A short note on symmetric distributed arbitration.

This is only a short note. Firstly, my problem is only partly solved; secondly, a pending holiday precludes a longer note now. The results obtained so far, however, deserve recording.

We consider the classic mutual exclusion problem of N cyclic processes, each with its critical section and to be synchronized in such a fashion that at any moment at most one of the processes is engaged in its critical section. The equally classic solution employs one binary semaphore, traditionally called "mutex" and initialized at 1 . Each of the processes can then be coded as follows:

```
do true → P(mutex);
          critical section ;
          V(mutex);
          noncritical section
od .
```

Here the semaphore mutex is a synchronization tool equally accessible to all N processes; for that reason the above solution is called "centralized"; as used, the semaphore mutex accomplishes N-fold arbitration. In this short note we occupy ourselves with the problem of achieving N-fold arbitration when twofold arbitration (between identified partners) is the only available primitive.

The classic solution is as follows. Place each of the processes in a distinct node of the complete N-graph, and associate with each edge a binary semaphore. Each "critical section" is preceded by the N-1 P-operations on the binary semaphores associated with the edges meeting at the node in which the process in question is placed; the critical section is followed by the compensating V-operations.

All solutions of this form obviously ensure N-fold arbitration, since no two processes can be simultaneously engaged in their critical sections. The only remaining problem is to prescribe for each node the order in which the N-1 P-operations occur in such a fashion that the danger of deadlock is absent.

Note. Because each binary semaphore is accessed by two processes, it blocks at any moment at most one process. As a result the question doesn't arise which of the blocked processes is allowed to complete its P-operation when the V-operation on it is performed. Furthermore, the fact that each semaphore blocks at most one process precludes the danger of infinite overtaking, and hence a solution which is free of the danger of deadlock is also free of the danger of individual starvation. (End of Note.)

A very simple way of avoiding the danger of deadlock is the following. Assign to each edge a "rank" such that no two edges that have an endpoint in common have the same rank, and let each process complete the P-operations in the order of increasing rank.

This note is wrong (17/10/80)

Note. The introduction of the ranks is no constraint at all. If in every node the order of the N-1 P-operations has been prescribed in such a way that the danger of deadlock is absent, ranks can be assigned to the edges in such a way that in each node the P-operations occur in the order of strictly increasing rank. The latter requirement imposes a total ordering on the ranks of the N-1 edges meeting at a node; the absence of the danger of deadlock implies that these N total orderings of subsets of N-1 edges are compatible. (End of Note.)

All the above we knew for a long time; it only serves as an introduction for the following results.

* * *

Scholten's result. For each deadlock-free assignment of the edge ranks the number of different complete blocking states equals 2^{N-1}. A complete blocking state is one in which one of the processes is engaged in its critical section while among the others no further arbitrations are possible.

Proof. A complete blocking state is characterized by the process in its critical section and for each other process by the rank of the edge for which arbitration took place in its disfavour. By dealing with the edges in a fixed order of (weakly) ascending rank we can construct any complete blocking state by the following nondeterministic algorithm.

A variable c is initialized with the set of all N nodes; we then deal with the edges in the fixed order. When we encounter an edge with both endpoints still in c , one of the two endpoints is removed from c , otherwise c is left unchanged; a node removed is tagged with the rank of the edge in question.

Because each edge has two different endpoints, c contains at least 1 node all through this execution of the above algorithm; because we process all the edges of the complete N-graph the execution terminates with at most 1 node in c . We conclude that execution terminates with exactly 1 node in c ; this node identifies the one and only process in its critical section.

Because N-1 nodes have been removed from c, we have had N-1 binary choices; hence there are 2^{N-1} different complete blocking states possible. (End of Proof.)

<u>Corollary</u>. No symmetric ranking exists or N is a power of 2.

<u>Proof</u>. With a symmetric ranking the number of complete blocking states with the same process in its critical section is independent of the identity of that latter process; in that case N is a divisor of 2^{N-1}, i.e. N is a power of 2. (End of Proof.)

* * *

<u>Dijkstra's result</u>. A symmetric ranking exists or N is not a power of 2.

<u>Proof</u>. The above is proved by constructing a symmetric ranking in the case that N is a power of 2. Number the nodes from 0 through $N-1$. The digits of the binary representation of the rank of the edge connecting nodes x and y are defined as the sum <u>mod</u> 2 of the corresponding digits of the binary representations of x and y. Thus the ranks 1 through $N-1$ are assigned, and the $N-1$ edges meeting at a single node are all of different rank. Renumbering the nodes by inverting in the binary representation of each node number the k-th digit does not affect the ranks thus assigned. Hence, given the ranks, any node can be regarded as node 0, and the ranking is therefore symmetric. (End of Proof.)

<u>Acknowledgement</u>. With the aim of introducing the op-

portunity for recursive arguments, W.P. de Roever suggested to me to investigate first the situations in which N is a power of 2. (End of Acknowledgement.)

* * *

After having exchanged these two independently formed results, Scholten and I investigated the worst case delay possible with the symmetric ranking just described. We asked ourselves the following question. What is the maximum number of critical section traversals between the moment that a process has completed its noncritical section and the next completion of its critical section? In this investigation we made the weakest assumption that guarantees individual progress: when a process is blocked because of an arbitration decided in favour of a competitor, that arbitration is inverted when that competitor leaves its critical section. For $N = 2^k$ that maximum number of traversals equals

$$(\underline{P}\, i : 0 \leq i < k : 2^i + 1)$$

i.e. for N = 2, 4, 8, 16, ... the maximum number of traversals equals 2, 6, 30, 270 ....

Consider, for instance, the case N = 16. The nodes can then be divided into two sets A and B of 8 nodes each, such that the edges of rank ≤ 7 connect two nodes from the same set and the edges of rank ≥ 8 connect two nodes from different sets. Note that each of the two sets in isolation presents with its internal edges our standard ranking for the complete 8-graph. From a worst case history for set A in isolation — according to our induction hypothesis of length 30 — we can construct for the 16-graph a history of length $9 * 30 = 270$ by inserting in front of each traversal from A the eight traversals from B (each time in the proper order). It is not difficult to see that the resulting history represents a worst case history for the standard ranking of the complete 16-graph.

* * *

For $N = 2^k$ we have investigated the standard ranking, which is a special way of assigning symmetrically the ranks from 1 through $N-1$ to the edges such that no two edges of the same rank have an endpoint in common. How special is the standard ranking?

Select two distinct ranks p and q and travel from an arbitrary node V along the edges of ranks p and q alternatingly

$$V \underset{p}{\text{———}} W \underset{q}{\text{———}} X \underset{p}{\text{———}} Y \underset{q}{\text{———}} Z$$

If the ranking is to be symmetric in the nodes, the rank of VX equals that of XZ. Since only one edge of that rank ends in X, nodes V and Z are the same node! The edges of two ranks form cycles of length 4; its "diagonals" have a third rank. As a result, edges of a fourth rank connect our 4-tuples pair-wise and edges of three other ranks pair our 4-tuples in the same fashion. Etc.

From this it follows that any symmetric ranking of the edges such that no two edges of the same rank have an endpoint in common must be a permutation of a standard ranking. With N=8 I investigated the maximum number of traversals for an essentially different ranking and found 32 instead of the 30 for the standard ranking. At the time of writing I don't know whether the standard ranking gives rise to the best worse case behaviour.

Plataanstraat 5
5671 AL NUENEN
The Netherlands

13 July 1980
prof.dr. Edsger W. Dijkstra
Burroughs Research Fellow