Online paging

Theorem 1: Competitive ratio of any deterministic online paging algorithm is at least k. → cache size.

Proof: Let A be a deterministic online paging algorithm. Consider pages $1, \dots, k+1$. Let δ_i be the page in $\{1, \dots, k+1\}$ that is missing in A's cache after processing $\delta_1, \dots, \delta_{i-1}$. The number of misses that A suffers is n (i.e., length of the request stream).

Consider the algorithm (not online) $\underbrace{\text{Furthest-in-the-future}}_{\text{FITF}}$. Whenever a page not in the cache is requested, the page whose next (future) request is latest. Observe that FITF incurs at most one miss in any k consecutive requests. Assume that in a step, page l is brought in and page i is evicted from the cache. Before this step, the content of the cache is $\{1, \dots, k+1\} \setminus \{l\}$.

All this k pages will be requested before i. After i is evicted, there $k-1$ pages in the cache which will be requested before i. So, no miss for next $k-1$ steps.

$\therefore$ The total number of misses by FITF is at most $\lceil \frac{n}{k} \rceil$.

$\therefore$ The competitive ratio of our algorithm is at least k.

$$c(A,\sigma) = n, \quad OPT(\sigma) \leq \lceil \frac{n}{k} \rceil \leq \frac{n}{k} + 1.$$

Let α be the competitive ratio of A. Hence, there exists a constant c s.t.

$$c(A,\phi) \leq \alpha \cdot OPT(\phi) + c \quad \forall \text{ seq. } \phi.$$

$$\Rightarrow c(A,\phi) \leq \alpha \cdot OPT(\sigma) + c$$

$$n = c(A,\phi) \leq \alpha \cdot \left(\frac{n}{k} + 1\right) + c$$

$$\Rightarrow n\left(1-\frac{\alpha}{k}\right) \leq c. \quad -①$$

If $\alpha < k$, $1-\frac{\alpha}{k} > 0$ and ① is false for large enough n.

$$\Rightarrow \alpha \geq k.$$

<u>Conservative algorithms</u>: An algorithm is <u>conservative</u> if for every request sequence with at most k distinct requests, the number of misses is at most k. Example: LRU, FIFO.

<u>Exercise</u>: Check that LRU & FIFO are conservative algorithms.

<u>Theorem 2</u>: Any conservative algorithm is k-competitive.

<u>Proof</u>: Let $\sigma = \sigma_1 \sigma_2 \cdots\cdots$ be some request sequence. We will partition σ into phases.

Phase 1: $\sigma_1, \dots, \sigma_i$ where i is the maximum index for which $\sigma_1, \dots, \sigma_i$ has at most k distinct pages.

Phase 2: A maximal seq. of consecutive requests starting from σ_{i+1} which has at most k distinct pages.

In each phase, the algorithm incurs at most k misses (from the def$^{\underline{n}}$ of conservative algorithm).

Let $a = a_1 a_2 \dots$ be an optimum answer sequence. A shifted phase starts from the second distinct request of a phase and ends just before the second distinct request of the next phase.

$k = 4$ phase 1 phase 2

1 1 3 1 1 2 4 2 3 5 2 7 5 7 2 10 10 5 8 8 5

shifted phase 1 shifted phase 2

We claim that a must incur 1 miss in every shifted phase.

Let p be the first request in this page. Hence, p is present in the cache in the beginning of this shifted phase. There are $k-1$ distinct requests in this phase which are different from p. Also, the first request of the next phase is different from p. Thus, in this shifted phase we will see k requests which are different from p. Since the number of pages different from p that are in the cache in the beginning of the shifted phase is $k-1$, therefore a incurs at least one miss in this shifted phase.

Thus in any sequence with ℓ phases,

→ a conservative algorithm incurs at most ℓk misses

→ Any answer sequence incurs at least $\ell-1$ shifted phases.

$\Rightarrow$ Competitive ratio of a conservative algorithm is at most k. ∎