Partial non-underflow and non-overflow of an arithmetic stack

D. T. Goodwin  
Department of Computer Science, University of Keele, Newcastle, Staffs., ST5 5BG

This paper continues analysis of the model of arithmetic stack behaviour set up in Goodwin (1977). There conditions computable at high level language compile time were given under which run time stack overflow or underflow could not occur. Because the conditions are strict, it is worth considering the question 'Is it possible for no overflow nor underflow to occur, even if the conditions do not apply?' A number of results are presented here, including an algorithm for finding the smallest stack size, if any, which permits an infinity of different program runs to take place.

(Received August 1976; revised July 1978)

This paper is a sequel to Goodwin (1977) in which stack non-overflow and non-underflow conditions are given when strings of a CF language are loaded symbol by symbol on to a stack. A particular symbol exists whose instance each cause the deletion of the top cell of the stack, in some undefined way. This model can easily be applied to the operation of a run time arithmetic stack, where the particular source program acts as the grammar of the CF language, and transfer-to-store instructions or arithmetic operations correspond to cell removal. A disadvantage of the conditions is that they are strict and only answer the question 'Is the program sure to run?' The author suspects that many useful programs would not satisfy this test. This paper considers the question 'Might the program run?' in the sense of whether it can be decided for any CF language whether some finite or infinite number of its strings cause no overflow nor underflow. The answer to this question could conceivably justify a compile time decision either to reject the program outright, or to run it with code inserted for over- and underflow checks at execution time, if no suitable interrupt were available. The former article discussed what high level language programs could be analysed in this way. Certainly all ALGOL 60 and FORTRAN programs are included.

It must be stressed that the case where only a finite number of strings cause no overflow nor underflow is not unduly restrictive and would not by itself justify rejection of the program. Each language string is a stream of variable references, constants and operators which is generated at run time and corresponds to the complete execution of the program. However the language string disappears symbol by symbol as it is generated. Each variable reference at once signals either a fetch to the stack or a transfer-to-store from the stack. Similarly a number input function would place its result on the stack. The operators immediately stimulate operations on the top cells of the stack. It may now be seen that a single particular language string could be used in an 'almost infinite' number of program runs, depending on the particular values of data supplied. Here the word 'almost' is only necessary because in practice a numeric variable can only take one of a finite, though very large, number of possible values, depending on its initial representation.

The question 'Might the program run?' is carefully phrased, and cannot be equated with 'Does the program have a good chance of running?' or 'Is it likely to run?'. These are matters of probability which can only be settled by considering the meanings, and hence the patterns of usage, of strings in the language. This difficult problem is not dealt with at all.

The algorithms of the sections below are often laborious and are given just to demonstrate computability. No claims for efficiency are made.

Notation and assumptions

This paper relies so heavily on the concepts, terminology and notation of the previous article that no further introduction is made here. For ease of reference the numbering of the theorems and lemmas is continued from that paper. Unless stated to the contrary no assumptions are made in what follows as to the existence of $1^+(S)$, $1^-(S)$, $m^+(S)$ and $m^-(S)$.

Shortest stack length

First considered is the question 'What is the shortest stack length necessary to accommodate some string of $G$ throughout the whole of its deposition on the stack?' This problem turns out to be computable and hence answers the simpler question 'Can any string of $G$ be accommodated using a stack of given length $L$?' At the moment only the upper bound is considered —thus strings of $G$ which use only a few cells of the stack are deemed acceptable even though they may make the stack length negative.

Definitions

For an arbitrary string $s$ of terminals, let $m(U(s))$ be the maximum number of stack cells used at any time during the deposition of $s$. Then $m(U(s)) \geq 1$ for all $s$.

For an arbitrary non-terminal $N$, let $N \Rightarrow s_N$ where $s_N$ is a terminal string. Then let

$$m(U(N)) = \min\{m(U(s)) : s_N \in S_N\}.$$  

$m(U(N))$ clearly exists whatever the nature of $G$, and whether there are a finite or infinite number of $S_N$.

Similarly for any string $u$, where $u \Rightarrow s_u$, let

$$m(U(u)) = \min\{m(U(s)) : s_u \in S_u\}.$$  

Now let $m(U(N))$ be the least maximum stack length used by all strings $s_N$ where the derivation $N \Rightarrow s_N$ begins with the $i$’th rule of $N$. (Let this rule be $N \Rightarrow c_1 \cdots c_j \cdots c_k$). Then

$$m(U(N)) = \min\{m(U(c_1)) : c_1 \cdots c_j \cdots c_k \in S_u\}.$$  

Consider how this least maximum can arise in the deposition of terminal strings derived from $c_1 \cdots c_j \cdots c_k$. From an argument similar to that used to derive the formula for the $m^-(N)$, $m(U(N)) = \min\{m(U(c)) : c \in S_u\}$

where

$$m(U(c)) = \begin{cases} m(U(c_1)) + m(U(c_2)) & \text{if } c = c_1 \cdots c_2 \in S_u \\ m(U(c_1) + \cdots + m(U(c_k)) & \text{if } c = c_1 \cdots c_k \in S_u \\ \max\{m(U(c_1) + \cdots + m(U(c_j)) : j \leq k \} & \text{if } c = c_1 \cdots c_j \cdots c_k \in S_u \end{cases}$$  

153
This minimum is harder to compute than \( m^*(N) \). In that case the \( l^*(c_j) \) are maxima which are known beforehand, while here the minimum may arise for any length \( l(c_j) \) of a terminal string of \( c_j \). It is desired to vary all the lengths involved so that the maximum stack length arising from any line of the formula is minimised.

Consider the \( n \)th line \( l(c_1) + \ldots + l(c_{n-1}) + mU(c_n) \). No other line of the formula involves \( c_n \), so the minimum \( mU(c_n) \) can be used, and regarded as a constant when considering other \( c_j \). Now the only lines to involve \( c_{n-1} \) are the \((n - 1)\)th and \( n \)th, viz.

\[
\sum_{k=1}^{n-2} l(c_k) + mU(c_{n-1}) \sum_{k=1}^{n-2} l(c_k) + l(c_{n-1}) + mU(c_n)
\]

So it is necessary to choose \( s_{n-1} \) and \( c_{n-1} \) so that max \( mU(s_{n-1}), l(s_{n-1}) + mU(c_n) \) is minimised, where \( mU(c_n) \) is a constant. For such \( s_{n-1} \) it is not always true that \( mU(s_{n-1}) = mU(c_n) \). This suggests a new function defined as follows:

\[
\text{Let } mU(u, k) = \min \{ \max \{ mU(s_u), l(s_u) + k \} \}. \\
\text{It follows that } mU(u, 0) = mU(u).
\]

Now if \( u \Rightarrow s_u \) and \( v \Rightarrow s_v \), then

\[
mU(uw, k) = \min \{ \max \{ mU(s_u), l(s_u) + k \} \}.
\]

From the definition of \( mU(u, k) \) it is clear that it increases monotonically with \( k \), so that the above minimum is attained by minimising the second parameter, i.e. \( mU(u, k) = mU(u, \min \{ \max mU(s_u), l(s_u) + k \} \) for all \( s_u \).

A third parameter \( r \) for the \( mU \) function is now introduced. This restricts the terminal strings of \( u \) to those whose generation trees are no more than \( r \) rule-applications deep. Then if \( N \Rightarrow c_1 \ldots c_j \ldots c_n \) is the \( i \)th rule of \( N \),

\[
mU(N, k, r + 1) = mU(c_1 \ldots c_j \ldots c_n, k, r) = mU(c_1 \ldots c_{n-1}, mU(c_n, k, r), r)
\]

and

\[
mU(N, k, r + 1) = \min \{ mU(N, k, r), mU(N, k, r + 1) \}, r \geq 0.
\]

(For low \( r \) some of the \( mU \) function values may not exist since they may eventually call for the evaluation of some \( mU(N, k, 0) \) —impossible since \( N \) is a non-terminal.)

Now since the \( c_j \) for all \( j, i \) and \( N \) are either terminals or non-terminals of \( G \), the above relations define a simple recursive procedure for finding the \( mU(N, 0, r + 1) \), and as many of the \( mU(N, k, r + 1) \) as are required for computing the \( mU(N, 0, r + 2) \). The use of the \( r \) parameter guarantees termination of the procedure, since it specifies that strings of a certain finite measure of complexity are analysed in terms of simpler ones, until eventually terminal strings are reached. No doubt a much faster iterative method could be found which preserved all intermediate \( mU(c_j, \ldots) \) values found so that no duplicated evaluations were made.

It still needs to be shown how to terminate the algorithm, having arrived at values for the \( mU(N) \). Certainly

\[
mU(N) = \lim_{r \to \infty} [mU(N, 0, r)]
\]

because by taking sufficiently large \( r \) every desired terminal string derived from \( N \) will be considered. Also as \( r \) increases, \( mU(N, k, r) \) decreases monotonically and \( mU(N, k, r) \geq mU(N, 0, r) \). Termination takes place at a certain \( r \) value when for each \( N, k \),

\[
mU(N, k, r + 1) = mU(N, k, r),
\]

then by the algorithm would give the same values for the \( mU(N, k, r + 2) \) as it had for the \( mU(N, k, r + 1) \). However for this condition to be useful the \( k \) values that can arise must be shown to be bounded for all \( N, i, r \). This is true for each \( N, i \) because as \( r \) increases \( mU(c_{i, 0, r}) \) is bounded, \( mU(c_{i, 0, r}) \) is bounded and so on. To avoid unnecessary evaluations just to test for termination the condition could be expressed as

\[
mU(N, 0, r + 1) = mU(N, 0, r)
\]

for all \( N, r \) and all \( k \) already evaluated.

It is also easy to see that by following the progress of the above algorithm one may construct a string of \( G \) which uses only the minimum stack length \( mU(S) \). This is by taking note of:

1. the \( i \) value used when evaluating each \( mU(N, k, r + 1) \);
2. the lowest \( r \) for which each of the \( mU(N, k, r) \) attains its minimum value.

Definition

Let \( mL(N) \) for each \( N \) be the maximal lower bounds analogous to the \( mU(N) \). The algorithm to compute them would clearly be similar to the one above. Of course it would be unusual for a single string \( s \) of \( G \) to have the property that \( mU(s) = mL(S) \) and \( mU(s) = mL(S) \). These results are summarised in the following theorem.

Theorem

1. It is possible to compute the shortest stack length necessary to accommodate some string of \( G \) through the whole of its deposition on the stack. A string which uses just this stack length can be found.
2. Suppose the strings of \( G \) are to be deposited on a stack where some cells are occupied already. Then it is possible to compute the least inroad that a string of \( G \) can make upon the cells already occupied. A string of \( G \) which does this can be found.

Corollary

Given a positive integer \( L \) it is decidable whether there is any string of \( G \) which can be deposited on a stack of length \( L \). If there is, then such a string can be found.

A consequence of this theorem is that for many high level language programs which at run time use an arithmetic stack, whether known or accessible to the programmer or not, it is decidable at compile time what the minimum stack length must be for any program to run.

The above discussion is particularly relevant to FORTRAN and ALGOL programs since for these languages \( m^*(S) = 0 \) so that only upper bounds need be investigated.
Upper and lower bounds together

Definition
A string \( s \) of \( G \) is 'stack-bounded in \([L_1, L_2]\)' if the length of the stack during the deposition of \( s \) remains in the finite interval \([L_1, L_2]\). It is only necessary to consider intervals which include the integer zero, since it is normally assumed that the stack length is zero before deposition of a string of \( G \).

Consider first a string \( s \) of terminals, which may be part of a complete terminal string of \( G \). Let the stack length before loading \( s \) be \( L_1 \) and afterwards \( L_2 \). Let the Boolean function \( sb(s, L_1, L_2) \) be true if \( s \) is stack bounded in the interval \([L_1, L_2]\) and false otherwise. Then for any \( s_1, s_2 \),

\[
\begin{align*}
sb(s_1, s_2, L_1, L_2) &= \bigwedge_{l \in \{L_1, L_2\}} \bigvee_{l \in \{l_1, l_2\}} sb(s_1, l_1, l_2) \\
&= sb(s_1, l_1, l_2)
\end{align*}
\]

where \( l \) is the stack length after the loading of \( s_1 \). Now let \( sb(s) = sb(s, L_1, L_2) \), the Boolean matrix of \( sb(s, L_1, L_2) \) values, where \( L_1, L_2 \) may each take all the values in \([L_1, L_2]\). Then

\[
\begin{align*}
\forall s \in S, L_1, L_2, \exists \forall s \in S, L_1, L_2, \exists
\end{align*}
\]

Furthermore, let \( sb(N) = \bigwedge_{l \in \{L_1, L_2\}} sb(s(N), l, l) \), where \( l_1, l_2 \) is the \( l_1, l_2 \)th element of this matrix is true if there is some \( s_0 \) such that if deposition of loading starts at stack length \( L_1 \), it finishes at length \( L_2 \) without exceeding the bounds \( L_1, L_2 \).

Introduce now a second parameter \( r \) which restricts the \( s_0 \)s considered to those who generation trees are no more than \( r \) rule-applications deep. Also denote by \( sb(N, r) \) the matrix arising from the application of the \( r \)th rule of \( N, N \rightarrow c_1 \ldots c_j \ldots c_m \). Then

\[
\begin{align*}
sb(N, r + 1) &= sb(N, r), \cup sb(N, r + 1) \\
&= sb(N, r), \cup sb(N, r + 1), \forall l \in \{L_1, L_2\}
\end{align*}
\]

Some of the \( sb(N, r) \) may not exist for low \( r \), since their evaluation may depend on some \( sb(N, 0) \) which cannot exist since \( N \) is a non-terminal. One could set \( sb(N, 0) = \text{('False')} \).

The above rules define a simple though laborious procedure for finding the \( sb(N, r) \) for all \( N \), and so \( sb(S, r) \), in particular. The procedure must terminate, i.e. there is an \( r \) such that for every \( N sb(N, r + 1) = sb(N, r) \). This is because the number of elements in each \( sb(N, r) \) is fixed, and as \( r \) increases the number of true elements can only increase. If the number of distinct non-terminals is \( m \), then there can be at most \( m(L_2 - L_1 + 1)^2 \) iterations before termination.

Once termination has been reached the truth-values of the matrix elements \( sb(S, 0, l) \) for any \( l \) in \([L_1, L_2]\) show whether any string of \( G \) is stack-bounded in \([L_1, L_2]\).

For any element of the matrix which performs in this way.

Theorem 8
It is decidable whether any string of \( G \) is stack-bounded in a given interval \([L_1, L_2]\). Where it is shown to exist, such a string can be found.

Theorem 9
It is possible to find a least interval \([L_1, L_2]\) in which any string of \( G \) can be stack-bounded.

Proof
Choose any string \( s \) of \( G \) and find \( mL(s) \) and \( mU(s) \). Then there are only a finite number of smaller intervals than \([mL(s), mU(s)]\) which include zero. By examining these in turn as in Theorem 8, a least interval can be found.
Example
Consider the simple grammar whose rules are $S \rightarrow aSC$, $S \rightarrow a$, and the interval $[0, 2]$.

$$sb(a, 0) = \left( \begin{array}{ccc} F & N & F \\ F & F & N \\ F & F & F \end{array} \right)$$

$$sb(c, 0) = \left( \begin{array}{ccc} F & F & F \\ N & F & F \\ F & F & F \end{array} \right)$$

Then $sb(S, 1) = sb_2(S, 1) = sb(a, 0) = \left( \begin{array}{ccc} F & N & F \\ F & F & N \\ F & F & F \end{array} \right)$

Now $sb_1(S, 2) = sb(a, 1) \cdot sb(S, 1) \cdot sb(c, 1)$

$$= \left( \begin{array}{ccc} F & O & F \\ F & F & O \\ F & F & F \end{array} \right) \cdot \left( \begin{array}{ccc} F & N & F \\ F & F & N \\ F & F & F \end{array} \right) \cdot \left( \begin{array}{ccc} F & F & F \\ F & F & F \\ F & F & F \end{array} \right)$$

$$= \left( \begin{array}{ccc} F & N & F \\ F & F & F \\ F & F & F \end{array} \right)$$

and $sb_2(S, 2) = sb(a, 1) = O[sb(a, 0)] = \left( \begin{array}{ccc} F & O & F \\ F & F & O \\ F & F & F \end{array} \right)$

Hence $sb(S, 2) = O[sb(S, 1), sb_1(S, 2), sb_2(S, 2)] = \left( \begin{array}{ccc} F & N & F \\ F & F & O \\ F & F & F \end{array} \right)$

Similarly $sb(S, 3) = \left( \begin{array}{ccc} F & O & F \\ F & F & O \\ F & F & F \end{array} \right)$

from which clearly if $r \geqslant 3$, then $sb(S, r)$ has no ‘new’ elements and there are no more stack-bounded strings.

Proof of Theorem 10 continued
As with the former $sb$ matrices let evaluation continue until $r = r_1$ (say), after which no more ‘false’ elements are over-written with ‘old’ or ‘new’. For the ‘old’ and ‘new’ elements in the $sb(N, r_1)$ for all $N$, one of three possibilities remains:

1. The ‘new’ elements are all replaced by ‘old’ elements as $r$ increases, so that for some $r_2 \geqslant r_1$ no ‘new’ elements remain in any $sb(N, r)$ for $r \geqslant r_2$.

2. The ‘new’ and ‘old’ elements become fixed for $r \geqslant r_3 \geqslant 1$

This is a special case of 3.

3. While the number $k$, (which is no more than $m(L_2 - L_1 + 1)^2$) and identity of the ‘true’ elements are fixed, they oscillate in possessing ‘old’ or ‘new’ values, the only restriction being that for any $r$, they cannot all be ‘old’, for otherwise they would all continue to be ‘old’ for all higher $r$. Thus there are $2^k - 1$ possibilities for the set of ‘old’ and ‘new’ values. Cases 1 and 3 can now be distinguished by examining the $sb(N, r_1 + 2^k)$ matrices for all $N$. Either all elements will be ‘false’ or ‘old’, or there will exist distinct $r_4$, $r_5$, in the interval $[r_1, r_1 + 2^k]$ for which $sb(N, r_4) = sb(N, r_5)$ for all $N$. The former situation is case 1, showing that $G$ is not infinitely stack-bounded in $[L_1, L_2]$, and the latter is case 3. Then provided an element $sb(S, 0, l)$, for some $l$, has the ‘new’ value for some $r$ in the oscillating sequence, then there will be an infinite number of such $r$, and $G$ is infinitely stack-bounded in $[L_1, L_2]$. Hence the theorem is proved.

Theorem 11
If $G$ is known to be infinitely stack-bounded in some interval $[L_1, L_2]$ it is possible to find a least interval in which it is infinitely stack-bounded.

Proof
Since zero must be in any such interval, there are only a finite number of intervals shorter than $[L_1, L_2]$ which need to be examined. The procedure of Theorem 10 can be used for each such interval.

The remainder of the article is devoted to proving computable conditions under which $G$ is infinitely stack-bounded in some interval, and determining a least such interval in the general case. Various trivial ‘theorems’ arise on the way which have some slight value of their own—it is useful to know that the $l(N)$ can be generated parametrically as well as with unwieldy recurrence relations.

Some Diophantine equations
Some Diophantine equations are used a number of times in the succeeding sections. Their relevant properties are therefore summarised here. One is:

$$a_1x_1 + a_2x_2 + \ldots + a_ix_i + \ldots + a_nx_n = a^*$$

where the $a_i$ and $a^*$ are any given integers and solutions in the $x_i$ are sought for integers $x_i > 0$, for all $i$. The results given in Mordell (1969) may be summarised and trivially extended as follows:

1. A solution can only exist if the highest common factor of all the $a_i$ divides into $a^*$.

2. Given 1, then there are always an infinite number of solutions if the $a_i$ do not all have the same sign. The $x_i$ can then be expressed as linear functions of a finite number of parameters.

3. If all the $a_i$ do have the same sign then solutions exist only if $a^*$ also has the same sign. If so then there are either no solutions or a finite number of solutions.

4. Given fixed $x_i$ and integers $A+$ and $A-$ it is computable whether there exists any $a^* < A+$, or alternatively $a^* > A-$, for which there is a solution of (1), and if so what values it may respectively take.

Now let

$$y_j = b_{ij} + b_{ij1}x_1 + b_{ij2}x_2 + \ldots + b_{ijk}x_j + \ldots + b_{ijm}x_n$$

for $i = 1$ or $2$ and $1 \leq j \leq n$, and consider the two simultaneous equations

$$z_{11}y_1 + z_{12}y_2 + \ldots + z_{1j}y_j + \ldots + z_{1m}y_m = 0,$$
$$z_{21}y_1 + z_{22}y_2 + \ldots + z_{2j}y_j + \ldots + z_{2m}y_m = 0$$

(2)

to be solved for non-negative $y_j$ and $x_{jk}$ in terms of the arbitrary integers $b_{ijk}$. The needed result is:

5. It is computable whether there is a solution to the equations (2) for given values of the $b_{ijk}$. The author has a proof of computability, but for brevity it is omitted from this paper.

N-non-recursive subtrees
In subsequent sections it will be shown how to compute whether a particular subset of the basic cycles of $G$ can appear in the derivation of a single string of the language. These subsets will in turn be used to provide formulae for generating all the $l(N)$. As a first step $N$-non-recursive subtrees are defined and their maximum number of possible occurrences is shown to be computable.

Definitions
A generation tree is ‘$N$-non-recursive’ if its base-node is the non-terminal $N$ and it contains no other instance of $N$. The base-node of such a tree is termed ‘non-recursive’ for brevity in what follows. However, note that $N$-non-recursiveness does not necessarily imply zero-recursiveness.

The marks $N \Rightarrow u$ are defined to mean either that $N = u$, or that $N \Rightarrow u$.

For each pair of non-terminals $N'$ and $N$ of $G$, let the maxi-
num number of N-non-recursive subtrees which can exist in the
tree of any expansion of N’ be I(N’, N). Trivially also
define I(t, N) = 0 for any terminal t.

Examples
All zero I(, ) values are omitted.

1. \[ S \rightarrow NN \]
\[ I(S, N) = 2, \]
\[ I(S, S) = 1, \]
\[ I(N, N) = 1. \]

2. \[ S \rightarrow NS \]
\[ I(S, S) = 1, \]
\[ I(S, N) = 1, \]
\[ I(N, S) = \infty, \]
\[ I(N, N) = 1. \]

3. \[ S \rightarrow MSS \]
\[ I(S, S) = \infty, \]
\[ I(S, M) = \infty, \]
\[ I(M, M) = 1, \]
\[ I(N, N) = 1. \]

4. \[ S \rightarrow PS \]
\[ I(S, S) = 1, \]
\[ I(S, M) = 1, \]
\[ I(S, N) = 1, \]
\[ I(N, N) = 1, \]
\[ I(S, P) = \infty. \]

It is easy to see that for any N, either I(N, N) = 1 or I(N, N) = \infty.
Also if N1 \Rightarrow uN2v for some u, v then I(N, N1) \leq I(N, N2).
Note that it does not follow that if N’ is in a cycle then
I(N’, N) will be infinite. The following lemma gives a precise
condition for infiniteness.

Lemma 5
I(N’, N) is infinite if and only if there exists N1 such that
N’ \Rightarrow uN1v for some u, v.
and
u1v1 \Rightarrow u2N2v for some u2, v2.

Proof
1. Sufficiency
Given that exists with the above property, then taking any
integer j and executing the N1 cycle j times,
N’ \Rightarrow uN1v \Rightarrow uuv1N1v, \]
and so ignoring the orders of symbols,
N’ \Rightarrow uuv1N1v. \]
Thus the number of derivable instances of N in an expansion of
N’ is unbounded. Each such N-instance either is or can be used to
derive a non-recursive instance of N.

2. Necessity
Let the largest number of non-terminals on the righthand side
of any rule of G be q. Consider an expansion of some non-
terminal N’ of G such that its tree contains no path between
two non-terminals longer than n arcs, where n is the number of
distinct non-terminals of G. Such a tree can have no more than
q^n zero-recursive non-recursive instances. Now if I(N’, N) is
infinite there must be an expansion of N’ with more than q non-
recursive instances of N, each of which has the property that
the path from N’ to it passes through no other non-recursive
instances of N. Thus these instances may be regarded as the
leaves of a tree which is derived as follows: take the original
tree of the expansion and at each node remove every subtree
which does not contain a non-recursive instance of N. On this
new tree there can still be no more than q subtrees attached to
any one node so that from the argument above there must be a
path from N’ to a non-recursive instance of N which is more than
n arcs long. Hence there exists a non-terminal N’1, not
necessarily distinct from N’ or N, with more than one instance
on a path from N’ to N.

i.e. N’ \Rightarrow uN1v \text{ for some } u, v \]
and
N’1 \Rightarrow u1N1v1 for some u1, v1.

Now suppose that there are no u2, v2 such that u1v1 \Rightarrow u2N2v2,
so that use of the N1 cycle introduces no new instances of N.
Consider the expansion of N’ which is as before except that
now the N1 cycle (together with one instance of N1) is
removed. Then either the path considered still has more than n
non-terminal instances on it, in which case the argument can be
repeated, or no more than n instances. Repeating for each such
path, a tree can be derived which has more than q^n non-
recursive instances of N with no path more than n arcs long—
an impossibility. Therefore there exist u2, v2 such that
u2v2 \Rightarrow u2N2v2.

The condition of this lemma is clearly computable, for it is
equivalent to determining whether certain paths exist between
nodes of a subgraph of G, which is finite.

Now let R(N') = N' -> C_{i_1}C_{i_2} \cdots C_{i_j}C_{i_p}.
Then if I(N’, N) is finite,
I(N’, N) = 1 if N’ = N, and
I(N’, N) = \max \left( \sum_{i=1}^{p} I(C_{i_j}, N) \right) otherwise,
from its definition. From Lemma 5 above the formula is never
usefully recursive in any quantity I(N’, N), which makes its
computation finite and straightforward.

Co-existent basic cycle subsets
Definition
A subset of the basic cycles of G is ’co-existent’ if there exists a
string of the language which uses each of the cycles of the
subset at least once, and no others at all.

In this section an algorithm is developed for determining
whether a basic cycle subset is co-existent. Hence all co-
existent subsets can be found.

There are three possible co-existence relationships between
two basic cycles:

1. The use of one cycle may prohibit the use of the other.
   \[ S \rightarrow N \quad S \rightarrow M \]
e.g. \[ N \rightarrow aN \quad M \rightarrow bM \]
The strings are \{a^n \mid n \geq 1\} \{b^m \mid m \geq 1\}

2. The cycles may be used independently.
   \[ S \rightarrow MN \quad M \rightarrow bM \]
e.g. \[ N \rightarrow aN \quad M \rightarrow bM \]
The strings are \{b^ma^n \mid m, n \geq 1\}

3. One cycle may be used only if the other one is also used.
   \[ S \rightarrow N \quad M \rightarrow bM \]
e.g. \[ N \rightarrow a \quad M \rightarrow g \]
The strings are \{a \mid b^mga^d \mid m \geq 0, n \geq 1\}.

Note that in the simple grammar \{S \rightarrow SS, S \rightarrow a\}, the first
rule defines two basic cycles. A particular application of this
rule, however, may be regarded as the use of either cycle, but
not both. Thus these cycles may be used independently, and this
grammar formally has three co-existent basic cycle subsets,
according to the selection of either or both of these cycles.

Consider the tree of a string of a language in which two cycles
are used. Let N1 and N2 be non-terminal instances, one in each of
the two cycle occurrences. Consider the paths from the
base-node $S$ to $N_1$ and to $N_2$. These paths must diverge at some node of the tree, perhaps at the base-node, and when they do must be where a particular rule

$$N \rightarrow u'N_1'v'N_2'w'$$

is applied, for some $u'$, $v'$, $w'$, where

$$S \rightarrow uN_0, \quad N_1' \rightarrow u_1N_1v_1 \quad \text{and} \quad N_2' \rightarrow u_2N_2v_2$$

for some strings $u$, $u_1$, $v_1$, $u_2$, and $v_2$. Conversely, the existence of such a rule of $N$ guarantees that the two cycles can coexist.

More generally, suppose paths on $G$ from $S$ to $N_1$, $N_2$, ..., $N_s$ in distinct basic cycles all diverge at a particular node, and it is desired to find out whether all these basic cycles can coexist. The cycles can be grouped depending on which rule of $N$ can be used to connect each $N_j$ to $S$. Then all the cycles can coexist if sufficient instances of $N$ can occur to allow all the necessary rules of $N$ to be used. Some of these rules may be parts of cycles, and if so can appear as many times as desired (providing of course that the cycles involved are in the subset currently being investigated). Once any such rules have been allowed for, the desired number $D(N)$ of instances of $N$ must all be taken from the non-reverse instances of $N$. Hence the condition for co-existence of the cycles when the paths join at $N$ is

$$D(N) \leq I(S, N)$$

This condition can now be made the basis of the algorithm to determine whether a subset of basic cycles is co-existent. Firstly, for each cycle determine all the non-cyclic paths on $G_R$ from $S$ to each non-terminal in the cycle. Now also allow paths with circuits to be added provided they introduce further distinct rule instances.

Now select one path for each cycle out of the finite set just established and determine as above whether this selection of paths can all appear in any one string of the language. Repeat until a co-existent selection is found, or otherwise until all such selection combinations have been exhausted to prove non-coexistence.

An analogous algorithm can be applied to find the basic cycle subsets which may co-exist in some expansion of a non-terminal $N$ distinct from $S$. Here the computation described would use not $G_R$ and $G$ but the subgraph of $G_R$ containing only nodes accessible from $N$ with the associated sub-grammar of $G$. Of course, the cycles of this subgrammar would be a subset of the cycles of $G$.

**Generating functions for the $l(N)$**

Since the recurrence relations

$$l(N) \leftarrow \sum_{j=1}^{r} l(C_j)$$

are not easy to work with, some linear functions $f_N^i$ are now introduced whose arguments can take all the positive integer values and whose domains, taken together, constitute all the $l(N)$.

**Definitions**

For any recursive chained derivation $N \Rightarrow uNv$ call the string $uv$ the 'result' of the derivation, or of the cycle which is involved. Here the order of the symbols in $uv$ is ignored.

Now let all the $B_N^i$, the co-existent basic cycle subsets of $N$, be denoted by

$$B_N^1, \ldots, B_N^i, \ldots, B_N^s,$$

and let the basic cycles in $B_N^i$ be specially numbered

$$C_N^1, \ldots, C_N^i, \ldots, C_N^s.$$ 

Now for fixed $i$ and $j$ there may well be more than one zero-cyclic length $l_0(C_N^j)$. Call them

$$l_1^j, \ldots, l_k^j, \ldots, l_t^i$$

where $t$ varies depending on $C_N^j$.

In the algorithm for determining the co-existence of the subset $B_N^i$, terminal expansions were found which each included a path selection but had no further cycles. Call them

$$e_1^i, \ldots, e_N^i, \ldots, e_N^i,$$

and let the number of instances of

$$C_N^j \quad \text{in} \quad e_N^i \quad \text{be} \quad n_N^{jb}.$$

for any particular $j$, $1 \leq j \leq s$.

The generating functions for the $l(N)$ can now be presented.

**Lemma 6**

1. Every terminal expansion of $N$ has a length

$$f_N^i(x_{i,j}) = l(e_N^i) + \sum_{j=1}^{r} \sum_{k=1}^{s} x_{i,j} l_k^i$$

for some $i$, $h$ and a vector $x_{i,j}$ of non-negative integers $x_{i,j}$ provided that for fixed $i$, $h$ and $j$

$$\sum_{k=1}^{r} x_{i,j} = 0 \text{ only if } n_N^{jh} > 0.$$

2. Conversely, for every selection of $i$, $h$ and $x_{i,j}$, (according to the rule just given), the above expression gives the length of some expansion of $N$.

Before giving the proof an explanation of the $f_N^i$ is useful: $N$ fixes the cycle subset and $h$ specifies a particular minimal expansion of $N$ which will permit all the cycles of $B_N^i$ to be used. The $x_{i,j}$ fix the number of additional times a particular cycle is to appear, with its result yielding a particular zero-reducible tree structure. Recursiveness in the expansion of the result of the cycle is allowed for by varying some of the $x_{i,j}$. The condition of 1 specifies that each cycle in the subset must appear at least once. Only when the $e_N^i$ concerned contains a particular cycle $C_N^j$ can

$$\sum_{k=1}^{r} x_{i,j} = 0,$$

be zero.

**Proof of Lemma 6**

1. Consider a terminal expansion $e_N$ of $N$. Then either it is 0-recursive or its tree contains a basic cycle whose result is 0-recursive. In the former case the lemma is proved. In the latter case it means that a node $N'$ has a proper subtree with base-node $N'$ such that

$$N' \Rightarrow vN'v \Rightarrow uvw,$$

where $uv$ is 0-recursive. Now form a new tree and a corresponding new expansion $e_{NT}$ by putting the latter instance of $N'$, with its subtree ($N' \Rightarrow w$), in place of the former instance of $N$. This has the effect of removing the basic cycle with the 0-recursive result:

$$l(e_N) = l(e_{NT}) + l(uv)$$

for some $i$, $j$, $k$.

The next step is to reduce $e_{NT}$ in the same way, ensuring when removing cycles that for each non-terminal in the removed cycles some instance is retained in the new reduced subtree. Eventually no further cycle instances can be removed, bearing the above restriction in mind, and the resultant tree's expansion is one of the $e_N^i$. This establishes the formula.

2. From the construction of any expansion $e_N^i$, it is possible to add as many instances of cycles $C_N^j$ with 0-recursive results as desired. Hence any particular formula certainly has a corresponding expansion of $N$.
Computability of the $I(N)$

**Theorem 12**

1. It is computable whether any expansion of $N$ has a given length $L$, and also whether there is a finite or an infinite number of such expansions. In the former case they can all be determined. In the latter case formulae can be derived from which as many such expansions can be generated as required.

2. As 1, but with the expansions required to be in a finite interval $[L_1, L_2]$. 3. As 2 but with the expansions in infinite intervals $[-\infty, L_2]$ or $[L_1, +\infty]$.

**Proof**

1. From Lemma 6 the problem is equivalent to solving all the equations

$$f_N^k(x_{i,j,k}) = L$$

in the $x_{i,j,k}$, for all $i$ and $h$. These are not simultaneous equations. Each separately may contribute solutions for the $x_{i,j,k}$ and hence expansions of $N$. Such an equation is an example of the Diophantine equation (1) given above. A solution in the $x_{i,j,k}$ specifies the number of instances of each basic cycle that each corresponding satisfactory expansion may have. There are only a finite number of terminal expansions for each such solution. They can be found exhaustively by developing generation trees in which all combinations of alternative expansions of the non-terminals are taken.

2. Repeat 1 for all $L$ in $[L_1, L_2]$.

3. Derive all suitable $a^*$ in $[-\infty, L_2]$ or in $[L_1, +\infty]$ using Diophantine equation result 4 above (Mordell, 1969). Then continue as just above.

**Theorem 13**

As Theorem 12, but considering any string $u$ rather than $N$.

**Proof**

1. Permute the symbols of $u$ until $u = vw$, where $v$ consists of all the terminals of $u$, and

$$w = N_1 \ldots N_y \ldots N_b,$$

being all the non-terminals of $u$. Then the problem is equivalent to solving separately all the equations

$$L = l(v) + \sum_{y=1}^{b} f_N^k(x_{i,j,k})$$

for all possible combinations of the $i$ and $h$ varying separately for each $N$. These equations have the same Diophantine form as those of Theorem 12 above.

2. As for Theorem 12.

**Is there any infinite stack-boundedness?**

**Theorem 14**

$G$ is infinitely stack-bounded in some interval $[L_1, L_2]$ if and only if there exists a cycle $C$, not necessarily basic, such that

$$l(C) = 0 \quad \text{and} \quad lh(C) = 0.$$  

**Proof**

1. **Sufficiency**: take a string of $G$ whose derivation includes a non-terminal of the cycle for which the conditions are true. Then apply the cycle to generate an infinite of stack-bounded strings.

2. **Necessity**: suppose it is known that $L_1, L_2$ exist for which an infinity of strings are stack-bounded in $[L_1, L_2]$. The idea of the proof is to show that there is a suitable cycle with zero left-hand length, and then that there are sufficient of these to guarantee that one of them must have zero righthand length as well.

Consider the infinite set

$$B = \{b_1, \ldots b_i, \ldots \}$$

which consists of all the finite strings $b_i$ of $G$ which are stack-bounded in $[L_1, L_2]$. Since there are only a finite number of rules $S \rightarrow w$, then there must exist one of them, say $S \rightarrow w_1$, which is the first rule applied in the derivations of an infinite subset $B_1$ of $B$. Similarly there must exist a rule $N_i \rightarrow w_2$, where $S \rightarrow w_1 = u_1 N_i v_1$, for some $u_1, v_1$ such that an infinite subset of $B_1$ start off their derivations with $S \rightarrow w_1, N_i \rightarrow w_2$, so that $S \Rightarrow u_1 w_1 v_1$. Similarly for any integer $k$, there exists an infinite set $B_{k+1}$ each of whose members has a derivation which begins

$$S \Rightarrow u_1 u_2 \ldots u_k w_{k+1} v_{k+1} = v_1 \equiv y_1 \ (say)$$

$$\Rightarrow u_1 u_2 \ldots u_i u_{i+1} \ldots u_{i+k} N_i v_{i+1} \ldots v_{i+k+1} \equiv y_{i+1} \ (say).$$

Let $u_1 u_2 \ldots u_k \equiv y_0, v_1 \equiv y_0$, and let $u_{i+1} \ldots v_{i+k+1} \equiv y_i, \ldots u_{i+k+1} \equiv y_i$. Then $\exists y_1 u_1 = u_1, \ldots u_{i+k} N_i v_{i+1} \ldots v_{i+k+1}$ and $S \Rightarrow u_0 u_1 u_2 \ldots u_i u_{i+1} \ldots v_1 \equiv y_{i+1}$.

Consider the loading of a particular terminal derivation of $y_{i+1}$ on to the stack, and let the stack-length after the deposition of the terminal string derived from $u_i$ be $l_{n_i}$ for $0 \leq n \leq f$. Choose $f > L_2 - L_1 + 1$. Then because the terminal string does not exceed the bounds $L_1$ and $L_2$, there must be distinct integers $r$ and $s$ such that

$$l_{1} = l_{r}, \quad \text{i.e.} \quad l(u_1 u_2 \ldots u_f) = 0,$$

and there exists a chained recursive derivation

$$N_{j_{r+1}} \Rightarrow u_{r+1} \ldots u_{j_{r+1}} N_{j_{r+1}} v'_{r+1} \ldots v''_{r+1}$$

whose left-hand length is zero. More generally, for sufficiently large $f$ there exists $l, l_1 \leq i \leq L_2$, and distinct integers $r_1, r_2, \ldots r_{L_2 - L_1 + 2}, \ldots f$, such that

$$l = l_{r_1} = l_{r_2} = \ldots = l_{r_{L_2 - L_1 + 2}}.$$  

Consider now the loading of a terminal derivation of $v_{r_1} \ldots v_{r_{L_2 - L_1 + 2}}$ on to the stack. Let $l'$ be the stack length just before $v_{n}$ is loaded. Consider the stack lengths $l'_{r_n}$ for $1 \leq p \leq L_2 - L_1 + 2$. Then there exist integers $a, b, a > b$, such that

$$l'_{r_{n}} = l'_{r_{n+1}}, \quad \text{i.e.} \quad l(v_{r_{n}} v_{r_{n+1}} = 0,$$

and from above

$$l(u_{r_{n+1}} \ldots u_{r_{n+1}} u_{r_{n+1}}) = 0.$$  

Hence the cycle corresponding to the chained recursive derivation

$$N \Rightarrow u_{r_{n+1}} \ldots u_{r_{n+1}} N_{r_{n}} v_{r_{n+1}} \ldots v_{r_{n+1}}$$

has a zero left-hand length and zero righthand length, and hence zero overall length, as required.

**Theorem 15**

It is computable whether there exist integers $L_1, L_2$ such that $G$ is infinitely stack-bounded in $[L_1, L_2]$. If so such an interval can be found.
Proof
From Theorem 14, it is sufficient to show that it is computable whether there exists a cycle $C$ for which $l(C) = 0$ and $lh(C) = 0$. Let $\{C_1, \ldots, C_t\}$ be a composite subset of the basic cycles of $G$. Then consider a cycle $C$ composed from $y_j$ instances of $C_j$, for each $j$, $1 \leq j \leq t$. Then

$$l(C) = \sum_{j=1}^{t} y_j \cdot l(C_j)$$

$$lh(C) = \sum_{j=1}^{t} y_j \cdot lh(C_j).$$

In general there are a finite number of alternative expressions for $l(C)$ and $lh(C)$, each on the lines of the proof of Theorem 13. For each combination of alternatives the simultaneous equations

$$l(C) = 0$$

$$lh(C) = 0$$

take the form of Diophantine equations (2) and the Diophantine result 5 shows it to be computable whether a solution exists.

If one of these sets of equations proves to be solvable, then the algorithm for deciding that also leads readily to one solution at least. This means that one cycle with the desired properties is known. Suppose that $N$ is a non-terminal in the cycle, say $N \Rightarrow u_iN_v$, where $u_i$ and $v_i$ are terminal strings. Then choose any derivation $S \Rightarrow u'Nv'$ for any terminal $u'$ and $v'$, and a non-cyclic derivation $N \Rightarrow s_N$. One can then show that an infinity of terminal strings generated by the cycle is $u'u_rN_vw'$, $n \geq 1$, whose upper and lower bounds are the same as those for $u'u_rN_vw'$, where $n = 1$. Thus one suitable interval $[L_1, L_2]$ can be found if any exist at all.

**Theorem 16**
It is possible to find a smallest interval, if any, in which $G$ is infinitely stack-bounded.

**Proof**
This theorem just brings together the results of Theorems 11 and 15.

References

---

**Book reviews**


The text is drawn from an advanced course on software portability given at the University of Kent at Canterbury in Spring 1976 and comprises the copious session material of the individual speakers. The course was sponsored by SRC under the auspices of the Informatics training group of the EEC Scientific and Technical Research Committee. Speakers were drawn from both sides of the Atlantic and appeared very much as a listing in ‘who’s-who’ in Software Portability.

The book is divided into eight ‘parts’ each being subdivided into chapters representing individual speaker’s contributions. Each chapter has appended a list of references and these define the state-of-the-art at the time of writing. Parts 1 and 2 are an introduction which outlines the format of the text and a basic concepts review which is really a justification of what follows. Part 3 concerns itself with mechanisms which affect and aid portability. It includes descriptions of the roles played by verifiers and filters, computer architecture and microprogramming, and microprocessors. The problems associated with portable compilers and the translation between high level languages are discussed. Part 4 considers pragmatics, the practical minutiae which so greatly affect portability as a useful tool. Considerations include engineering, the system interface, performance and optimisation. This ‘part’ is very practical and informative. Part 5 has a novelty as well as great general use. It addresses itself to legal aspects and provides a valuable insight into the legal standing of ‘information’. The session was driven by the portability concept but turns out to be of more general interest. In part 6 a number of case studies are presented, ranging from the portability of ALGOL 60, through SNOBOL 4, BCPL, commercial software, data, FORTRAN and GENESYS, an operating system, and NAG’s approach, to the view of a manufacturer (ICL). The experiences recounted certainly ring true to the reviewer. In Part 7 two of the leading lights in the field (Waite and Griswold) combine to discuss current research in the area and take the possibly dangerous step of predicting the future. Part 8 presents two study group reports on portability, those of CNRS/SRC and EEC.

The text is a wide ranging survey of the nature, problems and mechanisms involved in software portability and succeeds in this area. Since it is formed of a series of papers from different contributors the book suffers to a certain degree because of variations in style, depth and general presentation (including typesetting); however the effect is not as bad as that experienced in many similarly organised texts.

The fragmentation of the book does have an advantage: each chapter is largely self contained and therefore the text easily lends itself as a reference document with respect to the different facets of portability. In general the level of the material is at least final year undergraduate. The contents will be of interest to the software industry at large, since many practical lessons may be learnt from within its covers, without the pain of finding out on a ‘real’ system.

ROY NEWTON (Middlesbrough)

*The Challenge of Microprocessors* by M. G. Hartley and Anne Buckley, 1979; 200 pages. (Manchester University Press, £7.95)

Hartley and Buckley are the editors of the *International Journal of Electrical Engineering Education*, and this book consists mainly of papers from a special issue of this journal. It is divided into four main sections: Overall Perspective; General Laboratory Activity; Laboratory Experiments; and Courses and Services. Unfortunately, the editors seemed unable to find sufficient material about microprocessors and have padded out the book with chapters on the use of programmable calculators and a description of a large computerised education system, which are totally unrelated to the rest of the book.

An attempt is made to provide a coverage of a wide range of microprocessor topics, which succeeds only in generating confusion. There is little coherence between the chapters, with the result that some are aimed at educators who are totally unfamiliar with microprocessors, while there are three chapters on microprogramming and bit-slice processors, which seem to be aimed at budding minicomputer designers. The typical educational institution approach of costing equipment merely by the value of the electronic components is noticeable, and highlights the lack of information about the commercial side of microprocessors, such as the development costs of hardware and software, project management and so on.

There is little in this book to commend it to the general reader, and since it does not aim at an identifiable part of the microprocessor arena, or of the educational spectrum, it can also not be recommended to any specific audience.

HARLEY QUILLIAM (Guildford)