Randomized Online-Algorithms

~~4/16/14~~
4/25/16
15-750

Online Problems

1) The Paging Problem
2) Server Problem
3) Cat/Mouse Games

Paging Prob

N pages in slow memory
K pages in fast memory $K < N$ FM

Request seq $\equiv \sigma = \sigma_1 \sigma_2 \ldots \sigma_m$

Cost model

request	cost	
$\sigma_i \in FM$	0	
$\sigma_i \notin FM$	1	swap in σ_i and evict a page.

On-line Strategy (Deterministic)

LRU ≡ Least Recently Used
(Evict page not used in longest time)

Off-line Strategy (Det)

LFD ≡ Longest Forward Distance
(Evict page not needed for longest time)

Note LRU & LFD are lazy alg.
Eager alg move before needed.

Know Results

Thm Lazy algs suffice

Thm LFD is off-line opt

no pfs

<u>Recall</u> Metric Space $\equiv$

1) Set S

2) Distance measure $d(\, , \,)$

st

$\forall u \in S \quad d(u,u) = 0$

$\forall u,v \in S \quad d(u,v) \geq 0$

$\forall u,v,w \in S \quad d(u,v) + d(v,w) \geq d(u,w)$

(triangle inequality)

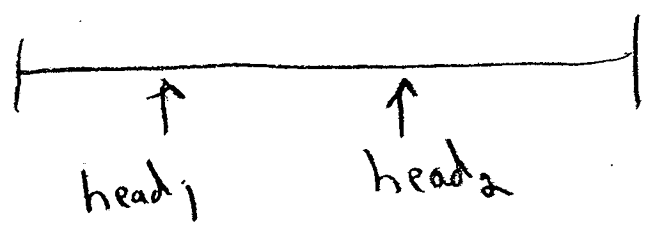
K-server Prob

1) Metric space $|S| \geq K+1$
2) K-servers $\{h_1 \cdots h_K\} = H \subseteq S$
3) Request sequence $\sigma = \sigma_1 \cdots \sigma_m \quad \sigma_i \in S$

Cost Model

request	Cost
$\sigma_i \in H$	0
$\sigma_i \notin H$	Move some server h_j to σ_i Cost $\equiv d(h_j, \sigma_i)$

A 2-server example $\equiv$ 2-headed disk. Prob.



$$d(\text{head}_1, \text{head}_2) = \text{distance}$$

Paging Prob as a K-server Prob

$S \equiv$ pages of slow memory

Fast memory as $\subseteq S$

$$d(u,v) = \begin{cases} 0 & \text{if } u = v \\ 1 & \text{o.w.} \end{cases}$$

Known Thms

Thm $\forall$ K-server prob the competitive factor $\geq K$.

Thm $\forall$ K-server probs $\exists$ 2K-comp. alg

Conj $\forall$ K-server probs $\exists$ K-comp alg.

Back to Paging Prob

The competitive factor for LRU verses LFD.

Consider case $K+1 = N$

Request $\sigma = 1, 2, 3, \cdots, N, 1, 2, 3, \cdots, N, 1, 2, \cdots$

Not: After request $1, \cdots, K$
LRU has a page fault per request.

While
LFD has a page fault every $(K-1)$th request.

	FM	requests	cost
es	$[1, \cdots, N-1]$	$1, 2, \cdots, N-1$	0
	$[1, \cdots, N-2, N]$	N	1
	$[1, \cdots, N-2, N]$	$1, \cdots, N-2$	0
	$[1, \cdots, N-3, N-1, N]$	$N-1$	1
	$[\quad]$	$N, 1, \cdots, N-3$	0
	$[1, \cdots, N-4, N-2, N-1, N]$	$N-2$	1

<u>Thus</u> LRU is at most $(k-1)$-competitive

Goal: Get better cont-factor using randomization.

<u>Def</u> A randomized alg A is <u>c-competitive</u> if $\exists$ constant a $\forall$ request seq τ $\forall$ alg B off-line

$$\text{Expect}\left[C_A(\tau)\right] \leq c \cdot C_B(\tau) + a$$

<u>We will show</u>

$\exists$ paging alg (Randomized) that is $O(\log N)$-Competitive.

Yet Another online Prob.

The Cat/Mouse Game

1) 1 Cat & 1 Mouse
2) N-hiding places

Cat ≡ Seq of probes looking for mouse

$$\text{Cost} \equiv \begin{cases} 1 & \text{is mouse found} \\ 0 & \text{o.w.} \end{cases}$$

Note Cat/Mouse just Paging with $K+1 = N$

Question: Find a good randomized strategy for mouse

First Try

RAND: If found move to random new home.

Claim RAND not good!

Suppose

Cat visits: 1, 2, ..., N-1, 1, 2, ... , N-1, 1

She does not probe N!

Opt off-line: Mouse moves to N

Total cost is 1.

RAND: 1) If at N cost is 0

2) If not at N

What is expected # of moves to land at N?

This is the same as:

Expect # of rolls of an N-sides die

to get, say, N.

Let $E \equiv$ The expected # of rolls

E satisfies recurance:

$$E = \frac{1}{N}(1) + \left(\frac{N-1}{N}\right)(1+E)$$

$$= 1 + \left(\frac{N-1}{N}\right)E$$

$$\frac{1}{N}E = 1 \Rightarrow E = N$$

Back to RAND

$\text{Expect}[\text{RAND}] = N$

Thus RAND is $\Omega(N)$-competitive!

Claim All alg are $\Omega(\log N)$-Competitive.

Pf

Cats Alg: Probe randomly for t times

where $t = N \log N$.

On-line alg: Expect cost $\equiv t/N$

thus $= N\log N / N = \log N$

Off-line $\equiv$ Looking into future for a place to hide!

Question: # of probes for cat to inspect every square?

Let X be a random variable = # probes.

Let $P_i \equiv$ Prob of seeing a new sq after seeing i squares.

Thus $P_i = \frac{N-i}{N}$

let X_i = Random variable # of probes to see a new sq after seeing i sq.

Note $X = \sum_{i=0}^{N-1} X_i$ & $E(X_i) = \frac{N}{N-i}$

Thus

$$E(X) = \sum_{i=0}^{N-1} E(X_i) = \sum_{i=0}^{N-1} \frac{N}{N-i} = N \sum_{i=1}^{N} 1/i$$

$$= \Theta(N \log N)$$

Expect cost of off-line = $O(1)$

$\therefore$ $\Omega(\log N)$ - Competitive

A Better on-line Alg

MARKING:
1) Start at random place.
2) Mark each probed place.
3) When found move to random unmarked place.
4) When all places marked unmark and restart.

Claim Marking is $O(\log N)$-competitive.

pf

Def Phase $\equiv$ time from a restart to a restart.

2 Types of probes

1) probing marked spot (no cost)

2) probing unmarked spot

Since Cat knows your strategy no type 1 probes.

Let M = # of moves per phase.
(random variable)

$$M_i = \begin{cases} 1 & \text{if found at probe } i. \\ 0 & \text{o.w.} \end{cases}$$

$$M = \sum M_i \quad \& \quad E(M_i) = 1/N-i+1$$

Thus $E(M) = \sum E(M_i) = \Theta(\log N)$

Since every place probed per phase

$Opt \geq 1$

Thus Marking is $O(\log N)$ - Competitive

Marking for Pages

Init $\equiv$ no marks

1) Mark each requested page.
2) Eject a random unmarked page.
3) When all page in fast memory are marked restart.

Known: This alg is $2H_K$- Competitive