Logarithmic Space

REACHABILITY Is NL-Complete

- REACHABILITY $\in$ NL (p. 95).
- Suppose L is decided by the $\log n$ space-bounded TM N.
- Given input x, construct in logarithmic space the polynomial-sized configuration graph G of N on input x (see Theorem 21 on p. 176).
- G has a single initial node, call it 1.
- Assume G has a single accepting node n.
- $x \in L$ if and only if the instance of REACHABILITY has a "yes" answer.

2SAT Is NL-Complete

- 2SAT $\in$ NL (p. 265).
- As NL = coNL (p. 191), it suffices to reduce the coNL-complete UNREACHABILITY to 2SAT.
- Start without loss of generality an *acyclic* graph G.
- Identify each edge (x, y) with clause $\neg x \vee y$.
- Add clauses (s) and $(\neg t)$ for the start and target nodes s and t.
- The resulting 2SAT instance is satisfiable if and only if there is no path from s to t in G.

The Class RL

- REACHABILITY is for *directed* graphs.
- It is not known if UNDIRECTED REACHABILITY is in L.
- But it is in randomized logarithmic space, called **RL**.
- RL is RP in which the space bound is logarithmic.
- We shall prove that UNDIRECTED REACHABILITY $\in$ RL.[a]
- As a note, UNDIRECTED REACHABILITY $\in$ coRL.[b]

---

[a] Aleliunas, Karp, Lipton, Lovász, and Rackoff (1979).
[b] Borodin, Cook, Dymond, Ruzzo, and Tompa (1989).

Random Walks

- Let $G = (V, E)$ be an undirected graph with $1, n \in V$.
- Add self-loops $\{ i, i \}$ at each node i.
- The randomized algorithm for testing if there is a path from 1 to n is a **random walk**.

The Random Walk Framework

$x := 1$;
while $x \neq n$ **do**
    Pick y uniformly from x's neighbors (including x);
    $x := y$;
end while

Some Terminology

- v_t is the node visited by the random walk at time t.
- In particular, $v_0 = 1$.
- d_i denotes the degree of i (including the self-loops).
- Let $p_t[\,i\,] = \text{prob}[\,v_t = i\,]$.

A Convergence Result

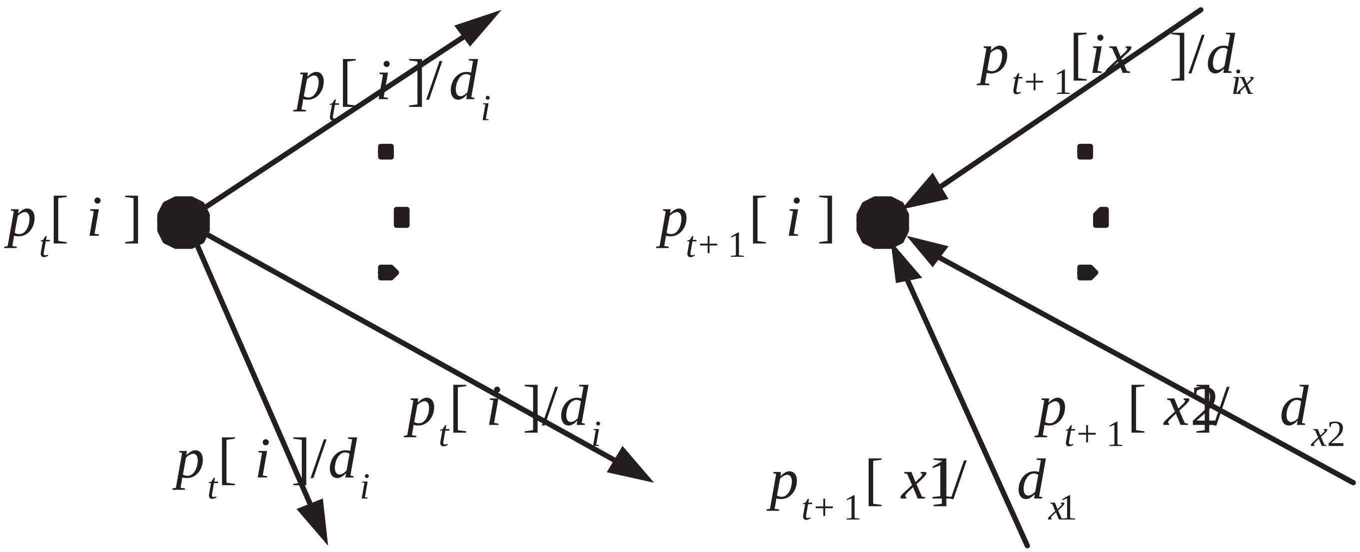
Lemma 102 *If $G = (V, E)$ is connected, then* $\lim_{t \to \infty} p_t[\, i \,] = \frac{d_i}{2 \cdot |\, E \,|}$ *for all nodes i.*

- Here is the intuition.
- The random walk algorithm picks the edges uniformly randomly.
- In the limit, the algorithm will be well "mixed" and forgets about the initial node.
- Then the probability of each node being visited is proportional to its number of incident edges.
- Finally, observe that $\sum_{i=1}^{n} d_i = 2 \cdot |\, E \,|$.

Proof of Lemma 102

- Let $\delta_t[\,i\,] = p_t[\,i\,] - \frac{d_i}{2 \cdot |\,E\,|}$, the deviation.
- Define $\Delta_t = \sum_{i \in V} |\delta_t[\,i\,]|$, the total absolute deviation.
- Now we calculate the $p_{t+1}[\,i\,]$'s from the $p_t[\,i\,]$'s.
- Each node divides its $p_i[\,t\,]$ into d_i equal parts and distributes them to its neighbors.
- Each node adds those portions from its neighbors (including itself) to form $p_i[\,t+1\,]$.

The Flows



Proof of Lemma 102 (continued)

- $p_t[\,i\,] = \delta_t[\,i\,] + \frac{d_i}{2\cdot|\,E\,|}$ by definition.
- Splitting and giving the $\frac{d_i}{2\cdot|\,E\,|}$ part does not affect $p_{t+1}[\,i\,]$ because the same $\frac{1}{2\cdot|\,E\,|}$ is exchanged between any two neighbors.
- So we only consider the splitting of the $\delta_t[\,i\,]$ part.
- The $\delta_t[\,i\,]$'s are exchanged between adjacent nodes.

Proof of Lemma 102 (continued)

- Clearly $\sum_i \delta_{t+1}[\,i\,] = \sum_i \delta_t[\,i\,]$ because of conservation.
- But $\Delta_{t+1} = \sum_i |\delta_{t+1}[\,i\,]| \leq \sum_i |\delta_t[\,i\,]| = \Delta_t$.
  - If $\delta_t[\,i\,]$'s are all of the same sign, then $\Delta_{t+1} = \sum_i |\delta_{t+1}[\,i\,]| = \sum_i |\delta_t[\,i\,]| = \Delta_t$.
  - When $\delta_t[\,i\,]$'s of opposite signs meet at a node, that will reduce $\sum_i |\delta_{t+1}[\,i\,]|$.
- We next quantify the decrease $\Delta_t - \Delta_{t+1}$.

Proof of Lemma 102 (continued)

- There is a node i^+ with $\delta_t[\, i^+ \,] \geq \frac{\Delta_t}{2 \cdot |\, V \,|}$, and there is a node i^- with $\delta_t[\, i^- \,] \leq -\frac{\Delta_t}{2 \cdot |\, V \,|}$.
  - Recall that $\sum_i \delta_t[\, i \,] = 0$ and $\sum_{i \in V} |\delta_t[\, i \,]| = \Delta_t$.
  - So the sum of all $\delta_t[\, i \,] \geq 0$ equals $\Delta_t / 2$.
  - As there are at most $|\, V \,|$ such $\delta_t[\, i \,]$, there must be one with magnitude at least $(\Delta_t / 2) / |\, V \,|$.
  - Similarly for $\delta_t[\, i \,] \leq 0$.

Proof of Lemma 102 (continued)

- There is a path $[\, i_0 = i^+, i_1, i_2, \ldots, i_{2m} = i^- \,]$ with an even number of edges between i^+ and i^-.
  - Add self-loops to make it true.
- The positive deviation $\delta_t[\, i^+ \,]$ from i^+ will travel along this path for m steps, always subdivided by the degree of the current node.
- Similarly for the negative deviation $\delta_t[\, i^- \,]$ from i^-.

Proof of Lemma 102 (continued)

- At least a positive deviation equal to $\frac{1}{|V|^m}$ of the original amount will arrive at the middle node i_m.
- Similarly for a negative deviation from the opposite direction.
- So after $m \leq n$ steps, a positive deviation of at least $\frac{\Delta_t}{2 \cdot |V|^n}$ will cancel an equal amount of negative deviation.
- We do not need to care about cases where numbers of the same sign meet at a node; they will not change Δ_t.

Proof of Lemma 102 (concluded)

- So in n steps the total absolute deviation decreases from Δ_t to at most $\Delta_t(1 - \frac{1}{|V|^n})$.
- But we already knew that Δ_t will never increase.[a]
- So in the limit, $\Delta_t \to 0$ (but exponentially slow).

[a]Contributed by Mr. Chih-Duo Hong (R95922079) on January 11, 2007.

First Return Times

- Lemma 102 (p. 783) and theory of Markov chains[a] imply that the walk returns to i every $2 \cdot | E |/d_i$ steps, *asymptotically and on the average.*
- Equivalently, if $v_t = i$, then the expected time until the walk comes back to i for the first time after t is $2 \cdot | E |/d_i$, asymptotically.
  - This is called the **mean recurrence time**.

[a]Particularly, theory of homogeneous Markov chains on first passage time.

First Return Times (concluded)

- Although the above is an asymptotic statement, the said expected return time is *the same* for any t—including the beginning $t = 0$.
- So from the beginning and onwards, the expected time between two successive visits to node i is exactly $2 \cdot |\, E \,|/d_i$.

Average Time To Reach Target Node n

- Assume there is a path $[\,1, i_1, \ldots, i_m = n\,]$ from 1 to n.
  - If there is none, we are done because the algorithm then returns no false positives.
- Starting from 1, we will return to 1 every expected $2 \cdot |\,E\,|/d_1$ steps.
- Every cycle of leaving and returning uses at least *two* edges of 1.
  - They may be identical.

Average Time To Reach Target Node n (continued)

- So after an expected $\frac{d_1}{2}$ of such returns, the walk will head to i_1.
  - There are d_1^2 pairs of edges incident on node 1 used for the cycles.
  - Among them, d_1 of them leave node 1 by way of i_1 and d_1 of them return by way of i_1.
- The expected number of steps is

$$\frac{d_1}{2} \frac{2 \cdot |\, E \,|}{d_1} = |\, E \,|.$$

Average Time To Reach Target Node n (concluded)

- Repeat the above argument from $i_1, i_2, \ldots$
- After an expected number of $\leq n \cdot |E|$ steps, we will have arrived at node n.
- Markov's inequality (p. 410) suggests that we run the algorithm for $2n \cdot |\, E \,|$ steps to obtain the desired probability of success, 0.5.

Probability To Visit All Nodes

Corollary 103 *With probability at least 0.5, the random walk algorithm visits all nodes in* $2n \cdot |E|$ *steps.*

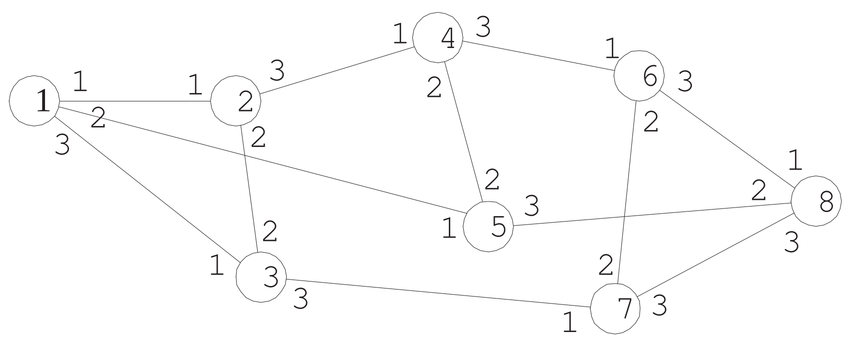
- Repeat the above arguments for this particular path: $[1, 2, \ldots, n]$.

The Complete Algorithm

$x := 1$;
$c := 0$;
while $x \neq n$ and $c < 2n \cdot |\, E \,|$ **do**
    Pick y uniformly from x's neighbors (including x);
    $x := y$;
    $c := c + 1$;
end while
if $x = n$ **then**
    "yes";
else
    "no";
end if

Some Graph-Theoretic Notions

- A d-regular (undirected) graph has degree d for each node.
- Let G be d-regular.
- Each node's incident edge is labeled from 1 to d.
  - An edge is labeled at both ends.



Universal Sequences[a]

- A sequence of numbers between 1 and d results in a walk on the graph if given the starting node.
  - E.g., $(1, 3, 2, 2, 1, 3)$ from node 1.
- A sequence of numbers between 1 and d is called **universal** for d-regular graphs with n nodes if:
  - For *any* labeling of *any* n-node d-regular graph G, and for *any* starting node, all nodes of G are visited.
  - A node may be visited more than once.
- Useful for museum visitors, security guards, etc.

[a] Attributed to Cook.

Existence of Universal Sequences

Theorem 104 *For any n, a universal sequence exists for the set of d-regular connected undirected n-node graphs.*

- Enumerate all the different labelings of d-regular n-node connected graphs and all starting nodes.
- Call them $(G_1, v_1), (G_2, v_2), \ldots$ (finitely many).
- S_1 is a sequence that traverses G_1, starting from v_1.
  - A spanning tree will accomplish this.
- S_2 is a sequence that traverses G_2, starting from the node at which S_1 ends *when applied to* (G_2, v_2).

The Proof (concluded)

- S_3 is a sequence that traverses G_3, starting from the node at which $S_1 S_2$ ends when applied to (G_3, v_3), etc.
- The sequence $S \equiv S_1 S_2 S_3 \cdots$ is universal.
  - Suppose S starts from node v of a labeled d-regular n-node graph G'.
  - Let $(G', v) = (G_k, n_k)$, the kth enumerated pair.
  - By construction, S_k will traverse G' (if not earlier).

A $O(n^3 \log n)$ Bound on Universal Sequences

Theorem 105 *For any n and d, a universal sequence of length $O(n^3 \log n)$ for d-regular n-node connected graphs exists.*

- Fix a d-regular labeled n-node graph G.
- A random walk of length $2n \cdot |\, E \,| = n^2 d = O(n^2)$ fails to traverse G with probability at most $1/2$.
  - By Corollary 103 (p. 797).
  - This holds wherever the walk starts.
- The failure probability for G drops to $2^{-\Theta(n \log n)}$ if the random walk has length $\Theta(n^3 \log n)$.

The Proof (continued)

- There are $2^{O(n \log n)}$ d-regular labeled n-node graphs.
  - Each node has $\leq n^d$ choices of neighbors.
  - So there are $\leq n^{d+1}$ d-regular graphs on nodes $\{1, 2, \ldots, n\}$.
  - Each node's d edges are labeled with unique integers between 1 and d.
  - Hence the count is

$$\leq n^{d+1}(d!)^n = n^{O(n)} = 2^{O(n \log n)}.$$

The Proof (concluded)

- The probability that there exists a d-regular labeled n-node graph that the random walk fails to traverse can be made at most $1/2$.
  - Lengthen the length of the walk suitably.
- Because the probability is less than one, there *exists* a walk that traverses all labeled d-regular graphs.

Finis