Lower bounds for paging

**Theorem 1** No deterministic algorithm can be better than $k$-competitive.

Proof: let $A$ be a deterministic algorithm. Take a universe of size $k + 1$. Define your sequence inductively: at time $t$, given the sequence so far, consider the current set of the cache, and request next the one item that is not in the cache. This defines a sequence.

$A$ has a page fault on every request.

Partition the sequence into phases as usual. Each phase has length at least $k$ and in each phase $OPT$, which is LFD, has only one page fault, hence the bound.

**Theorem 2** Let $y$ be a distribution on input sequences $\sigma_y$ such that for every deterministic algorithm $A$, the expectation over $y$ of the number of pages faults of $A$ is at least $\alpha E(OPT(\sigma_y)) + c$. Then no randomized marking algorithm can be better than $\alpha$-competitive.

Let $R$ be a randomized algorithm. $R$ is nothing else than a distribution $x$ over deterministic algorithms $A_x$. For each $A_x$, $E_y(cost(A_x, \sigma_y)) \geq \alpha E_y(OPT(\sigma_y)) + c$

so that’s also true on average over $x$: $E_x E_y(cost(A_x, \sigma_y)) \geq \alpha E_y(OPT(\sigma_y)) + c$

By Fubini’s theorem we can exchange the order of expectations and write

$E_y E_x(cost(A_x, \sigma_y)) \geq \alpha E_y(OPT(\sigma_y)) + c$

$E_y(E_x(cost(A_x, \sigma_y)) - \alpha E_y(OPT(\sigma_y)) - c) \geq 0$

By the probabilistic method, there exists a $y$ such that

$E_x(cost(A_x, \sigma_y)) - \alpha OPT(\sigma_y) - c \geq 0$

and therefore algorithm $R$ has competitive ratio at least $\alpha$.

**Theorem 3** No randomized algorithm can be better than $H_k$-competitive.

Here is a distribution over sequences: the universe has size $k + 1$, and at each step we pick a page uniformly at random among the $k + 1$ pages.

Let $A$ be a deterministic algorithm. the expected number of faults of $A$, averaged over $\sigma_y$’s, is $m/(k + 1)$ because at each step there is a page fault with probability $1/(k + 1)$.

As for $OPT$, its number of faults is the number of phases in the sequence. In expectation, by independence (and finite expectation) and the elementary renewal theorem, that’s $m$ divided by the expected length of a phase. How long does a phase take? As long as it takes to see every page at least once. That’s a coupon collector problem and it is easy to see (and well-known) that it’s $(k + 1)H_{k+1}$.

Hence the ratio.