Let $G$ be an undirected graph, $G = (V, E)$, $V = \{1, \ldots, n\}$. Let $d(i)$ be the degree of vertex $i$. Let $m = |E| = \#\text{of edges}$.

A random walk starts at vertex $i_0$. When at vertex $i$, transitions to a randomly chosen neighbor $j$ of $i$.

Transition probability:

$$p_{ij} = \begin{cases} \frac{d(i)}{2m} & \text{if } (i, j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Transition matrix $P = (p_{ij})$ is an $n \times n$ matrix.

$P$ has row sums equal to 1, a stochastic matrix.

**Definition:** The hitting time (also known as the first passage time) $h_{ij}$ is the expected time for a random walk starting at vertex $i$ to reach vertex $j$.

$C_i := \text{expected time for a random walk starting at } i \text{ to reach every vertex of } G \text{ at least once}$.

The cover time $C := \max_i C_i$.

The commute time $C_{ij} := h_{ij} + h_{ji}$ is the expected time for random walk, starting at vertex $i$, to reach vertex $j$ and return to vertex $i$.

We assume $G$ is connected so these values are well-defined and finite.

**Goal:** Give good upper bounds on these quantities.
A probability distribution $\pi$ is a function $\pi : [n] \to [0, 1]$ s.t. $\sum \pi(i) = 1$.

A stationary distribution is a distribution $\pi$ s.t. $\pi \cdot P = \pi$.

Intuition: If $\tau$ is the distribution of position vertex $i$ at time $t$,
then $\tau \cdot P = \tau$ at time $t+1$.

**Defn:** Let $\pi(i) = \frac{d(i)}{2m}$.

Then $\pi = (\pi(0), \ldots, \pi(n))$ is the (unique) stationary distribution.

**Proof** $\pi$ is stationary:

$$(\pi \cdot P)_i = \sum_j P_{ij} \pi_j = \sum_j \sum_{j \in N(i)} \frac{d(j)}{2m} \frac{i}{2m} = \sum_{j \in N(i)} \frac{d(i)}{2m} = \pi_i,$$

where $N(i) = \{j : (i,j) \in E\}$.

6. $\sum \pi_i = \sum \frac{d(i)}{2m} = \frac{\text{(# of edges out of $i$)}}{2m} = \frac{2m}{2m} = 1$.

C. $\pi$ is unique:

Let $\tau$ be another stationary distribution.

Define $r_i = \frac{\tau(i)}{\pi(i)}$. Let $B(V) = \{i : r_i = \max_j r_j\}$. Note $r_i > 1$.

Since $G$ is connected and since $\sum r_i = 1$, there is an $i_0$ s.t. $r_{i_0} = \max_j r_j$ but not all $i_0$'s neighbors satisfy this property.

This gives a contradiction.

qed.
Fundamental Theorem: Let \( \pi \) be the stationary distribution.

(a) \( h_{ii} = \frac{1}{\pi_i} \).

(b) \( \pi_i \) gives the expected frequency of visit to vertex \( i \), i.e.

for any initial distribution \( \pi \), let \( N_{\pi}(i,t) \) = expected \# of visits to vertex \( i \) in \( t \) steps, then

\[
\lim_{t \to \infty} \frac{N_{\pi}(i,t)}{t} = \pi_i.
\]

Proof (sort of)

(a) Let \( H = (h_{ij}) \) an non matrix.

Note \( h_{ij} = 1 \cdot p_{ij} + \sum_{k \neq j} p_{ik} (h_{kj} + 1) \)

Claim: \( H = \pi (H - H_{\text{diag}}) + E \)

where \( H_{\text{diag}} = (h_{ij} \delta_{ij})_{ij} \) - \( \delta_{ij} \) delta function, \( H_{\text{diag}} \)-diagonal

\( E = (1)_{ij} \) all 1's matrix.

Proof: \( h_{ij} = \sum_{k \neq j} p_{ik} (h_{kj} + 1) + p_{ij} = \sum_{k \neq j} p_{ik} h_{kj} + 1 \)

\( = \sum_{k} p_{ik} h_{kj} - p_{ij} h_{jj} + 1 \)

So \( H = \pi H - \pi H_{\text{diag}} + E \) good claim.

Thus \( \pi H = \pi P (H - H_{\text{diag}}) + \pi E \)

\( = \pi P (H - H_{\text{diag}}) + \pi E = \pi (H - H_{\text{diag}}) + \pi \text{I} \)

where \( \text{I} \) = all 1's row vector

So \( \pi H_{\text{diag}} = \pi \text{I} \), i.e.

\( \pi_i \cdot h_{ii} = 1 \), so \( h_{ij} = \frac{\pi_j}{\pi_i} \).
(b): By linearity of expectation.

\[ N_x(i; t) = x + \tau P + \tau P^2 + \tau P^3 + \ldots \]

\[ = \tau (I + P + P^2 + P^3 + \ldots) \]

and \( \lim_{t \to \infty} \tau (I + P + P^2 + \ldots P^{t-1}) \) is a stationary distribution.

and hence equal to \( \tau \). (Omit rest of proof)

Q.E.D.

**Corollary:** The expected frequency of transitions from vertex \( i \) to vertex \( j \) along (directed) edge \((ij)\) equals

\[ T_i \cdot P_j = \frac{d(i)}{2m}, \frac{1}{d(i)} = \frac{1}{2m} \]

Thus all directed edges are traversed with the same expected frequency.

**Lemma:** There is an edge from \( i \) to \( j \),

\[ h_{ij} + h_{ji} \leq 2m \]

**Proof:** By linearity, \( h_{ij} + h_{ji} = \sum \) (expected # of occurrence of transition from \( i \) to \( j \) in a commute from \( i \) to \( j \) or back to \( i \)).

By corollary, and since any infinite random walk is made up of \( i-j-i \) commutes (with probability 1), each term of the summation is equal. Thus

\[ h_{ij} + h_{ji} = 2m (\text{expected # of transitions from } i \text{ to } j \text{ in an } i-j-i \text{ commute}) \]

for all \((i,j) \in E\).

Since \((i'=i, j'=j)\), edge \((ij)\) is traversed at most once

\[ h_{ij} + h_{ji} \leq 2m \]

q.e.d.
Theorem. The cover time $C$ is $\leq 2m(n-1) \leq 2n^2$.

Pf. For $i$ a vertex, need to show $C_i \leq 2m(n-1)$.
Let $T$ be a spanning tree of $G$ with root $i$.
$T$ has $n-1$ edges.
Do a depth-first traversal of $T$ at $i' = i_0, i_1, i_2, \ldots, i_{2n-2}$
when $i_{2n-2}$ is the last child of $i$.

By the lemma, the cover time $C_i$ satisfies

$$C_i \leq T(i_0,i) + T(i_1,i_2) + \ldots + T(i_{2n-3}, i_{2n-2})$$

$$\leq \sum_{(i,j) \in E} (T(i,j) + T(j', i'))$$

(by reading terms twice)

$$\leq (n-1)2m$$

g.e.d.

Defn. RLP = set of languages $L$ for which there is a TM $M$ s.t.
$M$ runs leg. opt and polynomial time and, for all $x$,

$x \in L \Rightarrow M(x)$ accepts with probability $\geq \frac{1}{2}$

$x \notin L \Rightarrow M(x)$ does not accept (i.e., accepts with probability $0$).

[No longer assume $G$ is connected.]

Defn. USTCON is the decision problem

$$(G, s, t) \in \text{USTCON} \Leftrightarrow \text{exists path from } s \text{ to } t \text{ in } G$$

$$\Leftrightarrow s, t \text{ in the same connected component of } G.$$ 

Theorem: USTCON \in RLP.

Pf: Algorithm: Starting from $s$, do random walk for $\frac{4n^3}{\epsilon}$ steps
Accept iff vertex $t$ is encountered.

By Markov inequality, $(s, t, G) \in \text{USTCON} \Rightarrow \Pr[M \text{ accepts}] \geq \frac{1}{2}$
Let $G$ be a $d$-regular graph. A walk on $G$ is specified by its initial vertex and a sequence $o \cdot s_0, \ldots, s_{d-1}$, where $s_i$ is the choice of outgoing edge chosen at step $i$.

Definition: A walk is $n$-universal if for all graphs $G$, all $s, t$ in some connected component, the walk specified by initial vertex $s$ and by $o$ contains $t$.

Intuition: "derandomization"

Theorem: There is an $n$-universal sequence of length $O(n^3 \log n)$

Proof: The cover time for a $d$-regular graph is $2dn(n-1) \leq 2dn^2 - 1$.

The total number of $d$-regular graphs is $\leq \left(\frac{n}{d}\right)^n \leq (\log n)^n = 2^{\log n \cdot \log n}$.

Hence, there is a collection of $\leq 2n \cdot \log n + 1$ many transversal sequences of length $\geq 4dn^2 - 1$.

Every $d$-regular graph and any given choice of initial vertex is covered by at least one transversal sequence in the collection.

The concatenation of these.

Above argument doesn't quite work.

Instead consider a random choice of $dn \cdot \log n + 1$ many transversal sequences and concatenate them.

For a fixed $G$, probability that this does not traverse $G$ is $\leq \left(\frac{1}{2}\right)^{dn \cdot \log n + 1}$.

Since there are $\leq 2dn \cdot \log n$ many $G's$, the probability it does not traverse all of them is $\leq \left(\frac{1}{2}\right)^{dn \cdot \log n + 1} \cdot 2 \leq \frac{1}{2}$.

QED.
Directed graphs

AKLLR show: Similar results for strongly connected, directed graphs which have indegree $d$ and outdegree $d$ at each vertex.

For non-fixed degree, here is an example of how things can fail:

![Graph diagram]

Probability of reaching $t$ from $s$ before returning to $s$ is $2^{-d}$.
Hence random walk needs time $\approx 2^{-d}$ to reach $t$ from $s$.

Worse, the graph might not be strongly connected (hence not ergodic).

For this case consider following algorithm

Input: $G,s,t$ - $n$ vertices

Loop: Starting at $s$, do random walk of length $n$

If encounter $t$, accept.
Else, continue.

This has (worst-case) exponential runtime if $t$ is not reachable from $s$.
If it is not reachable from $s$, it never halts.

To make it halt, want to run in many times than the loop.
Then reject if it is not encountered.

This makes a log space, exponential time algorithm that satisfies condition $\epsilon$ for $n \in \mathbb{N}$.

Problem: Counting to $n^n$ takes space $\log^m n \approx n \log n$, not log space.
Fix! Instead of counting to $n^2$, instead select (and reject) $\log_2(n^2)$ many random bits. If they are all equal to 1, halt (and reject). Else continue.

New algorithm

Loop

Start by $2n$ do random walk of length $n$

If encounter 1, accept.

Flip $n \log_2(n)$ many coins.

If all equal heads, reject.

Else continue.

This does satisfy condition (A):

$x \in L \Rightarrow \Pr[\text{M1x accepts}] \geq \frac{1}{2}$

$x \notin L \Rightarrow \Pr[\text{M1x accepts}] = 0$. 