EECS 495: Randomized Algorithms **Lecture 14**
Random Walks

Reading: Motwani-Raghavan Chapter 6

[[*Powerful tool for sampling complicated distributions since use only local moves to explore state space.*]]

2SAT

Given 2SAT formula,

- Fix satisfying assignment A
- Pick random unsatisfied clause, and flip one of its vars at random
- Let $f(k)$ be expected time to get all n variables to match A if k currently match
  - $f(n) = 0$, $f(0) = 1 + f(1)$
  - $f(k) = 1 + \frac{1}{2}(f(k+1) + f(k-1))$
  - Rewrite: $f(0)-f(1) = 1$ and $f(k)-f(k+1) = 2 + f(k-1) - f(k)$
  - Conclude: $f(k) - f(k+1) = 2k+1$ so $f(0) = \sum_{k=0}^{n-1}(f(k) - f(k+1)) = 1 + 3 + \ldots + (2n-1) = n^2$

    [[*recall geometric argument*
- Find with probability 1/2 in time $2n^2$ by Markov
- Find whp in $O(n^2 \log n)$ time

[[*Intrepret as walk on a line.*]]

Markov Chain

Given:

- state space S
- initial distribution of states
- matrix P of transition probabilities p_{ij} for $i, j \in S$ as prob. transition from i to j

[[*Interpret as directed graph. Compare to 2SAT example.*]]

Note: Properties:

- $\sum_j p_{ij} = 1$
- *memoryless*:

  $\Pr[X_{t+1} = j | X_0 = i_0, \ldots, X_{t-1} = i_{t-1}, X_t = i] = \Pr[X_{t+}$

  where X_t is rand var of state at time t
- If X_t has dist q (q_i is prob of state i), then X_{t+1} has dist qP
- $\Pr[X_{t+r} = j | X_t = i] = P^r_{ij}$

Def:]] The *stationary distribution* is a π s.t. $\pi P = \pi$ (i.e., left eigenvector with eigenvalue 1).

[[*Stationary distribution is sample from state space, so to sample from a set of objects, define chain with correct stationary dist. When does this work?*]]

Question: Stationary distributions for

- 2-cycle? no stationary dist.
- disconnected graph? multiple stationary dist.

Def: A Markov chain is *irreducible* if any state can reach any other state. I.e.,

- path between any two states
- single strong component

Persistent/Transient states:

- r_{ij}^t is prob. first hit j at t given start in state i
- f_{ij} is prob. eventually reach j from i, so $\sum_t r_{ij}^t$
- expected time is hitting time

$$h_{ij} = \begin{cases} \sum_t t r_{ij}^t & : f_{ij} = 1 \\ \infty & : f_{ij} < 1 \end{cases}$$

Def: If $f_{ii} < 1$, state i is *transient*, else *persistent*. If $h_{ii} = \infty$, *null persistent*, else *non-null persistent*.

Note: In finite irreducible chain, all states non-null persistent.

Periodicity:

- max T s.t. state only has non-zero prob. at times $a + Ti$ for integer i
- chain *aperiodic* if no state has periodicity more than 1

Example: bipartite graph periodic, graph with self loops aperiodic

Def: A state is *ergodic* if it is aperiodic and non-null persistent.

[[*Chance of being in state at any sufficiently far future time.*]]

Claim: Fundamental theorem of Markov chains: Any irreducible, finite, aperiodic Markov chain satisfies:

- All states ergodic

  Since finite irreducible implies non-null persistent.
- unique stationary distribution π with $\pi_i > 0$ for all i
- $f_{ii} = 1$ and $h_{ii} = 1/\pi_i$ for all i

  Since hit every $1/h_{ii}$ steps on average.
- number of times visit i in t steps approaches $t\pi_i$ in limit of t

  From linearity of expectation.

Random Walks

Markov chains on (connected, non-bipartite) undirected graphs:

- states are vertices $\{u\}$ with degrees $d(u)$
- move to uniformly chosen neighbor so $p_{uv} = 1/d(u)$ for every neighbor v of u

Claim: unique stationary dist.: $\pi_v = d(v)/2m$

Proof: System of equations:

$$\pi_v = \sum_u \pi_u P_{uv}$$

and

$$\sum_u \pi_u = 1$$

has soln as stated, unique by fundamental theorem of markov chains.

Claim: $h_{vv} = 1/\pi_v = 2m/d(v)$

Def: *Commute time* is $h_{uv} + h_{vu}$.

Def: *Cover time* is $\max_u C_u(G)$ where $C_u(G)$ is expected length of random walk that starts at u and ends after visiting each vertex once.

Question: What do you expect to have bigger commute/cover times?

- clique, line, lollipop

Note for clique, like coupon collector, commute $O(n)$, cover $O(n \log n)$.

Note: Adding edges can increase cover time though improves connectivity!

Claim: For edge (u, v), $h_{uv} + h_{vu} \leq 2m$.

Proof: Define new MC:

- $2m$ states: pair of edge, direction – edge most recently traversed, direction traversed in
- transitions $Q_{(u,v),(v,w)} = P_{vw} = 1/d(v)$

Note Q is *doubly stochastic*:

- col/row sums are 1 since $d(v)$ edges transit to (v, w) each with prob. $1/d(v)$

so uniform stationary dist $\pi_e = 1/2m$, so $h_{ee} = 2m$. Hence in original chain:

- if arrived via (u, v), will traverse (u, v) again in $2m$ steps
- conditioned on arrival edge, commute time $2m$
- memoryless, so can remove conditioning

Note: Bound is for an *edge* of chain.

Claim: Cover time $O(mn)$.

Proof: Consider dfs of spanning tree:

- gives order on vertices
- time for two adj. vertices to be visited in this order $O(m)$ by bound on commute times
- total time $O(mn)$

Claim: Tighter analysis $C_{uv} = 2mR_{uv}$ where R_{uv} is effective resistance in electrical network of graph with 1 unit of resistance on each edge.

Claim: Kirchhoff's Law: conservation of current

Claim: Ohm's Law: voltage across resistance equals product of resistance and current

Def: *Effective resistance* between u and v is voltage diff when one ampere injected into u and removed from v.

Proof: (of claim):

- put $d(x)$ amperes into every x, remove $2m$ from v
- ϕ_{uv} voltage at u w.r.t. v
- Ohm: current from u to neighbor w is $\phi_{uv} - \phi_{wv}$
- Kirchoff: $d(u) = \sum_{w \in N(u)} \phi_{uv} - \phi_{wv} = d(u)\phi_{uv} - \sum \phi_{wv}$
- Also $h_{uv} = \sum (1/d(u))(1 + h_{wv})$ so $d(u)h_{uv} = d(u) + \sum_w h_{wv}$ so $\phi_{uv} = h_{uv}$
- $h_{vu} = \phi_{vu}$ when insert $2m$ at u and remove $d(x)$ from every x
- $h_{uv} + h_{vu}$ is voltage diff. when insert $2m$ at u and remove at v

Result follows from Ohm's law.

Corollary 0.1 *Effective resistance at most shortest path, so $C_{uv} \leq n^3$ for any connected graph.*

$\left[\left[\text{A drunk man gets home visiting every bar in town in time } n^3.\right]\right]$

Example:

- *line graph:* $h_{0n} = h_{n0}$ *and* $h_{0n} + h_{n0} = 2mR_{0n} = 2n^2$, *so* $h_{0n} = n^2$.
- *lollipop:* $h_{uv} + h_{vu} = 2\Theta(n^2)\Theta(n) = \Theta(n^3)$ *and from line,* $h_{uv} = \Theta(n^2)$ *so* $h_{vu} = \Theta(n^3)$ *(so extra factor* n *is "latency" getting started on line).*

Hitting/cover times not monotonic w.r.t. adding edges: line to lollipop to clique.

Applications

Randomized st-connectivity

In log-space

- *walk randomly for* $O(n^3)$ *steps*
- *need to store, current vertex, destination vertex, number of steps*

Note: *In deterministic log-space by Reingold (STOC'05)! Uses ideas from derandomization, e.g., expanders.*

$\left[\left[\text{Same year, best student paper of Vladimir Trifonov did deterministic } O(\log n \log\log n) \text{ space, now at UIC.}\right]\right]$

Card shuffling

Random transposition: Pick two cards i *and* j *and switch.*

- *irreducible? yes, any perm is product of transpositions*
- *aperiodic? yes, self loops*
- *also reversible, i.e.,* $P_{xy} = P_{yx}$ *so doubly stochastic so stationary dist. is uniform*

shuffle cards if repeat enough.

Top-to-random: Take top card, insert at random place.

- *irreducible and aperiodic*
- *not reversible, but each perm has in-degree* n *and out-degree* n *so doubly stochastic and* π *is uniform*

shuffle cards if repeat enough.

Riffle shuffle:

- *split deck into two parts using binomial dist.*
- *drop cards in sequence where card comes from left hand w/prob.* $\frac{|L|}{|L|+|R|}$ *(random interleave)*

Also has stationary dist.

Key issue is mixing time (time to get close to stationary dist from any starting state):

- *top-to-random has* $O(n \log n)$ *mixing time*
- *riffle has* $O(\log n)$ *mixing time – seven shuffles theorem*

Technique coupling: particles move on chain and if ever appear together, then get joined forevermore

- *pair of processes* (X_t, Y_t)

- *each of* $(X_t, \cdot)$ *and* $(\cdot, Y_t)$ *look like MC* ($\Pr[X_{t+1} = j | X_t = i] = P_{ij}$)

- *if* $X_t = Y_t$ *then* $X_{t+1} = Y_{t+1}$

Mixing time is related to time to couple:

- *max dist to stationary at time* t *is* $\Delta(t)$

- *max dist between dist starting at* x *and* y *is* $||p_x^t - p_y^t||$

- *this is at most prob.* X_t *and* Y_t *haven't coupled given start at* x *and* y

Example: *top-in-at-random:*

Define reverse chain:

pick card c *uniformly at random and move to top*

same mixing time.

Coupling: X_t *and* Y_t *pick same* c *(which may be at different positions)*

Fact: *Once card chosen in coupling, always in same position in both decks.*

So coupling/mixing time $O(n \log n)$ *by coupon collector.*

Sampling colorings

Markov chain: pick vertex and color at random and recolor if legal.

- *symmetric, aperiodic, irreducible if at least* $\Delta + 2$ *colors*

Important conjectures:

1. *Random sampling polytime whenever* $q \geq \Delta + 1$
2. *Above chain mixing time* $O(n \log n)$ *whenever at least* $\Delta + 2$ *colors*

Claim: *Mixing time* $O(n \log n)$ *if at least* $4\Delta + 1$ *colors.*

Proof: *Coupling is both chains pick same vertex* v *and color* c. *Let* d_t *be number of vertices where disagree,* q *be number colors.*

- *Good moves: color of* v *disagrees,* c *legal in both graphs. at least* $d_t(q - 2\Delta)$ *good moves.*

- *Bad moves: chosen vertex doesn't disagree, but neighbors disagreeing vertex* v' *and color* c *is color of* v' *in one of graphs and not other. at most* $2d_t\Delta$ *bad moves*

- *Neutral moves: everything else*

Diff between good and bad at least $d_t(q - 4\Delta)$, *so expect distance to decrease when* $q \geq 4\Delta + 1$.

Counting perfect matchings