

页面置换算法

ljt12138

2019 年 4 月 2 日

1 概述

在现代页式虚拟存储计算机，内存被看作是磁盘的缓存，操作系统通过将常用的虚拟内存页放入物理内存页帧中，将不常用虚拟内存页放入磁盘中来提供更大、更自动化的虚拟存储。实现这种机制的方式是缺页中断。当访问的页不在物理内存中时，MMU会触发缺页中断，操作系统根据一定条件将这个虚拟内存页加载到内存中来。如果内存已经没有多余的页帧空闲，就需要将一些页面换出以空出空间。

页面置换算法决定哪一个空闲页被换出。一个更加通用的问题可以被这样描述：给定一台支持 M 个逻辑页面、拥有 N 个页面大小的内存的机器，选择每次的置换页面使得总的置换次数最少。同样的问题也适用于高速缓存和磁盘缓存的情景，许多算法是通用的。

页面置换算法可以分为离线算法（offline）和在线算法（online）。离线算法是指整个访存序列已经预先给定，在这种情况下拥有高效的最优算法；在线算法是指访存序列是动态给出的，只有解决了上一次的置换问题才会给出下一次访存操作。很显然，真实情景下的页面置换算法必须是在线的。一般来说，评价在线算法的优劣可以用其竞争性分析（Competitive analysis）的方法。给定在线算法与离线下的最优答案的比的上界称为在线算法的竞争比（Competitive ratio）。竞争比越接近1，说明算法与最优算法越接近。

2 经典确定性算法

Belady在其 [3]中最早给出了称为MIN算法的最优算法，当访存序列已经预先给定时可以最小化缺页次数。另一种简单的算法FIF (Further-in-future)，其总是换出下一次访问时间最晚的页面。

引理 2.1 *FIF*算法是最优的。

证明 2.1 证明省略，请参考 [1]。

虽然最优算法需要使用未来的信息，但其可以作为分析算法的一个标准。通过一个算法缺页次数与最优算法缺页次数的比可以简单评价一个在线算法的优劣。设一个算法 A 在访存序列 S 上的缺页次数为 $Miss(A, S)$ ，如果对于任意访存序列 S ，都有：

$$Miss(A, S) \leq f(k, h) \times OPT(S) + O(1)$$

我们可以称，算法 A 是 $f(k, h)$ -Competitive的。其中 k 为物理页面个数， h 为总页面数。

2.1 确定性在线算法竞争比的下界

定理 2.1 任何确定性在线算法 C 的竞争比至少是 k 。

证明 2.2 考虑仅有 $k + 1$ 个逻辑页面的长度为 l 的访存序列，每次访问的页面恰好是 C 上一次换出的页面，那么每一次访问都会发生缺页；在最优算法中，每次缺页发生之后，下一次发生缺页一定是在 k 次之后（因为换出的是下一次访问最晚的页面），因而总的缺页次数不会超过 $k + l/k$ 。有：

$$Miss(C, S) \geq k \times OPT(S) + O(1)$$

2.2 标记算法

标记算法 (Marking Algorithm) 是一类置换算法的总称。假设每个在内存中的页面都有一个标记位 M 。初始时所有页面的标记位都为0。当一个页面被引用时，这个页面的标记位被标记为1。如果算法 C 只会将标记位为0的页面置换出去。当所有页面标记位都为1时，就会将所有页的标记位设置为0。

许多根据访问排序的页面置换算法都是标记算法，例如最近未使用算法（NRU）与最近最久未使用算法（LRU）。

引理 2.2 LRU是标记算法。

证明 2.3 考虑使用反证法。如果LRU不是标记算法，换言之在某个时刻其换出了一个未标记的页 p ，并且存在一个标记的页 q 没有被换出。在最后一次执行所有页标记置0后， q 被访问了至少一次， p 没有被访问过，因而 q 上一次访问比 p 要近。根据LRU算法的定义， p 不会被置换出，矛盾。

定理 2.2 任何标记算法都是 k -Competitive的。

证明 2.4 将序列划分为 t 个块 b_1, b_2, \dots, b_t ，每个块 b_i 中仅包含至多 k 种不同的页面，且 b_1 尽可能长，在此前提下 b_2 尽可能长，以此类推。用归纳法很容易证明，在每一块末尾，标记算法总会使得所有块被标记。那么在下一块开头所有块都被清除标记，且这些未标记的块要么被重新访问，要么被换出，这一块内的换出次数不会超过 k 。那么总的换出次数不超过 $t \times k$ 。

接下来我们说明最优算法至少会发生 t 次缺页。考虑每一个块开头的页面 p_1, p_2, \dots, p_t 。如果 p_i 没有发生缺页，说明 p_i 在内存中。根据划分块的定义，上一次 p_i 被引用一定不在块 b_{i-1} 中，因而 b_{i-1} 可以使用的内存页面只有 $k-1$ 个。那么除掉 p_{i-1} 发生的缺页，根据鸽巢原理，还会至少发生一次缺页。所以总的缺页次数至少是 t 。由于：

$$Miss(Marking, S) \leq t \times k \leq k \times OPT(S)$$

任何标记算法都是 k -Competitive的。

其他确定性算法的竞争比结论如下 [5]。

- LRU算法是 k -Competitive的；这是上面结论的平凡推论。
- FIFO, Clock算法是 k -Competitive的；他们都是所谓Conservative algorithms [6]。
- LIFO, LFU算法不是Competitive的，即算法缺页次数和最优解的缺页次数之比可以无限大。

3 随机算法

类似的，我们定义随机算法R是 $f(h, k)$ -Competitive的，如果：

$$E[Miss(R, S)] \leq f(h, k) \times Miss(FIF, S) + O(1)$$

参考文献

- [1] <https://piazza.com/class/i5j09fns17k5x0?cid=1296>
- [2] <https://riptutorial.com/algorithm/topic/8022/online-algorithms>
- [3] Belady, L. A . A study of replacement algorithms for a virtual-storage computer[J]. IBM Systems Journal, 1966, 5(2):78-101.
- [4] Fiat A , Karp R M , Luby M , et al. Competitive paging algorithms[J]. Journal of Algorithms, 1991, 12(4):685-699.
- [5] <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-854j-advanced-algorithms-fall-2005/lecture-notes/n23online.pdf>
- [6] https://en.wikipedia.org/wiki/Page_replacement_algorithm#Conservative_algorithms