The Paging Algorithm

An Overview

Xiaoxu Guo

March 2, 2021

Deterministic Algorithms i

In **sleator1985amortized**, Sleator and Tarjan considered the **Paging** problem.

In **Paging**, we have a set C (the cache) and a parameter $k \in \mathbb{N}^+$ (the capacity of the cache).

A paging algorithm receives a sequence of accesses $\sigma = \sigma_1, \sigma_2, \ldots \sigma_n$. The elements of the set C and σ are pages.

When σ_i is received,

- If $\sigma_i \in C$, there is no cost.
- If $\sigma_i \notin C$ (a page fault), there is some cost to add σ to C. Also, if $|C| = k$, one of the pages currently in C has to be discarded. In the basic Sleator and Tarjan model, the cost of a page fault is a constant (say page fault cost $= 1$).

Deterministic Algorithms ii

Let $\mathrm{cost}_{\mathcal{A},k}(\sigma)$ be the number of page faults for the algorithm $\mathcal{A}$, the parameter k, and the sequence σ. And OPT is the optimal offline paging algorithm.

sleator1985amortized showed that if $\mathcal{A}$ is an online deterministic algorithm,

$$\mathrm{cost}_{\mathcal{A},k_{\mathrm{ON}}}(\sigma) \geq \frac{k_{\mathrm{ON}}}{k_{\mathrm{ON}} - k_{\mathrm{OPT}} + 1} \cdot \mathrm{cost}_{\mathrm{OPT},k_{\mathrm{OPT}}}(\sigma).$$

where $k_{\mathrm{ON}} \geq k_{\mathrm{OPT}}$.

Finally, they showed that if $\mathcal{A}$ is FIFO (First in First out) or LRU (Least Recently Used),

$$\mathrm{cost}_{\mathcal{A},k_{\mathrm{ON}}}(\sigma) \leq \frac{k_{\mathrm{ON}}}{k_{\mathrm{ON}} - k_{\mathrm{OPT}} + 1} \cdot \mathrm{cost}_{\mathrm{OPT},k_{\mathrm{OPT}}}(\sigma) + k.$$

Randomized Algorithms i

Fiat et al (**fiat1991competitive**) introduced a randomized paging algorithm, RMA (Randomized Marking Algorithm).

RMA works as follows:

- Initially, all pages are **unmarked**.
- When a page p is accessed, it is **marked**.
- When the cache is full, an uniformly chosen **unmarked** page is discarded. If all pages in the cache are **marked**, unmark all.

Randomized Algorithms ii

They showed that

$$\mathbb{E}[\mathrm{cost}_{\mathrm{RMA},k}(\sigma)] \leq 2H_k \cdot \mathrm{cost}_{\mathrm{OPT},k}(\sigma) + O(1)$$

where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k} \approx \ln k$.

And, for any randomized paging algorithm $\mathcal{A}$, the competitive ratio

$$c_{\mathcal{A},k} = \limsup \frac{\mathbb{E}[\mathrm{cost}_{\mathcal{A},k}(\sigma)]}{\mathrm{cost}_{\mathrm{OPT},k}(\sigma)} \geq H_k$$

McGeoch et al (**mcgeoch1991strongly**) gave a H_k-competitive randomized paging algorithm, matching the lower bound.

Locality of Reference i

From **sleator1985amortized**, we know that $c_{\mathrm{FIFO},k} = c_{\mathrm{LRU},k} = k$. However, in practice, LRU performs much better than FIFO.

To close the gap between theory and practice, researchers seek to model Locality of Reference, which means the pages being accessed consecutively are in some sense related.

One of theses work is **borodin1995competitive**. Borodin et al defined an **access graph** model for paging.

Locality of Reference ii

An **access graph** is a graph $G = \langle V, E \rangle$ where $V = \{1, 2, \ldots, n\}$ is the same as the set of pages. And a valid sequence of accesses σ holds that $\{\sigma_i, \sigma_{i+1}\} \in E$ for all i.

The authors showed that if the access graph is a tree, then $c_{\mathrm{LRU},k} = c_k = \min_{\mathcal{A}} c_{\mathcal{A},k}$ where the minimum is taken over all online paging algorithms.

The paper posted an open question asking if $c_{\mathrm{LRU},k} \leq c_{\mathrm{FIFO},k}$ for all access graphs, which is proved by **chrobak1999lru**.

Locality of Reference iii

The paper also proposed an algorithm, FAR. FAR is a marking algorithm, which discarded the unmarked page which is farthest to marked pages in the access graph.

In a subsequent paper **irani1996strongly**, it is proved that $c_{\mathrm{FAR},k} = O(c_k)$.

Locality of Reference iv

Subsequent development consists of two papers. The first one is **fiat1995randomized**.

The paper had two contributions.

1. It proposed an algorithm RAND where $c_{\mathrm{RAND},k} = O(c_k)$ for general access graph.
2. It considered the multi-pointer setting – there are (implictly) m pointers on the access graph and the next accessed page must be a neighbor of one of the pointer. The authors also presented a strongly-competitive algorithm for this setting.

So far, algorithms based on an access graph need the access graph known as prior.

The second paper **fiat1997truly** considered a truly-random setting, where the access graph is not given to the algorithm.

Locality of Reference v

Parameterizing the access sequence with a graph is a big burden. In **panagiotou2006adequate**, the authors used two numbers α and β to characterize the access graph, yielding a competitive bound in terms of these parameters. .

Locality of Reference vi

A setting similar to access graph is concerned in **karlin1992markov**.

Instead of the worst-case, the sequence of accesses is generated by a Markovian random walk in the access graph. The paper considered the average ratio of page faults in this setting.

Locality of Reference vii

Another attempt to model Locality of Reference is by **albers2005paging**.

Instead of an access graph, the Denning working-set model assumes the number of distinct pages in consecutive n pages is at most $f(n)$.

With respect to the working set model and concave functions f, Albers et al show:

- proved a lower bound for the competitive ratio of deterministic online paging algorithms.
- showed that LRU achieved the lower bound, and FIFO is sub-optimal by giving a lower bound and upper bound of the competitive ratio of FIFO.

Comparison i

The following series of papers compared two paging algorithms directly, instead of referring to a single optimal algorithm.

ben1994new proposed Max/Max model, which is to bound

$$\limsup_{n\to\infty} \frac{\max_{|\sigma|=n} \mathrm{cost}_{\mathcal{A},k}(\sigma)}{\max_{|\sigma|=n} \mathrm{cost}_{\mathcal{B},k}(\sigma)}$$

Sadly, in Max/Max model, LRU=FIFO.

Comparison ii

Similarly, **boyar2007relative** proposed the relative worst-order model.

Let $\mathcal{P}(n)$ is the set of permutations of $\{1, 2, \ldots, n\}$. The model is to bound

$$\sup_{\sigma} \frac{\max_{\phi \in \mathcal{P}(|\sigma|)} \mathrm{cost}_{\mathcal{A},k}(\phi(\sigma))}{\max_{\phi \in \mathcal{P}(|\sigma|)} \mathrm{cost}_{\mathcal{B},k}(\phi(\sigma))}.$$

Same, LRU = FIFO in relative worst-order model.

Comparison iii

A recent paper is **angelopoulos2007separation**.

They used a complicated "Average Analysis+ Concave analysis", with respect to which the ratio for LRU iis better than for FIFO.

File Caching i

So far we have deal with unit page sizes and unit cost.

In **young1994k**, Young considered a variant of Paging, where a page fault of the i-th page has a cost $c_i \geq 1$.

The author formulated the generalized version into a special case of k-server, and proposed a deterministic primal-dual based algorithm GreedyDual where

$$\mathrm{cost}_{\mathrm{GreedyDual},k}(\sigma) \leq \frac{k}{k-h+1} \cdot \mathrm{cost}_{\mathrm{OPT},h}(\sigma).$$

They also proposed a new measure called loosely-competitiveness.

File Caching ii

Later, in **young1998online**, **young1998online** considered a more generalized version with arbitrary page sizes.

He presented an algorithm Landlord whose competitive ratio is also $\frac{k}{k-h+1}$.