.1

There is a polynomial time algorithm which, given formula ϕ, produces ϕ' such that

(i) ϕ' is in prenex form,

(ii) all atomic formulas of ϕ' are of the form $s\approx t$ (i.e. ϕ' contains no relational symbols), and

(iii) for any v, $\mathcal{T}\models_v\phi$ iff $\mathcal{T}\models_v\phi'$.

Proof

The standard algorithm for converting any sentence to an equivalent one in prenex form, which can be found in any logic text (e.g. [6]) is polynomial in time and will suffice for our purposes. To dispose of the relational symbols, since every R_i^m is interpreted as universally false in $\mathcal{T}$, atomic formulas of the form $R_i^m t_1 \ldots t_m$ occurring in ϕ may be replaced by the formula $a_0\approx a_1$, which is also false in $\mathcal{T}$. Then (iii) may be verified by induction on the structure of ϕ. ∎

Henceforth all sentences of L we consider will be assumed to be in this form.

We have reduced the validity problem to the problem of truth in T of sentences of a special form. One useful consequence of this, which we will exploit fully, is that the subtle distinction between mention and use can now be conveniently ignored, since the semantic individuals (closed terms) are actually syntactic objects as well. More precisely,

Theorem 2.2

(i) $T \models s \sim t$ iff $s = t$, $s, t \in \tau$;

(ii) $T \models \forall x_i \phi(x_i)$ iff for all $t \in \tau$ $T \models \phi(t)$;

(iii) $T \models \exists x_i \phi(x_i)$ iff there is a $t \in \tau$ $T \models \phi(t)$.

Proof

(i) is a direct consequence of the fact that $s_T = s$ and $t_T = t$; (ii) and (iii) follow from the definition of $\models$ and Theorem 1.8. ∎

Thus we may write

$$Q_1 x_1 \ldots Q_k x_k \ \phi(x_1, \ldots, x_k) \qquad (*)$$

for

$$T \models Q_1 x_1 \ldots Q_k x_k \ \phi(x_1, \ldots, x_k); \qquad (**)$$

Here (**) is an assertion about truth of a sentence of L in T, whereas (*) is a metastatement about elements of τ. In (*), all $\sim$ have been changed to $=$, variables range over τ, and the Q_i are no longer symbols of L, but represent the English "for all" and "there is" in (ii) and (iii) of the previous theorem. Henceforth we shall in general not distinguish between (*) and the right side of the $\models$ in (**).

Now we show how to get rid of leading universal quantifiers.

Theorem 2.3

$$T\models\forall x_i\phi(x_i) \text{ iff } T\models\phi(a_j),$$

where $a_j\notin \mathrm{Sym}(\phi)$.

Proof

Let v be any valuation with $v(x_i) = a_j$.

$$(\rightarrow)\ T\models\forall x_i\phi(x_i) \rightarrow T\models_v\phi(x_i)$$

$$\rightarrow T\models\phi(a_j),$$

by Theorem 1.8.

$(\leftarrow)$ Let $T\models\phi(a_j)$. For arbitrary y, define

$h(a_i) = a_i$ for a_i Sym(ϕ)

$h(a_j) = y$

and let h map $\{a_i | a_i\notin \mathrm{Sym}(\phi) \text{ and } i\neq j\}$ onto τ. Extend h to a homomorphism $T\rightarrow T'$, where T' is just T with some of the a_i's not appearing in ϕ reinterpreted, as in the proof of Theorem 2.0.

Since h is surjective, by Corollary 1.11,

$$T'\models\phi(a_j),$$

thus

$$T'\models_{v[j\backslash y]}\phi(x_j),$$

by Theorem 1.8. By Theorem 1.5,

$$T\models_{v[j\backslash y]}\phi(x_j).$$

As y was arbitrary,

$$T\vdash\forall x_j\phi(x_j).$$ ∎

The above theorem indicates why universal quantifiers are so easy to eliminate in this setting: there are an infinite number of unused constant symbols which are ripe for reinterpretation. In [2] this was not possible, since the number of symbols was finite. The problem studied in [2], namely the truth of sentences of the form

$$Q_1x_1\ldots Q_kx_k\ s(\bar{x})\approx t(\bar{x})$$

in a finitely presented algebra, appears to correspond to the validity problem for sentences in L when a certain kind of bounded quantification is allowed, but the exact correspondence is unclear (see §3).

Let us further restrict our attention to formulas with conjunctive matrices. Let ϕ be in prenex form with no relational symbols besides $\approx$, i.e., ϕ looks like

$$Q_1x_1\ldots Q_kx_k\ \ B(\phi_1(\bar{x}),\ldots,\phi_n(\bar{x}))$$

where B is a monotone Boolean tree with leaves $\phi_1(\bar{x}),\ldots,\phi_n(\bar{x})$, each ϕ_i an atomic formula $s_i\approx t_i$, and $\bar{x} = \langle x_1,\ldots,x_k\rangle$.

<u>Lemma 2.4</u>

$$Q_1x_1\ldots Q_kx_k\ \ B(\phi_1(\bar{x}),\ldots,\phi_n(\bar{x}))$$

iff

there is a subset of the ϕ_i's, WLOG say $\phi_1,\ldots,\phi_m$, such that

(i) $B(\underbrace{\text{true},\ldots,\text{true}}_{m},\ \underbrace{\text{false},\ldots,\text{false}}_{n-m}) = \text{true}$, and

(ii) $Q_1x_1 \ldots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(\bar{x})$. ■

Proof

Induction on the number of quantifiers. The basis is easy. The induction step has two cases:

Case 1 leading existential quantifier.

$$\exists x_1\ Q_2x_2 \ldots Q_kx_k\ B(\phi_1(\bar{x}),\ldots,\phi_n(\bar{x}))$$

iff

for some $x_1 \epsilon \tau$, $Q_2x_2 \ldots Q_kx_k\ B(\phi_1(\bar{x}),\ldots,\phi_n(\bar{x}))$

iff (by induction hypothesis)

for some $x_1 \epsilon \tau$ and some subset $\phi_1,\ldots,\phi_m$ of the ϕ_i's,

(i) $B(\underbrace{true,\ldots,true}_{m}, \underbrace{false,\ldots,false}_{n-m})$, and

(ii) $Q_2x_2 \ldots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(\bar{x})$

iff

for some subset $\phi_1,\ldots,\phi_m$ of the ϕ_i's,

(i) $B(\underbrace{true,\ldots,true}_{m}, \underbrace{false,\ldots,false}_{n-m})$, and

(ii) $\exists x_1 Q_2x_2 \ldots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(\bar{x})$.

Case 2 leading universal quantifier.

$$\forall x_1 Q_2x_2 \ldots Q_kx_k\ B(\phi_1(\bar{x}),\ldots,\phi_n(\bar{x}))$$

iff (by Theorem 2.3)

$$Q_2x_2 \dots Q_kx_k\ B(\phi_1(a_j,x_2,\dots,x_k),\dots,\phi_n(a_j,x_2,\dots,x_k)),$$

where $a_j \notin \mathrm{Sym}(\phi)$,

iff (by induction hypothesis)

for some subset $\phi_1,\dots,\phi_m$ of the ϕ_i's,

(i) $B(\underbrace{true,\dots,true}_{m}, \underbrace{false,\dots,false}_{n-m})$, and

(ii) $Q_2x_2 \dots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(a_j,x_2,\dots,x_m)$

iff

for some subset $\phi_1,\dots,\phi_m$ of the ϕ_i's,

(i) $B(\underbrace{true,\dots,true}_{m}, \underbrace{false,\dots,false}_{n-m})$, and

(ii) $\forall x_1 Q_2x_2 \dots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(\bar{x})$. ∎

Lemma 2.4 is not as trivial as it first may appear; some of the variables are universally quantified, and different valuations of these variables could cause different atomic formulas of the matrix to be true. The object of the lemma is to uniformize the set of atomic formulas which can be true, so that our nondeterministic polynomial time algorithm can initially guess this set of atomic formulas, verify that B is true with those formulas true, and then verify the conjunctive formula

$$Q_1x_1 \dots Q_kx_k \bigwedge_{i=1}^{m} \phi_i(\bar{x}).$$

The following definitions and lemmas are simplified versions of ones appearing in [2], which the reader may consult for a more thorough treatment.

<u>Definition 2.5</u>

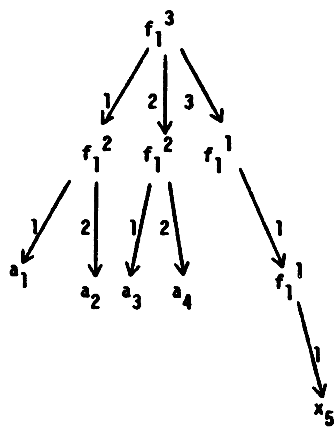
Let $\alpha \epsilon \{f_i^{m}k \mid 1 \le k \le m\}^*$ be a string of symbols on a path through the tree representation of a term. E.g. if s,t are terms,

$$s = f_1^3 f_1^2 a_1 a_2 f_1^2 a_3 a_4 f_1^1 f_1^1 x_5$$

$$t = f_1^1 x_5,$$

then their tree representation are

s =



t =

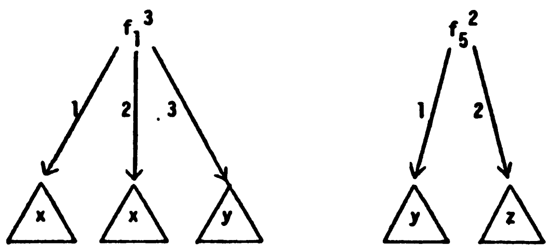
and the path from the root of s to the root of t is

$$\alpha = f_1^3 3 f_1^1 1.$$

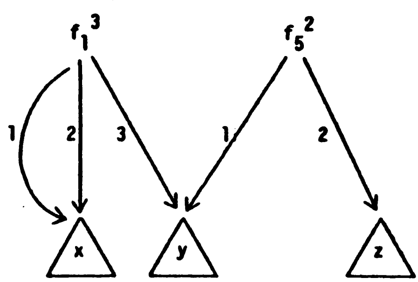
We write $s\alpha t$ to indicate that term t appears as a subterm of term s at the position specified by α.

The empty string is denoted λ; thus $s\lambda t$ iff $s=t$. ∎

As in [2], we will allow terms to be represented by dags instead of trees, by "factoring out" common subterms; e.g.



can be represented more concisely by



The reason for this representation, as opposed to a tree representation, is that sometimes we will want to replace all occurrences of some variable with some term; the dag representation allows us to do this by readjusting edges, so that the representation does not grow any bigger.

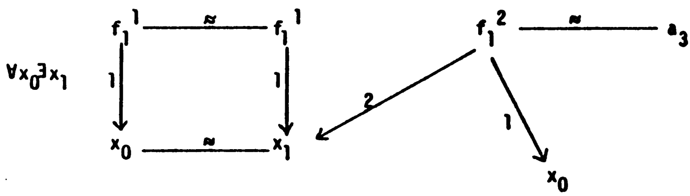
Let the sentence

$$Q_1x_1 \ldots Q_kx_k \bigwedge_{i=1}^{m} s_i \approx t_i$$

be so represented. Extra undirected edges between terms may be used to represent $\approx$. E.g., the sentence

$$\forall x_0 \exists x_1 \; f_1^1 x_0 \approx f_1^1 x_1 \wedge f_1^2 x_0 x_1 \approx a_3 \wedge x_0 \approx x_1$$

could be represented by



In the following, let

$$\phi = Q_1x_1 \ldots Q_kx_k \bigwedge_{i=1}^{n} s_i \approx t_i$$

be given.

Definition 2.6

$\sim$ is the smallest equivalence relation on terms satisfying

(i) $s_i \sim t_i$, $1 \le i \le n$

(ii) if $f^m u_1 \ldots u_m \sim f^m v_1 \ldots v_m$ then

$u_i \sim v_i$, $1 \le i \le m$. ∎

Lemma 2.7

If $\bar{y} \in \tau^k$ is such that $\bigwedge_{i=1}^{n} s_i(\bar{y}) = t_i(\bar{y})$, and if $u \sim v$, then $u(\bar{y}) = v(\bar{y})$.

Proof

Induction on definition of $\sim$. ∎

<u>Definition 2.8</u>

Let x_i, x_j be variables. Define $x_i \underline{\alpha} x_j$ if either

(i) $\exists u\ x_i \sim u\ \&\ u\alpha x_j$, or

(ii) $\exists x_k, \beta, \gamma\ \alpha = \beta\gamma,\ x_i \underline{\beta} x_k,\ \&\ x_k \underline{\gamma} x_j$.

A variable x_i is <u>principal</u> if $x_i \underline{\alpha} x_j$ implies $\alpha = \lambda$. ∎

<u>Lemma 2.9</u>

If $\bar{y} = \langle y_1, \ldots, y_k \rangle \in \tau^k$ such that $\bigwedge_{i=1}^{n} s_i(\bar{y}) = t_i(\bar{y})$, and if $x_i \underline{\alpha} x_j$, then $y_i \alpha y_j$.

<u>Proof</u>

Induction on definition of $x_i \underline{\alpha} x_j$. ∎

<u>Definition 2.10</u>

$R^+ = \{\text{subterms of } s_i \text{ and } t_i,\ 1 \leq i \leq n\}$.

$R = R^+ \cap \tau$. ∎

<u>Lemma 2.11</u>

If x_i is not principal then there is a proper term $u \in R^+$ (a <u>proper term</u> is one that is not a variable or a constant) such that $x_i \sim u$. Moreover, there is a polynomial time algorithm to determine whether x_i is principal, and if not, supply a proper $u \in R^+$ such that $x_i \sim u$.

<u>Proof</u>

The first part follows from the definition of $x_i \underline{\alpha} x_j$. For the second part, construct the relation $\sim$ inductively on the dag representation of $\bigwedge_{i=1}^{n} s_i \approx t_i$. ∎

<u>Lemma 2.12</u>

If x_1 is principal, $y_1 \notin R$, and $z_1 \notin R$, then

$$Q_2x_2\ldots Q_kx_k \bigwedge_{i=1}^{n} s_i(y_1,x_2,\ldots,x_k) = t_i(y_1,x_2,\ldots,x_k)$$

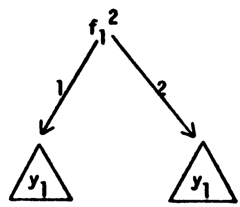
iff

$$Q_2x_2\ldots Q_kx_k \bigwedge_{i=1}^{n} s_i(z_1,x_2,\ldots,x_k) = t_i(z_1,x_2,\ldots,x_k).$$

<u>Proof</u>

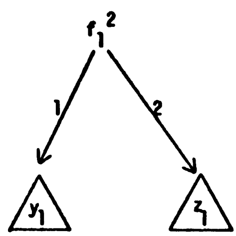
Suppose x_1 is principal, $\bar{y}\epsilon\tau^k$ with $y_1 \notin R$, and $\bigwedge_{i=1}^{n} s_i(\bar{y})=t_i(\bar{y})$. Let $z_1 \notin R$ be arbitrary. For $1\le i\le k$, define $g_i(y_i) = y_i$ with all occurrences of y_1 in y_i at a position α such that $x_i\underline{\alpha}x_1$ replaced by z_1. For example, if

$y_3=$



and $x_3(\underline{\underline{f_1}}^2 2)x_1$ but not $x_3(\underline{\underline{f_1}}^2 1)x_1$ then

$g_3(y_3) =$



Note that $g_1(y_1) = z_1$. Let $<g_1(y_1),\ldots,g_k(y_k)>$ be denoted by $g(\bar{y})$.

We claim that $\bigwedge_{i=1}^{n} s_i(g(\bar{y})) = t_i(g(\bar{y}))$. WLOG, it suffices to show that whenever $s_i(\bar{y})\alpha y_1$ and $s_i(g(\bar{y}))\alpha z_1$ then $t_i(\bar{y})\alpha y_1$ and $t_i(g(\bar{y}))\alpha z_1$, since then all the same occurrences of y_1 in $s_i(\bar{y})$ and $t_i(\bar{y})$ are replaced by z_1. Suppose $s_i(\bar{y})\alpha y_1$ and $s_i(g(\bar{y}))\alpha z_1$. We know $t_i(\bar{y})\alpha y_1$, since $s_i(\bar{y}) = t_i(\bar{y})$. It must be that $\alpha=\beta\gamma$, $s_i\beta x_j$, $x_j\underline{\gamma}x_1$, and $y_j\gamma y_1$, for some β,γ,x_j. If $t_i\alpha w$, w is a proper term, and $w\delta x_\ell$ for some x_ℓ, use the definition of ~ to show that $x_1\underline{\delta}x_\ell$, contradicting the principality of x_1. If $t_i\alpha w$ and $w\in R$, then $y_1\in R$, contradicting an assumption. The only possibility remaining is that $t_i\delta x_\ell$ and $\delta\eta=\alpha$, for some δ,η,x_ℓ. A case argument of two cases (one in which δ is a substring of β, the other in which β is a substring of δ) shows that $x_\ell\underline{\eta}x_1$, thus $y_\ell\underline{\eta}y_1$ by Lemma 2.9. Then $g_\ell(y_\ell)\eta z_1$ and $t_i(g(\bar{y}))\alpha z_1$, and the claim is verified.

Proceeding by induction on quantifiers, suppose for any $y_2,\ldots,y_\ell\in\tau$,

$$Q_{\ell+1}y_{\ell+1}\cdots Q_k y\phi(y_1,\ldots,y_\ell,y_{\ell+1},\ldots,y_k)$$

$$\rightarrow$$

$$Q_{\ell+1}y_{\ell+1}\cdots Q_k y_k\phi(g_1(y_1),\ldots,g_\ell(y_\ell),y_{\ell+1},\ldots,y_k),$$

where $\phi = \bigwedge_{i=1}^{n} s_i\approx t_i$. Then

$$\exists y_\ell Q_{\ell+1}y_{\ell+1}\cdots Q_k y_k\phi(y_1,\ldots,y_\ell,y_{\ell+1},\ldots,y_k) \qquad (*)$$

$$\rightarrow$$

$$\exists y_\ell Q_{\ell+1}y_{\ell+1}\cdots Q_k y_k\phi(g_1(y_1),\ldots,g_{\ell-1}(y_{\ell-1}),y_\ell,\ldots,y_k); \qquad (**)$$

the y_ℓ satisfying (**) is obtained by applying g_ℓ to the y_ℓ satisfying (*). If

$$\forall y_\ell Q_{\ell+1}y_{\ell+1}\cdots Q_k y_k\phi(y_1,\ldots,y_{\ell-1},y_\ell,\ldots,y_k)$$

then it cannot be the case that $x_\ell\underline{\alpha}x_1$ for any α, by Lemma 2.9. Thus

$g_\ell(y_\ell) = y_\ell$ for all $y_\ell \in \tau$. Then

$$\forall y_\ell Q_{\ell+1} y_{\ell+1} \cdots Q_k y_k \phi(y_1, \ldots, y_\ell, y_{\ell+1}, \ldots, y_k)$$

$$\leftrightarrow$$

$$\forall y_\ell Q_{\ell+1} y_{\ell+1} \cdots Q_k y_k \phi(g_1(y_1), \ldots, g_\ell(y_\ell), y_{\ell+1}, \ldots, y_k)$$

$$\leftrightarrow$$

$$\forall y_\ell Q_{\ell+1} y_{\ell+1} \cdots Q_k y_k \phi(g_1(y_1), \ldots, g_{\ell-1}(y_{\ell-1}), y_\ell, \ldots, y_k).$$

We have shown

$$Q_2 x_2 \cdots Q_k x_k \bigwedge_{i=1}^{n} s_i(y_1, x_2, \ldots, x_k) = t_i(y_1, x_2, \ldots, x_k)$$

$$\leftrightarrow$$

$$Q_2 x_2 \cdots Q_k x_k \bigwedge_{i=1}^{n} s_i(z_1, x_2, \ldots, x_k) = t_i(z_1, x_2, \ldots, x_k),$$

and the converse follows from symmetry. ∎

Now we are ready to show how to eliminate leading existential quantifiers.

<u>Theorem 2.13</u>

There is a nondeterministic polynomial time algorithm which, given

$$\exists x_1 Q_2 x_2 \cdots Q_k x_k \bigwedge_{i=1}^{n} s_i(\bar{x}) = t_i(\bar{x}), \qquad (*)$$

produces a true formula of the same size as (*) but with one fewer quantifier iff (*) is true.

<u>Proof</u>

Given (*), if x_1 is principal, guess whether some $y_1 \in R$ will satisfy (*) when substituted for x_1. If guessed yes, guess which one, replace all occurrences of x_1 in s_i and t_i with y_1 (this is done by redirecting all edges into occurrences of x_1 to the root of y_1, thus the size of the representation does not increase) and output

$$Q_2x_2 \cdots Q_kx_k \bigwedge_{i=1}^{n} s_i(y_1,x_2,\ldots,x_k) = t_i(y_1,x_2,\ldots,x_k).$$

If guessed no, output

$$Q_2x_2 \cdots Q_kx_k \bigwedge_{i=1}^{n} s_i(a_j,x_2,\ldots,x_k) = t_i(a_j,x_2,\ldots,x_k)$$

for some $a_j \notin \mathrm{Sym}(\phi)$. This suffices, by Lemma 2.12. If x_1 is not principal, Lemma 2.11 guarantees us a proper term $u \in R^+$ such that $x_1 \sim u$. Thus (*) is equivalent to

$$\exists x_1 Q_2x_2 \cdots Q_kx_k \bigwedge_{i=1}^{n} s_i(\bar{x})=t_i(\bar{x}) \wedge x_1 = u(\bar{x}),$$

by Lemma 2.7. If x_1 appears in u, the sentence is false, and we may immediately reject. Otherwise replace all occurrences of x_1 in s_i and t_i with u (redirect edges incident to x_1 to the root of u) and call the resulting terms s_i', t_i'. Now we have the equivalent formula

$$\exists x_1 Q_2x_2 \cdots Q_kx_k \bigwedge_{i=1}^{n} s_i'(\bar{x}) = t_i'(\bar{x}) \wedge x_1 = u(\bar{x}). \qquad (**)$$

Let $x_{j_1},\ldots,x_{j_r}$ be the variables appearing in u, and suppose y_1 satisfies (**). If any of the x_{j_i} are universally quantified, the sentence is immediately false. Otherwise, if $u\alpha x_{j_i}$, the only value of x_{j_i} which can satisfy (**) is the subterm of y_1 occurring at position α. As this is uniform in the universally quantified variables occurring before $\exists x_{j_i}$ in the quantifier list, the $\exists x_{j_i}$ may be moved to the front. Thus (**) implies

$$\exists x_1 \exists x_{j_1} \cdots \exists x_{j_r} Q_{r+1}'x_{r+1}' \cdots Q_k'x_k' \bigwedge_{i=1}^{n} s_i'(\bar{x}) = t_i'(\bar{x}) \wedge x_1 = u(\bar{x}), \qquad (***)$$

where $Q'_{r+1}x'_{r+1}\ldots Q'_k x'_k$ is the quantifier list $Q_2x_2\ldots Q_kx_k$ with all the $\exists x_{j_i}$ removed. Clearly the implication goes the other way as well, since for any $\phi(x_1,x_2)$,

$$\exists x_1 \forall x_2\ \phi(x_1,x_2) \rightarrow \forall x_2 \exists x_1\ \phi(x_1,x_2).$$

But (***) is equivalent to

$$\exists x_{j_1}\ldots\exists x_{j_r}\ [\exists x_1\ x_1 = u(\bar{x})]\ \wedge$$
$$Q'_{r+1}x'_{r+1}\ldots Q'_k x'_k \bigwedge_{i=1}^{n} s_i(\bar{x}) = t_i(\bar{x})\ , \qquad (+)$$

since x_1 does not occur in any s_i' or t_i'. But (+) is trivially equivalent to

$$\exists x_{j_1}\ldots\exists x_{j_r} Q'_{r+1}x'_{r+1}\ldots Q'_k x'_k \bigwedge_{i=1}^{n} s_i'(\bar{x}) = t_i(\bar{x}),$$

which is of the desired form. As all manipulations took polynomial time, we are done. ∎

Theorem 2.14

The validity problem is in NP.

Proof

Given ϕ, we need only check that $\vdash_T \phi$, by Theorem 2.0. We can eliminate relational symbols and convert to prenex form in polynomial time, by Lemma 2.1. By Lemma 2.4, we can convert to a formula with a conjunctive matrix in nondeterministic polynomial time. Theorems 2.3 and 2.13 allow us to eliminate quantifiers in nondeterministic polynomial time. Finally, we are left with a sentence of the form

$$\bigwedge_{i=1}^{n} s_i = t_i$$

where $s_i,t_i\in\tau$, which can certainly be verified in polynomial time. ∎

<u>Theorem 2.15</u>

The validity problem is $\le_{\log}$-complete for NP.

<u>Proof</u>

We reduce a well-known NP-complete problem, the satisfiability of Boolean formulas, to the validity problem.

Let B be a Boolean formula with variables $x_1,\ldots,x_k$. Let B′ be formed from B by replacing each literal x_i by $x_i\approx a_1$ and each literal $\neg x_i$ by $x_i\approx a_0$. E.g. if

$$B = (x_1 \vee \neg x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee \neg x_3)$$

then

$$B' = (x_1\approx a_1 \vee x_2\approx a_0 \vee x_3\approx a_1) \wedge (x_1\approx a_1 \vee x_2\approx a_1 \vee x_3\approx a_0).$$

If B is satisfiable over {true,false} then B′ is satisfiable over $\{a_0,a_1\}$ in the obvious way. If B′ is satisfiable over τ then B′ is satisfiable over $\{a_0,a_1\}$, by reassigning any variable in B′ not assigned to a_0 or a_1 to either a_0 or a_1. The monotonicity of B′ guarantees that the new assignment also satisfies B′. From this we get a satisfying assignment for B in the obvious way. Thus

B is satisfiable over {true,false}

iff

B′ is satisfiable over τ

iff (by Theorem 2.2)

$\mathcal{T} \models \exists x_1 \ldots \exists x_k\, B'$

iff (by Theorem 2.0)

$\exists x_1 \ldots \exists x_k\ B'$ is a valid sentence of L. ∎

Theorem 2.15 is a special case of a more general result:

Theorem 2.16

Let Σ be a finite set of sentences of the form $s\sim t$, $s,t\in\tau$, and let ϕ be a sentence of L. The problem,

"Is ϕ true in all models of Σ?"

is $\leq_{\log}$-complete for NP. ∎

A more extensive use of the techniques of [2] is needed, but all the main ideas are here. The Herbrand domain for the more general case is the quotient structure T/Σ, whose domain is the set of closed terms τ modulo the congruence relation induced by Σ.

It is conjectured that Theorem 2.16 holds even if Σ is allowed to contain atomic formulas $R_i^m t_1 \ldots t_m$, $t_i \in \tau$, $1\leq i\leq m$.

3. Problems

(i) Prove the conjecture at the end of §2. What other generalizations can be made?

(ii) Let Σ be given. Suppose that in addition to $\forall,\exists$ we allow bounded quantifiers of the form

$$\forall_{t_1,\ldots,t_n,f_1^{m_1},\ldots,f_k^{m_k}} \text{ and } \exists_{t_1,\ldots,t_n,f_1^{m_1},\ldots,f_k^{m_k}}.$$

The meaning of $\forall_{t_1,\ldots,t_n,f_1^{m_1},\ldots,f_k^{m_k}} x$ would be, "for all elements x of the substructure of A generated by $t_{1_A},\ldots,t_{n_A}$ under $f_{1_A}^{m_1},\ldots,f_{k_A}^{m_k}$, ..." We apparently now have enough power to force variables to range only over the algebra presented by Σ, instead of all of A, thus the validity problem is at least PSPACE-hard (see [2]). Use the fact that deciding membership in a finitely generated substructure is complete for P [2] to show that this is all the power you get; i.e., show that with bounded quantifiers, the validity problem for sentences with n alternations of quantifiers, the outermost a $\exists(\forall)$, is complete for $\Sigma_n^P(\Pi_n^P)$, and the validity problem in general is complete for PSPACE.

[1] Kozen, D., "Complexity of Finitely Presented Algebras," Proc. 9th ACM Symposium on Theory of Computing, May 1977, pp. 164-177.

[2] Kozen, D., "Finitely Presented Algebras and the Polynomial Time Hierarchy," Cornell Univ. Dept. of Computer Science TR77-303, March 1977.

[3] Pratt, V.R., "The Power of Negative Thinking in Multiplying Boolean Matrices," Proc. 6th ACM Symposium on Theory of Computing, April 1974, pp. 80-83.

[4] Stockmeyer, L.J., and A.R. Meyer, "Word Problems Requiring Exponential Time: Preliminary Report," Proc. 5th ACM Symposium on Theory of Computing, April 1973, pp. 1-9.

[5] Manna, Z., Mathematical Theory of Computation (New York: McGraw Hill, 1974), p. 105ff.

[6] Enderton, H.B., A Mathematical Introduction to Logic (New York: Academic Press, 1972).

[7] Shoenfield, J.R., Mathematical Logic (Reading, MA: Addison-Wesley, 1967).