

On-the-Fly Garbage Collection: An Exercise in Cooperation

Edsget W. Dijkstra, Leslie Lamport, A.J. Martin and E.F.M. Steffens

Communications of the ACM, 1978

Presented by Almog Benin
25/5/2014

This paper was the first one to suggest a concurrent garbage collection mechanism. The paper started by formalizing the problem in such a way it will be comfortable to describe the algorithms and prove their security. The suggested solution for concurrent garbage collection was presented through 3 milestones.

The first one is relatively intuitive but coarse-grained. It uses 3-colors for graph's nodes. The correctness of this solution is based on the invariant that there are no "black-to-white" edges in the graph through the whole execution of the algorithm.

The second one is less intuitive and also coarse-grained, to the same extent, but helps toward presenting the fine-grained solution. This solution modifies the former as little as changing the operations' order. This modification, although looks not important, allows weakening the "black-to-white" invariant of the former solution while keeping the algorithm correct.

The third one just breaks the atomic operations of the second solution. By doing so, they presented the desired solution for the concurrent garbage collection problem. For proving the correctness of this new solution, they weakened the invariants of the second solution even further, and showed that even this weak invariant is enough to prove the correctness of this solution.

As for my original idea – the original paper dealt with systems where there is just one job (process) which runs some threads concurrently. In modern computer systems, there are many processes, and each of them has its own threads. By applying the presented technique to these systems, we need to implement one collector thread of GC per process, which may be overkill and consume too many resources of the system. One way to overcome this issue that I thought about is to set just one collector thread, which will be shared between all of the processes. However, such a solution raises some implementation and security issues, so it may not be trivial.

No issues were raised from the audience because of time constraints.