4 Bookmark Algorithms for Paging

There is a randomized algorithm for the paging problem which is H_k competitive for cache size k. That algorithm, PARTITION, is rather complicated. Here, we give a highly simplified version, which we call BOOKMARKS, which works only for $k = 2$.

At any given time, the algorithm maintains a list of "bookmarked" pages in slow memory. There is no limit to the length of this list. Initially, there are no bookmarks.

We call the two pages in cache the "LRU" page and the "MRU" page. The MRU page is defined to be the page that was requested in the previous step, while the LRU page is simply the other one. Each time a request is made, those labels may change.

The Randomized Algorithm BOOKMARKS

1. If MRU is requested, do nothing.
2. If LRU is requested, erase all bookmarks.
3. If a bookmarked page is requested, eject LRU and erase all bookmarks.
4. If there are M bookmarks, and then an unbookmarked page in slow memory is requested, eject LRU with probability 1/(M+2), and eject MRU with probability (M+1)/(M+2). Bookmark the ejected page. (There will now be M+1 bookmarks.)

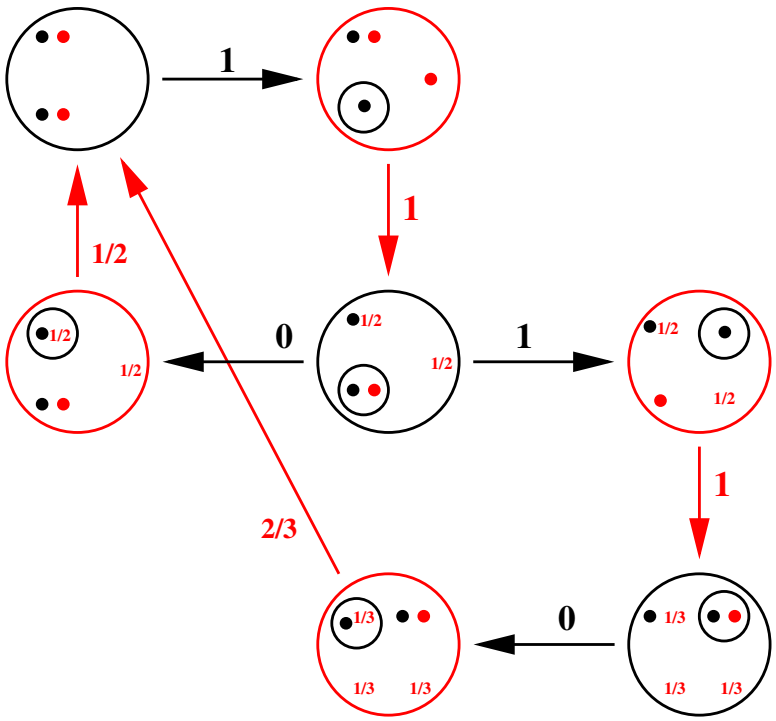


Figure 4.1: BOOKMARKS

Theorem 4.1 BOOKMARKS *is* $\frac{3}{2}$*-competitive.*

Proof: Define a "phase" as follows. Before the first request of a phase, the support of the work function consists of one configuration. The last request of a phase is when the support of the work function is once again just one configuration, except for a possibly incomplete phase at the end.

The proof consists of first verifying the following claims. We omit these verifications.

Claim 1: The cardinality of *Supp*(ω), the support of the work function ω, is $M+1$, where M is the number of bookmarks. If $b_1, ...b_M$ are the bookmarked pages, *Supp*(ω) consists of the pairs $(\text{MRU}, \text{LRU}), (\text{MRU}, b_1), ...(\text{MRU}, b_M)$.

Claim 2: If X in *Supp*(ω), the probability that $(\text{MRU}, \text{LRU}) = X$ is $\frac{1}{M+1}$.

Claim 3: If the length of a complete phase is d, then the optimal cost for that phase is at least $d-1$. For an imcomplete phase of length d, the optimal cost is d.

Claim 4: If the length of a phase is d, then BOOKMARKS' cost for that phase is $d - \frac{1}{d}$. For an incomplete phase of length d, then BOOKMARKS' cost is d.

The theorem follows immediately from the claims. □

The lowest competitiveness for any randomized algorithm for the paging problem is known to be H_k. The lowest competitiveness for any known randomized algorithm for the paging problem that uses no bookmarks is $2H_k - 1$. It is natural to ask whether there a tradeoff between the number of bookmarks and competitiveness.

Theorem 4.2 EQUITABLE *is an* H_k*-compeititive randomized algorithm that uses* $O(k^2 log k)$ *bookmarks.*

The actual formula is approximately $(5k^2 H_k)$, which works out to abut 30 if $k = 2$.

Theorem 4.3 *There is no* $\frac{3}{2}$*-competitive randomized algorithm for the paging problem for* $k = 2$ *that uses no bookmarks.*

Theorem 4.4 *There is a* $\frac{3}{2}$*-competitive randomized algorithm for the paging problem for* $k = 2$ *that uses at most one bookmark at any given time.*

Exercise 4.1 *Problem. What is the minimum number of bookmarks needed for the case* $k = 3$*, to achieve* $\frac{11}{6}$*-competitiveness? The number certainly does not exceed the number needed by* EQUITABLE*, but that number is ridiculously large.*

Exercise 4.2 *For general k, find a relationship between the number of bookmarks used and the competitiveness. More specifically, Borodin and El Yaniv as whether we can obtain* H_k *competitiveness with* $O(k)$ *bookmarks.*