Massachusetts Institute of Technology | Lecture 19
6.892: Theory of Parallel Systems | April 29, 1997
Professor Charles E. Leiserson | Scribe: Stuart Schechter, Ching Law

Lecture 19: Cilk Scheduler IV (Recycling Game)

1 Atomic Access Model

In the previous lectures we assumed that work-stealing requests could be satisfied in constant time, regardless of contention. On real machines, this assumpion is not true. In this lecture we show how to deal with this problem.

Our machine is modeled as follows:

- Asynchronous Parallel Computer with P Processors.
- Concurrent accesses to the same data structure are serially queued.
- If $n \geq 1$ accesses, then one is served immediately.
- Rest are queued for later steps
- Adversary chooses which are served. (must choose 1 to be served if $n \geq 1$)

The 'adversary' models a worst-case situation.

2 (P, M) Recycling Game

We now play the (P, M) *recycling game*, that models concurrent work-steal requests.

- The game is played by adversary
- P balls, P bins, M tosses, 1 reservoir
- Initially, all balls are in the reservoir
- For each step:
- 1. Adversary chooses subset of balls in reservoir and tosses each ball into a bin chosen at random
  2. For each bin that contains ≥ 1 ball, the adversary chooses one and moves it back to the reservoir

Let n_t = number of balls in bins at (immediately after) step t

$$\text{Total Delay} D = \sum_{t=1}^{T} n_t$$

The goal of the advesary is to maximize the total delay. We will now prove that, no matter what the adversary does, the total delay cannot be too big.

3 Recycling Game Theorem

No matter how the adversary plays the recycling game, we have

$$E[D] \leq M \ .$$

Moreover, for any $\epsilon > 0$, with probability $\geq 1 - \epsilon$,

$$D = O\left(M + P \lg P + P \lg \frac{1}{\epsilon}\right)$$

Application: Cilk Scheduling Use *work*, *steal*, and *wait* buckets. The recycling game theorem says that if *steal* bucket has \$M, the wait *bucket* has \$D = \$ $O(M + P\lg P + P\lg\frac{1}{\epsilon})$ with probability $\geq 1 - \frac{1}{\epsilon}$. Therefore $T_p = \frac{T_1}{P} + O(T_\infty + \lg P + \lg\frac{1}{\epsilon})$ with probability $\geq 1 - \frac{1}{\epsilon}$.

4 Proof of Recycling Game Theorem

Assume (without loss of generality) that the adversary's strategy is FIFO (first in first out). The assumption is reasonable because the number of balls in a given bin is the same no matter which ball is removed, and where balls are tossed is indpendent of the identity of the balls.

Let δ_r be the number of steps that ball r finishes in a bin. Therefore,

$$D = \sum_{r=1}^{p} \delta_r \ .$$

definition The ith cycle of a ball is the set of steps in which the ball remains in a bin from its ith toss to its removal. The ball's ith delay is the number of steps that it is in the ith cycle.

Let's focus on delay $\delta = \delta_1$ of ball 1 (since all balls are equivalent).

Let m be the number of times that ball 1 is tossed.

Let d_i be the random variable denoting the ith delay of ball 1.

$$\delta = \delta_1 = \sum_{i=1}^{m} d_i$$

We can say the ith cycle of ball 1 is *delayed* by ball r if ball 1 and ball r are in the same bin when ball r is removed. Because of the FIFO strategy, ball 1 is delayed by ball r once or not at all (in the same cycle). Let

$$X_{ir} = \begin{cases} 1 & \text{if the } i\text{th cycle of ball 1 is delayed by ball } r \\ 0 & \text{otherwise} \end{cases} \ .$$

[X_{ir} is a so-called *indicator random variable*, that is, a random variable which is either 0 or 1. Such variables turn our to be extremely useful in probabilistic arguments. The trick is to decompose a nasty problem into a sum of indicator random variables, find the distribution of the indicator variables and then use linearity of the espectation.]

Thus

$$d_i = \sum_{r=2}^{P} X_{ir}$$

and

$$\delta = \sum_{i=1}^{M} \sum_{r=2}^{P} X_{ir}$$

Consider any set

$$S = \{(i, r) \mid \text{the } i\text{th cycle of ball 1 is delayed by ball } r\} \ .$$

We claim that

$$\Pr\left\{ \bigwedge_{(i,r)\in S} (X_{ir} = 1) \right\} \le P^{-|S|} \ .$$

The proof of the claim is by induction on the size of S. Let $S' = S - \{(i_0, r_0)\}$. For the claim to be true, it is sufficient that

$$\Pr\{X_{i_0 r_0} = 1 \mid \bigwedge_{(i,r)\epsilon S'} (X_{ir} = 1)\} \le \frac{1}{P}$$

A priori, the i_0th toss of ball 1 lands in the same bin as r_0 with probability $\frac{1}{P}$ (r is in a bin) or 0 (r is in the reservoir).

Conditioning on events in S' cannot affect this probability:

1. If $i \neq i_0$, then (i, r) tells nothing about where the i_0th toss of ball 1 goes.
2. If $i = i_0$, but $r \neq r_0$, then (i, r) tells nothing about whether ball 1 actually goes into the bin containing r.

Moreover, "collusion" among indicator random variables provides no further information. So the claim holds. [This is an informal argument that we don't know how to formalize.]

Even without the claim (just by linearity of expectation) we can conclude that

$$E[\delta] = \sum_{i=1}^{m} \sum_{r=1}^{P} E[X_{ir}] \le \frac{1}{P}(P-1)(m) \le m \ .$$

By symmetry, $E[\delta_r] \le m_r$, where m_r = number of tosses of ball r

$$E[D] = \sum_{r=1}^{P} E[\delta_r] \le \sum_{r=1}^{P} m_r = M$$

In order to get a high-probability bound, however, we must use the above claim.

$$\Pr\{\delta > \Delta\} \leq \binom{m(P-1)}{\Delta} P^{-\Delta} \ .$$

Since

$$\binom{a}{b} \leq \left(e\frac{a}{b}\right)^b \ ,$$

we have

$$\Pr\{\delta > \Delta\} \leq \left(e\frac{mP}{\Delta P}\right)^\Delta = \left(e\frac{m}{\Delta}\right)^\Delta$$

Choose $\Delta = \max\left(2em, \lg P + \lg \frac{1}{\epsilon}\right)$

$$\Pr\{\delta \geq \Delta\} \leq \left(e\frac{m}{2em}\right)^{\lg P + \lg \frac{1}{\epsilon}} \leq \left(\frac{1}{2}\right)^{\lg P + \lg \frac{1}{\epsilon}} = \frac{\epsilon}{P}$$

Therefore, the probability that delay $\delta = \delta_1$ of ball 1 is larger than $\max(2em, \lg P + \lg \frac{1}{\epsilon})$ is $\leq \frac{\epsilon}{P}$. But all balls are symmetric (although the delays are not independent). Using Boole's inequality:

$$\Pr\{\text{Any ball has delay} \geq \max\left(2em, \lg \frac{1}{\epsilon}\right)\} \leq P\left(\frac{\epsilon}{P}\right) = \epsilon$$

Therefore, with probability $\geq 1 - \epsilon$

$$D = \sum_{r=1}^{P} \delta_r = \sum_{r=1}^{P} \max\left(2em, \lg P + \lg \frac{1}{\epsilon}\right) = \Theta\left(M + P\lg P + P\lg\frac{1}{\epsilon}\right) \ ,$$

where $\sum_{r=1}^{P} m_r = M$