竞争性页面置换调研报告

ljt12138

Department of CST

April 18, 2019

页面置换问题

最基本的页面置换问题可以被这样定义：物理页框共有k个，进程可能访问的页面数量是M，其中$M > k$。

每当一个页面被访问时，如果这个页面不在内存中，需要将其换入到内存中来；页框数量不足时，需要换出内存中的页面来获取页框。

- 决定哪一个页面被换出的算法被称为页面置换算法。

最优算法

Belady最早给出了最优的离线算法MIN，其主要思想是总是换出下一次访问最晚的页面。
真实的情景下使用的算法应当是在线的，但离线算法给了我们评价在线算法的标准。设$OPT(S)$为访存序列S最少的置换次数，$A(S)$为在线算法A的置换次数，常用：

$$\frac{A(S)}{OPT(S)}$$

来表示在线算法A在访存序列S上的效率。

竞争性分析

竞争性分析是分析在线算法的常用手段。竞争性分析将在线算法看作二人零和博弈，其中算法是“Defender”，为了让代价最小化；输入是“Hacker”，为了让代价最大化。竞争性分析的基本情景是分析Hacker可以任给输入的情景，即最坏情况。
对于一个算法A，如果对于任意访存序列S，都有：

$$A(S) \leq c \times OPT(S) + O(1)$$

我们可以称，算法A是c-Competitive的，c可以称为算法A的竞争比。

确定性算法竞争比的下界

Theorem 1

任何确定性在线算法C的竞争比至少是k。

Proof.

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，下一次发生缺页至少是在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l/k[1]。 □

[1]我们一般不考虑内存未满时发生的缺页

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，初始时所有页面都是未标记的。未标记的页面会在访问时被标记，当所有页面都被标记时，将所有页面重新设为已标记。标记算法是那些只会换出未标记的页面的算法。
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）。

LRU是标记算法

Lemma 2

*LRU*是标记算法。

Proof.

根据LRU的定义即可证明。 □

标记算法的竞争比

Theorem 3

任何标记算法都是k-*Competitive*的。

Proof.

将序列按顺序划分为t个段$b_1, b_2, \ldots, b_t$，每个段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。我们称这是访存序列的分割。只需要说明以下两个事实：

1. 标记算法在每一段中至多发生k次缺页。
2. 最优算法至少发生t次缺页。

这两点的证明都是简单的。 □

一些经典算法的竞争比

- 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。
- 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。

引入随机性

在双人零和博弈中，引入随机性可以改善最坏情况下的收益。
一个最简单的随机策略RAND就是从内存中随机地选择一页换出。根据wiki中的介绍，IBM的OS/390操作系统中曾使用随机置换算法作为LRU的补充，当LRU性能退化时使用RAND代替。
使用类似之前的方法，可以定义随机算法R是c-Competitive的，如果：

$$\mathbb{E}[R(S)] \leq c \times OPT(S) + O(1)$$

随机标记算法

一种有效的随机置换算法被称为随机标记算法RM。其思路类似于一般的标记算法，但每次需要换出时会从所有未被标记的页面中随机选出一个换出。
Fiat等人证明了RM是$2H_k$-Competitive的，其中H_k是调和级数前k项的和，$H_k = \log k + O(1)$。证明的主要思想在于访存序列分割中，相邻两段之间“新”的访问与OPT和RM缺页次数的关系。

极大极小原理

Theorem 4 (Minimax Principle)

我们可以将随机算法看成在一个确定性算法集合$\mathcal{A}$中以某种分布选出一个算法，设随机变量A为分布p下选出的算法；输入集合是$\mathcal{X}$，随机变量X为分布q下选出的输入，$c(a,x)$为算法a在输入x下的开销，那么：

$$\max_{x\in\mathcal{X}} \mathbb{E}[c(A,x)] \geq \min_{a\in\mathcal{A}} \mathbb{E}[c(a,X)]$$

利用极大极小原理构造可以证明，任何随机置换策略的竞争比不会低于$\lceil \log k \rceil/2$。Fiat还证明了另一个更紧的界H_{k-1}。

竞争性分析的局限性

竞争性分析与实践经验之间的距离是巨大的。

- 尽管LRU的竞争比达到了k，但在实际应用中，其竞争比通常在1到2之间
- 在实践中，LRU几乎总是比FIFO好，然而竞争性分析并不能给出他们之间的任何区别。
- ...

竞争性分析的扩展

为了弥补竞争性分析和实践之间的距离，许多新的模型被提出。

-
- 基于访问图（Access Graph）的模型：程序一段连续的访存行为应当是访问图上一个连续的路径。
- 基于工作集（Working Set）的模型：程序在一段时间内总是会频繁的访问一些内存页面。