Artificial Intelligence Project--RLE and MIT Computation Center
Memo 40 -- A Note on the Feasibility of Application of the Davis Putnam Proof Procedure to Elementary Number Theory

Donald Dawson

1.0 Introduction

In ref.1 Davis and Putnam present a computational proof procedure for quantification theory which they suggest might be applied to obtain proofs in mathematical domains. In ref.2 they give a finite axiom system for elementary number theory with the aim of applying the computational proof procedure to it. In ref. 3 Wang points out that as it stands this procedure would be far too inefficient to prove non trivial theorems and discusses how it might be made more efficient. In this note we will indicate that even the type of modification that Wang considered would not be sufficient to enable the system to prove non trivial theorems.

1.1 The Computational Proof Procedure

Since a complete presentation is given in ref.1 only a sketch will be given here. A w.f.f. R is proved by showing that $\sim R$ is inconsistent. $\sim R$ is first put into prenex normal form with the quantifier free part in conjunctive normal form. Then existential quantifiers are replaced by functional symbols. Then by substituting various constants for the variables in the canonical form of $\sim R$ a sequence of lines is generated. For example, the lines generated from $(x_1).(x_3)\ R(x_1,f(x_1),x_3)$ would be

$R(a,fa,a)$
$R(a,fa,fa)$
$R(fa,ffa,a)$
$R(fa,ffa,fa)$ etc.

Then a scheme is given which determines whether a given finite set of lines is inconsistent or not. Lines are generated until a set is found which is inconsistent in which case the theorem is proved

1.2 The Axiom System for Number Theory.

The non logical constants are as follows:

Individual symbols: $1, N, \sigma, \pi, I$

Binary predicate symbols: $=, \epsilon$

Singularly function symbols: $-, i, P, \surd, \S, \lambda$

Binary function symbols: $+, \cdot, J, \cap, X.$

Lower case Roman letters are used as individual variables.

Group A

1. $x = y \leftrightarrow (z)(x \epsilon z \leftrightarrow y \epsilon z)$

2. $x \epsilon y \rightarrow (x \epsilon N \wedge y \epsilon N)$

Group B

1. $1 \varepsilon N$
2. $x \varepsilon N \to (x+1) \varepsilon N$
3. $x+1 = y+1 \to x = y$
4. $1 \neq x+1$
5. $(1 \in z + (x)(x \varepsilon z \to (x+1) \varepsilon z)) \to (x)(x \varepsilon N \to x \varepsilon z)$

Group C

1. $J(x,y) = J(x',y') \to (x = x') + (y = y')$
2. $J(x,y) \varepsilon N \leftrightarrow x \varepsilon N + y \varepsilon N$
3. $x + (y+1) = (x+y) + 1$
4. $x \cdot 1 = x$
5. $x \cdot (y+1) = x \cdot y + x$

Group D

1. $x \varepsilon i y \leftrightarrow x = y + x \varepsilon N + y \varepsilon N$
2. $x \varepsilon (y \cap z) \leftrightarrow (x \varepsilon y) \cap (x \varepsilon z)$
3. $x \varepsilon -y \leftrightarrow (x \varepsilon N) + (x \not\varepsilon y)$
4. $J(x, J(y,z)) \varepsilon \sigma \leftrightarrow x = y+z + y \varepsilon N + z \varepsilon N$
5. $J(x, J(y,z)) \varepsilon \pi \leftrightarrow x = y \cdot z + y \varepsilon N + z \varepsilon N$
6. $J(x,y) \varepsilon I \leftrightarrow x = y + x \varepsilon N + y \varepsilon N$
7. $x \varepsilon P(y) \leftrightarrow (E z)(J(z,x) \varepsilon y)$
8. $J(x,y) \varepsilon \gamma z \leftrightarrow J(y,x) \varepsilon z$
9. $J(x, J(y,z)) \varepsilon \delta z' \leftrightarrow J(y, J(z,x)) \varepsilon z'$
10. $J(x, J(y,z)) \varepsilon \lambda z' \leftrightarrow J(x, J(z,y)) \varepsilon z'$

11. $J(u,v) \varepsilon (x \, X \, y) \leftrightarrow (u \varepsilon x) \& (v \varepsilon y)$

2.0 Direct Application of the Computational Procedure.

The canonical form obtained when proving $x \cdot y = y \cdot x$ is as follows:

$\forall x_1 \forall x_2 \forall x_4 \forall x_7 \Big([\sim x_1 \varepsilon f_1(x_1,x_2) \vee x_2 \varepsilon f_1(x_1,x_2) \vee x_1 = x_2$
$\& \sim x_1 = x_2 \vee (\sim x_1 \varepsilon x_4) \vee x_2 \varepsilon x_4 \& \sim x_1 = x_2 \vee (\sim x_2 \varepsilon x_4) \vee$
$(x_1 \varepsilon x_4) \& \sim x_1 \varepsilon x_2 \vee x_1 \varepsilon N \& \sim x_1 \varepsilon x_2 \vee x_2 \not\varepsilon N \& 1 \varepsilon N \&$
$\sim x_1 \varepsilon N \vee (x_1+1) \varepsilon N \& \sim x_1 + 1 = x_2 + 1 \vee x_1 = x_2 \& 1 \neq x_1 + 1 \& \sim 1 \varepsilon x_4 \vee$
$\sim f_2(x_1, x_2, x_4) \varepsilon x_4 \vee ((f_2(x_1, x_2, x_4) + 1) \varepsilon x_4 \vee \sim x_1 \varepsilon N \vee$
$x_1 \varepsilon x_4 \& J(x_1, x_2) \varepsilon N \vee x_1 \varepsilon N \& \sim J(x_1, x_2) \varepsilon N \vee x_2 \varepsilon N$
$\& \sim x_1 \varepsilon N \vee \sim x_2 \varepsilon N \vee J(x_1, x_2) \varepsilon N \& x_1 + (x_2+1) = (x_1+x_2)+1$
$\& x_1 \cdot 1 = x_1 \& x_1 \cdot (x_2+1) = (x_1 \cdot x_2 + x) \& \sim x_1 \varepsilon i x_2 \vee x_1 = x_2$
$\& \sim x_1 \varepsilon i x_2 \vee x_1 \varepsilon N \& \sim x_1 \varepsilon i x_2 \vee x_2 \varepsilon N \& \sim x_1 = x_2 \vee \sim x_1 \varepsilon N \vee \sim x_2 \varepsilon N$
$\vee x_1 \varepsilon i x_2 \& \sim x_1 \varepsilon (x_2 \cap x_4) \vee (x_1 \varepsilon x_2) \& (x_1 \varepsilon x_4) \vee \sim x_1 \varepsilon (x_2 \cap x_4)$
$\& \sim (x_1 \varepsilon x_2) \vee \sim (x_1 \varepsilon x_4) \vee (x_1 \varepsilon (x_2 \cap x_4)) \& \sim x_1 \varepsilon -x_2 \vee x_1 \varepsilon N$
$\& \sim x_1 \varepsilon -x_2 \vee x_1 \varepsilon x_2 \& \sim x_1 \varepsilon N \vee \sim x_1 \not\varepsilon x_2 \vee x_1 \varepsilon -x_2 \&$
$\sim J(x_1, J(x_2, x_4)) \varepsilon \sigma \vee x_2 + x_4 = x_1 \& \sim J(x_1, J(x_2, x_4)) \varepsilon \sigma$
$\vee \sim x_2 \varepsilon N \& \sim J(x_1, J(x_2, x_4)) \varepsilon \sigma \vee x_4 \varepsilon N \& \sim x_1 = x_2 + x_4 \vee$
$\sim x_2 \varepsilon N \vee \sim x_1 \varepsilon N \vee J(x_1, J(x_1, x_4)) \varepsilon \sigma \& \sim J(x_1, J(x_2, x_4)) \varepsilon \pi \vee$
$x_1 = x_2 \cdot x_4 \& \sim J(x_1, J(x_2, x_4)) \varepsilon \pi \vee x_2 \varepsilon N \& \sim J(x_1, J(x_2, x_4))$
$\varepsilon \pi \vee x_4 \varepsilon N \& \sim x_1 = x_2 \cdot x_4 \vee \sim x_2 \varepsilon N \vee \sim z \varepsilon N \vee J(x_1, J(x_2, x_4)) \varepsilon \pi$
$\& \sim J(x_1, x_2) \varepsilon I \vee x_1 = x_2 \& \sim J(x_1, x_2) \varepsilon I \vee x_1 \varepsilon N \&$
$\sim J(x_1, x_2) \varepsilon I \vee x_2 \varepsilon N \& \sim x_1 = x_2 \vee \sim x_1 \varepsilon N \vee \sim x_2 \varepsilon N$
$\vee J(x_1, x_2) \varepsilon I \& (\sim x_1 \varepsilon P(x_2) \vee J(f_3(x_1, x_2, x_4), x_1) \varepsilon x_2)$

$\& (\sim J(x_4, x_1) \varepsilon x_2 \vee x_1 \varepsilon P(x_2)) \& \sim J(x_1, x_2) \varepsilon \gamma x_4$
$\vee J(x_2, x_1) \varepsilon x_4 \& \sim J(x_2, x_1) \varepsilon x_4 \vee J(x_1, x_2) \varepsilon \gamma x_4$
$\& \sim J(x_1, J(x_2, x_4)) \varepsilon \delta x_7 \vee J(x_2, J(x_4, x_1)) \varepsilon x_7 \&$
$\sim J(x_2, J(x_4, x_1)) \varepsilon x_7 \vee J(x_1, J(x_2, x_4)) \varepsilon \delta x_7 \&$
$\sim J(x_1, J(x_2, x_4)) \varepsilon \delta x_7 \vee J(x_1, J(x_2, x_4)) \varepsilon \delta x_7 \vee$
$\sim J(x_2, J(x_4, x_1)) \varepsilon x_7 \vee J(x_1, J(x_2, x_4)) \varepsilon \delta x_7 \&$
$\sim J(x_1, J(x_2, x_4)) \varepsilon \lambda x_7 \vee J(x_1, J(x_4, x_2)) \varepsilon x_7$
$\& \sim J(x_1, J(x_4, x_2)) \varepsilon x_7 \vee J(x_1, J(x_2, x_4)) \varepsilon \lambda x_7 \&$
$\sim J(x_4, x_7) \varepsilon (x_1 X x_2) \vee x_4 \varepsilon x_1 \& \sim J(x_4, x_7) \varepsilon (x_1 X x_2)$
$\vee x_7 \varepsilon x_2 \& \sim x_4 \varepsilon x_1 \vee x_7 \varepsilon x_2 \vee J(x_4, x_7) \varepsilon (x_1 X x_2)]]$
$\& [\sim f_4(x_1, x_2, x_4, x_7) \varepsilon N \vee \sim f_5(x_1, x_2, x_4, x_7) \varepsilon N \vee \sim f_4(x_1, x_2, x_4, x_7) \cdot f_5(x_1, x_2, x_4, x_7)$
$= f_5(x_1, x_2, x_4, x_7) \cdot f_4(x_1, x_2, x_4, x_7)]]$.

Now the above system has:
5 constants (C)
6 two variable functions (B)
6 one variable functions (U)
2 three variable functions (T)
2 four variable functions (Q).

Considering the possible constants generated by these we have:
30 of type U(C)
180 od type U(U(C))
$6 \cdot 5 \cdot 5 = 150$ of type B(C, C)
$6 \cdot 180 \cdot 180 = 194,400$ of type B(U(C), U(C)), etc.

Moreover there are four variables to be substituted for so that using only those of type $U(C)$ there are 30^4. possible choices; those of type $U(U(C))$ there are $(180)^4$ possible choices; etc. We will later show that to prove $(y)\ (1+y=y+1)$ we require the set $P^2((\sigma \cap (Nx(i1xN)) \cap \lambda 6)$ which is of type
$U(U(B(B(C,B(C,B(C,C))),UCC))$. At this level there are of the order of 10^{40} choices which is of the same order as the number of move choices in checkers.

Now it has been suggested that by using some heuristic to generate only those terms which satisfy some criterion of usefulness such as using the functions appearing in the statement of the theorem or using Wang's method of sequential tables that the choices could be reduced to a manageable number. However for any simple minded scheme it is not difficult to show that the kind of exponential growth illustrated above is encountered. In the next section we will indicate a stronger result. In particular, we will consider the number of terms of the Davis Putnam system that occur when an informal proof is formalized in a minimal way in the system. Such a formalization of an informal proof is certainly more efficient than the above system of generating terms even with a very clever heuristic selection scheme.

3.0 Formalization of an Informal Proof

Let us first examine the appearance of a simple proof in the finite axiom scheme. Consider for example the proof of

$$x \in N \Rightarrow R(x).$$

Step1. Show that there exists a set S such that

$x \in S$ iff $R(x)$.

e.g. If $R(x)$ is $x+1=1+x$ then $S=P^2((\sigma\cap(N\times(11\times N)))\cap\lambda\sigma)$.

Step 2. Show that $1 \in S$, that is $R(1)$.

Step 3. Show that $x \in S \Rightarrow x+1 \in S$; that is, show that $R(x)$ implies $R(x+1)$.

Step 4. Then by the induction axiom B5, $S=N$, that is, if $x \in N$ then $x \in S$; that is, if $x \in N$ then $R(x)$.

Note that the induction step is essential unless the theorem can be obtained by direct substitution in one or more axioms, that is, a counterexample to the negation can be obtained by substitution into the axioms.

Example. Some of the simple properties of the predicates can be obtained by direct substitution in the axioms.

Proof of $x=y + z=y \Rightarrow x=z$.

$\sim f_4(1,1,1,1) = f_5(1,1,1,1) \vee \sim f_4(1,1,1,1) \varepsilon f_1(f_4(1,1,1,1),$
$f_6(1,1,1,1)) \vee f_5(1,1,1,1) \varepsilon f_1(f_4(1,1,1,1), f_6(1,1,1,1)) + \sim$
$(f_6(1,1,1,1) = f_5(1,1,1,1)) \vee \sim f_6(1,1,1,1) \varepsilon f_1(f_4(1,1,1,1), f_6(1,1,1,1))$
$\vee \sim f_5(1,1,1,1) \varepsilon f_1(f_4(1,1,1,1), f_6(1,1,1,1)) + \sim f_4(1,1,1,1)$
$\varepsilon f_1(f_4(1,1,1,1), f_6(1,1,1,1)) \vee f_6(1,1,1,1) \varepsilon f_1(f_4(1,1,1,1), f_6(1,1,1,1))$
$\vee (f_4(1,1,1,1) = f_6(1,1,1,1)) + f_4(1,1,1,1) \neq f_5(1,1,1,1)$
$+ f_6(1,1,1,1) \neq f_5(1,1,1,1) + \sim f_4(1,1,1,1) = f_6(1,1,1,1))$

which is inconsistent.

Note that whenever we need to apply this theorem to 3 constants a, b, c in the proof of another theorem we must write the above out in full, that is, we cannot use this result as a general theorem.

Herein lies one of the sources of inefficiency of the Davis and Putnam system as it now stands. The possibility of adding such general results to the axiom list will be discussed below.

Example: Outline of Proof of $x+1 = 1+x$.

Step 1. $x \varepsilon P^2((\sigma \cap (N \times (i/ \times N))) \cap \lambda\sigma) \Rightarrow x+1 = 1+x.$

$x \varepsilon P^2((\sigma \cap (N \times (i/ \times N))) \cap \lambda\sigma) \Rightarrow \exists z \ni.$

$J(z,x) \varepsilon P((\sigma \cap N \times (i/ \times N)) \cap \lambda\sigma) \Rightarrow \exists w \ni.$

$J(w, J(z,x)) \varepsilon ((\sigma \cap N \times (i/ \times N)) \cap \lambda\sigma) \Rightarrow$

$J(w, J(z,x)) \varepsilon N \times (i/ \times N) \Rightarrow z = 1.$

Also $J(w, J(1,x)) \varepsilon \sigma \Rightarrow w = 1+x$

$J(w, J(1,x)) \varepsilon \lambda\sigma \Rightarrow J(w, J(x,1)) \varepsilon \sigma$

$\Rightarrow w = x+1$

Using the theorem which was proved in the previous example we have $x+1 = 1+x$.

Step 2

$$x+1 = 1+x \Rightarrow x \varepsilon P^2((\sigma \cap (N \times (i/\times N)) \cap \lambda\sigma)$$

$$1+x = 1+x \Rightarrow J(x+1, J(1,x)) \varepsilon \sigma$$

$$x+1 = 1+x \Rightarrow J(x+1, J(x,1)) \varepsilon \sigma$$

$$\Rightarrow J(x+1, J(1,x)) \varepsilon \lambda\sigma.$$

Also $J((x+1), J(1,x)) \varepsilon N \times (i/\times N)$.

Hence, $J(x+1, J(1,x)) \varepsilon ((\sigma \cap (N \times (i/\times N)) \cap \lambda\sigma)$

Hence $x \varepsilon P^2((\sigma \cap (N \times (i/\times N)) \cap \lambda\sigma)$

Step 3. $1 \varepsilon S$

$$1+1 \varepsilon z \Leftrightarrow 1+1 \varepsilon z \Rightarrow 1+1 = 1+1$$

Step 4 $x \varepsilon S \Rightarrow x+1 \varepsilon S$

$$1+(x+1) = (1+x)+1$$

$$1+x = x+1 \Rightarrow 1+(x+1) = (x+1)+1$$

by a property of equality which we have not proven, viz, $a=b \Rightarrow a+1 = b+1$ and by the theorem proved in the previous example.

The above outline proof could easily be translated into the Davis and Putnam system giving an inconsistent set of terms to serve as a counterexample to the negation of the theorem. A ~~R~~ rough estimate of a lower limit to the number of terms in the formalization is 100 and this is probably about as low as will ever be encountered in a proof involving one induction step.

However most theorems are not this simple. In particular the steps $R(x)\Rightarrow R(x+1)$ often cannot be obtained by substitution in the axioms. For example, the step $R(x)\Rightarrow R(x+1)$ of many theorems will require the result $x+1=1+x$. In general we will have a chain of theorems:

$$\begin{tikzcd}[column sep=small]
& R(x) \arrow[dl, dash] \arrow[dr, dash] & & \\
R(1) & & R(x)\Rightarrow R(x+1)\equiv Q'(x) \arrow[dl, dash] \arrow[d, dash] & \\
& Q'(1) & Q'(x)\Rightarrow Q'(x+1)\equiv Q^2(x) & \\
& & \vdots & \\
& Q^n(1) & Q^n(x)\Rightarrow Q^n(x+1) &
\end{tikzcd}$$

where $Q^n(x)\Rightarrow Q^n(x+1)$ can be proved by substitution in the axioms.

For difficult theorem this may not lead to a $Q_n(x)$ whose base steps can be proved by substitution in the axioms. In this case it may be necessary to partition the theorem into parts. For example, to prove that $g(x)=h(x)$ it may be necessary to prove $g(x)=f_1(x)=f_2(x)\ldots=f_n(x)=h(x)$.

<u>**Example**</u> Consider the informal proof of

$$1^2+2^2+3^2+\ldots+r^2=r(r+1)\frac{(2r+1)}{6}$$

To prove $R(r)\Rightarrow R(r+1)$ we prove the chain of equalities

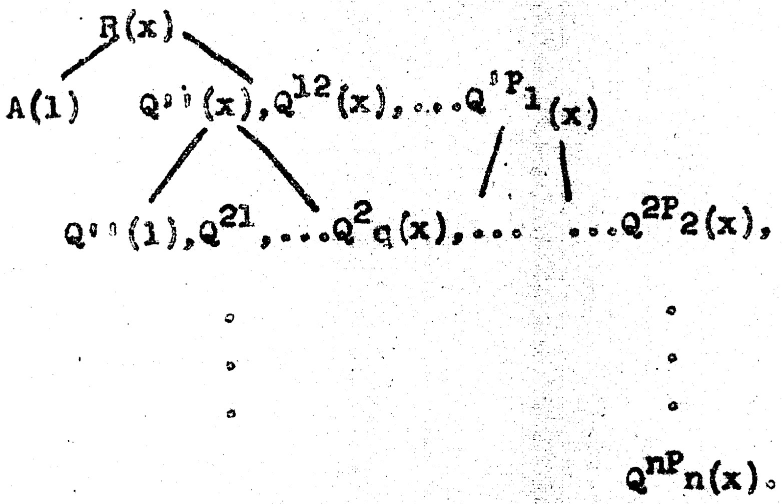
$$1^2+2^2+\ldots+r^2+(r+1)^2 = \frac{r(r+1)(2r+1)}{6}+(r+1)^2$$

$$= \frac{r(r+1)(2r+1)+6(r+1)^2}{6} = \frac{(r+1)[r(2r+1)+6(r+1)]}{6}$$

$$= \frac{(r+1)(2r^2+7r+6)}{6} = \frac{(r+1)(r+2)(2r+3)}{6}$$

$$= \frac{(r+1)((r+1)+1)(2(r+1)+1)}{6}$$

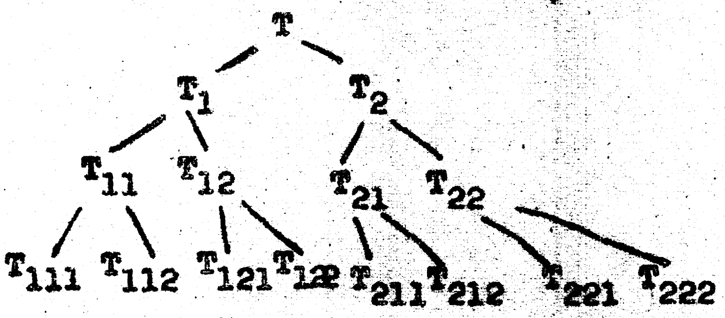
In the presence of partitioning the tree structure is of the form:--


$R(x)$
$A(1)$
$Q^{11}(x), Q^{12}(x), \ldots Q^{1P_1}(x)$
$Q^{11}(1), Q^{21}, \ldots Q^{2}q(x), \ldots$
$\ldots Q^{2P_2}(x),$
$Q^{nP_n}(x).$

We can consider the theorems $Q^{ij}(x)$ as theorems which logically preceed $R(x)$. Another of the sources of inefficiency of this system is that a complete proof of all the theorems which "logically preceed $R(x)$ must be included in a proof of $R(x)$. To see how inefficient this is consider a sequence of theorems $T_1, T_2, \ldots$ in which T_n is the only logical predecessor" of T_{n+1}. If $R_n \equiv T_n \supset T_{n+1}$ takes K lines, then T_n would require K lines in a logic theory type system if T_n were a previously proved theorem, but in the Davis Putnam system would require nK lines. The number of lines required for the proofs of the first n theorems would then nK for a logic theory type system and $\frac{n(n+1)K}{2}$ for a Davis Putnam system.

N.B. What is meant by a logic theory type system is the Davis Putnam system modified so that when a theorem is proved it is added to the list of axioms. Of course then the amount of time spent scanning the axioms becomes a serious problem as in the original studies of the logic theory machine.

Next consider the case in which we wish to prove T when T has two "immediate logical predecessors" which in turn each have two "immediate logical predecessors" and so on. Immediate logical predecessors are those arising in the first partition level,i.e. $Q''(x),\ldots,Q^{1P_1}(x)$; these can be considered to be the main lemmas used in the proof of the theorem. The following three structure is the exhibited:--

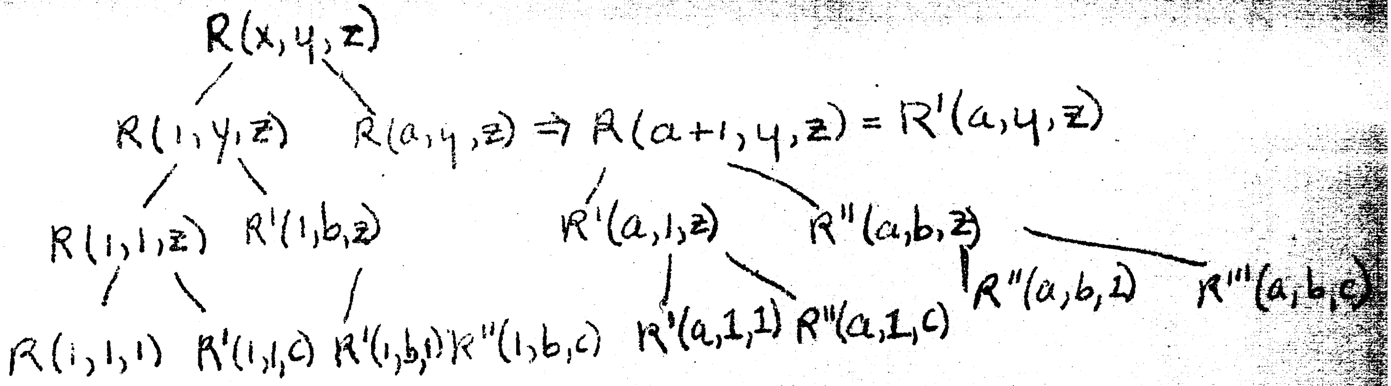


etc.

If there are n such levels and if we assume that $T\alpha 2$ ¢ $T\alpha 1 \Rightarrow T\alpha$ requires K lines, then T will require $\frac{n(n+1)k}{2}$ lines in the Davis Putnam system. In the logic theory type system if T_1 and T_2 were previously proved theorems, only K lines would be required. If there were 3 logical processors at each level the number of terms in the proof of T would increase with n^3 where n is the number of levels and so on.

So far we have restricted ourselves to cases in which there is induction on only one variable. Now consider the proof of $R(x_1,\ldots,x_n)$ of n variables.

Example: 3 variables.

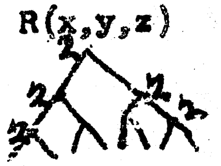


where $R''(a,b,z) = R'(a,b,z) \Rightarrow R'(a,b+1,z)$
etc.

If we have $R(x_1,\ldots,x_1)$ and assume that each step in the tree takes K terms in the Davis Putnam system, then $R(x_1,\ldots,x_n)$ will require $(2^n-1)K$ terms excluding the proofs of the last line of 2^n statements.

Now consider the n variable case when at each level there are 2 logical predecessors of the same type, that is with the same number of variables.

We obtain a tree of the form:--



where the integers at the vertices indicate the number of subtheorems in the partition at that level; i.e. the number of independent copies of the structure below it appearing in the complete structure.

For example if as in the above tree there are two subtheorems at each level, instead of $(Q^n-1)K$ terms, $2/3K(2^{2n}-1)$ terms would be required.

Example. Consider the theorem: pis a prime and plab, then pla or plb.

The canonical representation is

(Axiom System) $f_4(x_1,x_2,x_4,x_7,x_8,x_9,x_{10},x_{11},x_{12})$

$=1+f_8+f_6\cdot f_7=f_5\cdot f_4+\sim f_6=x_{11}\cdot f_4+$

$\sim f_7 \quad =x_{12}\cdot f_4+(x_{10}=1+x_9 \vee f_4=x_{10}+x_8 \cup x_8=f_8\cdot f_4)$

where $f_4=f_4(x_1,x_2,x_4,x_7,x_8,x_9,x_{10},x_{11},x_{12})$

$f_5=f_5(\qquad\qquad\qquad\qquad)$

$f_6=f_6(\qquad\qquad\qquad\qquad)$

$f_7=f_7(\qquad\qquad\qquad\qquad)$

$f_8=f_8(\qquad\qquad\qquad\qquad)$

This requires induction on the five variables $x_8,x_9,x_{10},x_{11},x_{12}$. Assuming the tree structure examined in the last paragraph and letting K=100 (a conservative figure) the number of terms required would be of the order of 10^5. However the tree structure for this theorem is probably more complicated than we have assumed so that this estimate is too low. Nevertheless, 10^5 terms is too long to be practical as a counterexample in contemporary machine programs of the Davis Putnam systems. Further analysis of specific theorems will be carried out in another paper.

The tree structures which would be encountered in practice would be much more complex than those studied above. Hence it seems clear that as the theorems become even moderately difficult the minimal Davis Putnam counterexample to the negation of the theorem becomes too long to be of practical use.

4.0 Conclusions

Hence we see that even a formalization of an informal proof becomes excessively long. As mentionned above we cannot expect the generating procedure even with a clever heuristic selection criterion to do better than this. The essential weakness of the system is that in proving any particular theorem it must also prove every theorem which logically precedes it. This is in contrast to informal mathematics or systems like the logic theory machine in which previous theorems may be referred to. This suggests that this system might be modified to add to the list of axioms those theorems that have already been proved and to use an appropriate metatheorem. Then in formalizing a field such as number theory it would merely be necessary to prove theorems in the right order to avoid excessively long proofs. But even then there is a difficulty in proving theorems whose informal proofs are not known. The only way to make this problem correspondingly feasible would be to get an informal plan but this is of course equivalent to proving the theorem informally. When the Davis and Putnam system was first studied it was suggested that it might be able to get a proof for difficult theorems such as Fermat's last theorem. However, what we have been saying above is that to prove any but the most trivial theorems we require some heuristic planning scheme. But this is equivalent to obtaining an informal proof by a system of the type of the logic theory machine and thus if we cannot prove a theorem informally by a heuristic problem solving program it is extremely doubtful if we can obtain a proof in a feasible length of time by falling back on a formal pseudo decision procedure.

References

1. Davis, M. and H. Putnam, "A Computational Proof Procedure for Quantification Theory," J.A.C.M., Vol. 7, No. 3, p. 201

2. Davis, M. and H. Putnam, "A Finitely Axiomatizable System for Elementary Number Theory," submitted to J. Sym. Log.

3. Wang, Hao, "Proving Theorems by Pattern Recognition," Comm. A. C.M., Vol. 3, No. 4.

CS-TR Scanning Project
Document Control Form

Date : 11 / 30 / 95

Report # AIM - 40

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
□ Laboratory for Computer Science (LCS)

Document Type:

□ Technical Report (TR) ☒ Technical Memo (TM)
□ Other:

Document Information

Number of pages: 16 (20 · images)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:
☒ Single-sided or
□ Double-sided

Intended to be printed as :
☒ Single-sided or
□ Double-sided

Print type:
□ Typewriter □ Offset Press □ Laser Print
□ InkJet Printer □ Unknown ☒ Other: MIMEOGRAPH

Check each if included with document:

□ DOD Form □ Funding Agent Form □ Cover Page
□ Spine □ Printers Notes □ Photo negatives
□ Other:

Page Data:

Blank Pages(by page number):

Photographs/Tonal Material (by page number):

Other (note description/page number):
Description : Page Number:
Ⓐ IMAGE MAP: (1-16) TITLE, 2-16
(17-20) SCANCONTROL, TRGT'S (3)
Ⓑ VERY POOR COPY

Scanning Agent Signoff:
Date Received: 11 / 30 / 95 Date Scanned: 12 / 11 / 95 Date Returned: 12 / 14 / 95

Scanning Agent Signature: Michael W. Cook