<u>On a problem posed by W.H.J.Feijen</u>.

Given, for $M \geq 0$ and $N \geq 0$, the strictly increasing sequences

$$F(0) < F(1) < \ldots < F(M-1)$$

$$G(0) < G(1) < \ldots < G(N-1)$$

and the desired final state

R: $Q(M, N)$

where the predicate Q has been given by

Q(m, n): $(\underline{N}\ (i, j): 0 \leq i < m,\ 0 \leq j < n: F(i) = G(j)) = k$

(the "<u>N</u>" to be read as: "the number of distinct pairs (i, j) with i and j in the ranges so and so, such that such and such".) W.H.J.Feijen posed the problem of giving a formal derivation, heuristics included, of an efficient program establishing R .

The standard way of solving this problem would be "Replacing constants by variables", i.e. would be to introduce two local variables --which we shall again denote by m and n -- , their role being given by the invariant

P: $Q(m, n)$ <u>and</u> $0 \leq m \leq M$ <u>and</u> $0 \leq n \leq N$ ,

easily initialized by

"m, n, k := 0, 0, 0"

and enjoying the obvious property

(P <u>and</u> m = M <u>and</u> n = N) => R . (1)

(It could be remarked that also initializations such as "m, n, k := 0, N, 0" or "m, n, k := M, 0, 0" would have done the job, but symmetry should never be destroyed lightly.)

Appealing to (1) implies a final state with (m, n) = (M, N) ; the crucial observation is that the transition to it from the initial state (m, n) = (0, 0) can be made via so many different paths (increasing m or/and n by 1 at a time) that we should raise the question whether we can exploit this freedom by <u>strengthening</u> P into P' and by <u>weakening</u> the condition for termination. (The weaker the condition for termination, the stronger the guards and the sooner we may expect the repetition to stop.) We can weaken the condition for termination by not requiring that both the local variables m and n have reached their maximum value, i.e. we find ourselves looking for a P', stronger than P, and such that

$$(P' \text{ \underline{and} } (m = M \text{ \underline{or} } n = N)) \Rightarrow R \qquad (2)$$

Investigating the case $m = M$, we find ourselves facing the question: what more should be known besides P such that we may conclude R ? In view of the monotonicity of the sequences F and G it would suffice to know that $F(M-1) < G(n)$ because

$$(Q(M, n) \text{ \underline{and} } F(M-1) < G(n)) \Rightarrow R \quad .$$

Because we are only interested in the inequality $F(M-1) < G(n)$ --itself to strong to be included in the invariant-- the inclusion of $F(m-1) < G(n)$ would suffice. Hence we find ourselves considering the invariant (symmetry!)

$$P': \quad P \text{ \underline{and} } F(m-1) < G(n) \text{ \underline{and} } G(n-1) < F(m) \qquad ,$$

an invariant that admits the same initialization if we define formally $F(-1) = G(-1) =$ minus infinity .

The standard routine of computing wp("m:= m + 1", P'), wp("n:= n + 1", P'), and wp("m, n, k := m + 1, n + 1, k + 1", P') and simplifying by omitting implied terms from the conjunctions gives the guards; the following program results:

```
m, n, k := 0, 0, 0; {P'}
do m ≠ M and n ≠ N →
     if F(m) < G(n) → m:= m + 1 {P'}
      ▯ G(n) < F(m) → n:= n + 1 {P'}
      ▯ F(m) = G(n) → m, n, k := m + 1, n + 1, k + 1 {P'}
     fi {P'}
od {R}          .
```

Plataanstraat 5
5671 AL NUENEN
The Netherlands

prof.dr.Edsger W.Dijkstra
Burroughs Research Fellow