CSE 6512 Lecture 14 Notes

Asa Thibodeau

October 13, 2011

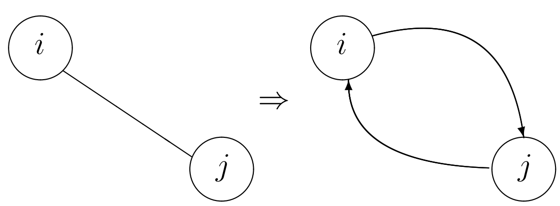
Let $G(V,E)$ be an undirected connected and non-bipartite graph on which we are interested in performing a random walk.

Definition.

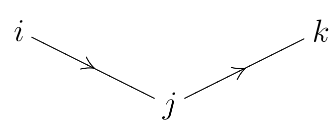
- h_{ij} is the expected time to visit node j starting from node i.
- Commute time between i and j is $h_{ij} + h_{ji}$.
- Let $C_i(G)$ be the expected time to visit each node at least once starting from node i.
- Cover Time, $C(G) = \max_{i \in V} C_i(G)$.

Lemma. *For any edge (i,j)* $\in E$*,* $h_{ij} + h_{ji} \leq 2m$ *where* $m = |E|$.

Proof. Let G' be a directed graph obtained from G by replacing every edge of G with two directed edges:



Construct a Markov Chain $M_{G'}$ where each edge of G' is a state.



$Q_{<i,j><j,k>} = \frac{1}{d_j}$, where d_j is the degree of node j in G.

$$Q_{<a,b><c,d>} = \begin{cases} \frac{1}{d_b} & \text{if } b = c \\ 0 & \text{otherwise} \end{cases}$$

Fact. Q is Doubly Stochastic, i.e., The sum of any row is 1 & the sum of any column is 1.

To see this, consider $column_{<j,k>}$

$$\text{Column Sum} = \sum_{\forall x,y \in V} Q_{<x,y><j,k>} = \sum_{<i,j> \in E'} Q_{<i,j><j,k>} = \sum_{<i,j> \in E'} \frac{1}{d_j} = 1.$$

Fact. If the transition probability matrix of a Markov Chain is doubly stochastic, then its stationary distribution is uniform, i.e., all the entries are the same.

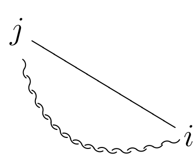
Note. # of states = 2m.

$\Rightarrow \Pi = (\frac{1}{2m}, \frac{1}{2m}, ...)$

$\Rightarrow$ The expected time to traverse the same edge $< j, i >$ twice is $2m$.

$\Rightarrow$ If we traverse the edge $< j, i >$ at some point in time, the expected time to traverse it again is $2m$.

Assume that the node i is reached via the edge $< j, i >$.



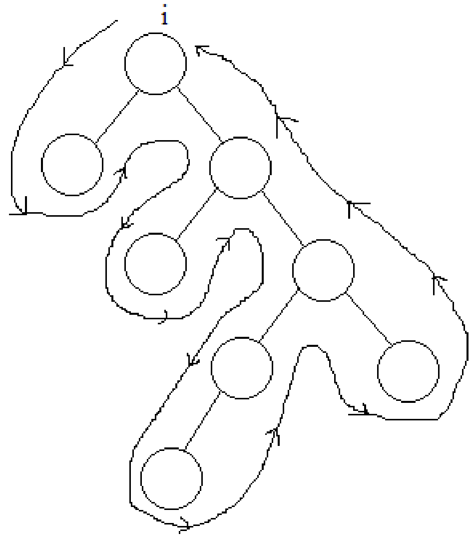
The Markov Chain is memoryless. Also, the node i could have been reached via some other edge as well.

$\Rightarrow$ The time from i to j to i is $\leq 2m$.

□

Lemma. $C(G) \leq 2m(n-1)$.

Proof. Let T be any spanning tree for G.
Consider the Euler Path for this tree:



Let this path be $i, i_1, i_2, \ldots, i_{2n-2} = i$

$$C_i(G) \leq h_{ii_1} + h_{i_1 i_2} + h_{i_2 i_3} + \ldots h_{i_{2n-3} i_{2n-2}}$$

$$= \sum_{(i,j) \in T} h_{ij} + h_{ji} = 2m(n-1).$$

$$\Rightarrow C(G) = \max_{i \in V} C_i(G) \leq 2m(n-1).$$

□

Corollary 1. *The expected run time of our 2SAT randomized algorithm is* $O(n^2)$*.*

Corollary 2. $C(G)$ *for* K_n *is* $O(n^3)$*.*

But we have proved a much better bound!

Corollary 3.
Problem: (Reachability)
Input: A Graph $G(V, E)$ and $a, b \in V$.
Question: "Is b reachable from a?"

Fact. Reachability $\in$ RLP.

Definition. A problem is in RLP if $\exists$ a randomized algorithm that uses only $O(\log n)$ work space such that Prob[x is correct] is $\geq \frac{1}{2}$ if $x \in L$ and Prob[x is correct] is 0 if $x \notin L$. Here the problem under concern is that of deciding if x is a member of the language L or not.

Proof. Do a random walk starting from a for $4m(n-1)$ steps.
Output "yes" if b has been seen, and "no" otherwise.
Using Markov's inequality, the probability of an incorrect answer is $\leq \frac{1}{2}$. □

Approximate Matrix Multiplication:

Consider the problem of adding n numbers $x_1, x_2, .., x_n$.
A randomized algorithm for this problem works as follows. Pick a random sample S, with $|S| = s$. Let the sample be $y_1, y_2, ..., y_s$.

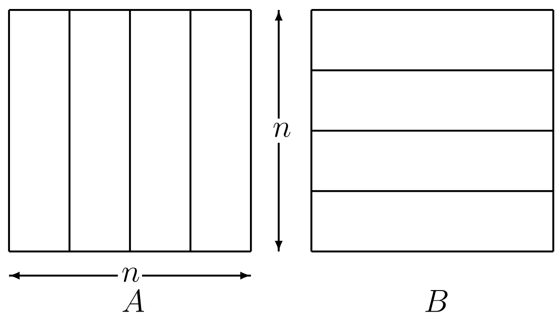
Output: $\frac{n}{s}\sum_{i=1}^{s} y_i$.

Analysis.

Expected value of a random element $= \frac{1}{n}\sum x_i$.

$$E[\sum_{i=1}^{s} y_i] = \frac{s}{n}\sum_{i=1}^{n} x_i. \Rightarrow E[\frac{n}{s}\sum_{i=1}^{s} y_i] = \sum_{i=1}^{n} x_i.$$

Approximate Matrix Multiplication (R. Kannan, et. al. 2005)



Pick s columns randomly from A;
Use the same rows in B.

Perform the multiplication of A & B only with respect to these columns & rows. Finally, multiply each element in the product by $\frac{n}{s}$.

Clearly, the answer output by this algorithm is correct in expectation!