So far looked at connectivity-like problems. Other graph problems
distance-based
have different reqs, e.g. symmetry breaking problems.

Eg. Maximal Independent Set

Input: Undirected graph $G(V,E)$

Output: A set $U \subseteq V$ where $\forall u,v \in U$, $(u,v) \notin E$ } Independent

and $\forall v \in V \setminus U$, v has some neighbor in U. } Maximal.

Sequentially very easy to find an MIS.
Just place the vertices in an arbitrary order, add one-by-one. Whenever a vertex is added, remove all neighbors from consideration

$\Rightarrow O(m+n)$ work.

We will see later that using the "right order" above also yields a fast parallel algorithm, although the process is seemingly sequential.

Lot of work (again in classic-era) on MIS. Culminated in a celebrated algorithm by Luby:

Idea is to use a simple random symmetry-breaking procedure to find an Independent set. Then we remove it, and its neighbors, and recurse using fresh randomness.

LubyMIS ($G(V,E)$)

if ($|V| == 0$) return {}

parfor $v \in V$: p_v = random_priority()

$I = \{ v \in V \mid p_v < \min_{u \in N(v)} \{p_u\} \}$ // independent set

$V' = V \setminus (I \cup N(I))$ // remove I and its neighbors

$G' = (V', \{(u,v) \in E \mid u \in V' \text{ and } v \in V'\})$

R = LubyMIS (G')

return $I \cup R$

Correctness

Not hard to see that this algorithm works correctly and returns a valid MIS. Can prove via induction. BC is an empty graph, which has no vertices and thus MIS = $\emptyset$. For the IS, we have that I is independent. Given I, no vertices in $N(I)$ can be in the MIS. Any remaining vertex in V' is independent of I. By induction, the set R is an MIS of $G[V']$, and there are none other from V' we can add while maintaining independence. $\Rightarrow$ $I \cup R$ is independent and maximal.

Cost. Will show: LubyMIS runs in $O(m+n)$ expected work, $O(\log^2 n)$ depth whp.

Cost analysis is trickier.

We will show that each round removes half of the edges in expectation.

Lemma: Each round of LubyMIS removes at least half of the edges in expectation.

Consider $(u,v) \in E$

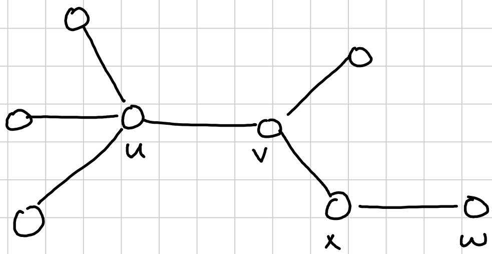
Let $A_{u,v}$ be the event that <u>u</u> has the largest priority among u, v and $N(u)$ and $N(v)$

Since the priorities are picked uniformly at random, we can compute this prob. directly by counting:

$$\Pr[A_{u,v}] \geq \frac{1}{d(u)+d(v)}$$

This is an inequality since u, v might share neighbors.

To avoid double-counting, we will (when adding a vertex u to MIS) focus on counting removal of its neighbors edges; not its own edges).



$Pr[A_{u,v}] = \frac{1}{4+3} = \frac{1}{7}$

if $A_{u,v}$ occurs:

- u will join the MIS
- all edges incident on v will be removed since $v \in N(u)$.
- u itself will be removed

We expect to remove $Pr[A_{u,v}]\, d(v)$ neighbors of v due to $A_{u,v}$.

Can we have multiple events remove edges incident to v on the same round?

Eg. $A_{u,v}$ and $A_{x,v}$ both occur?

Can't happen since $A_{u,v} \Rightarrow p_u$ has highest pri in $N(u) \cup N(v)$ and $x \in N(v)$.

This means if we count edges incident on v removed by $A_{u,v}$, we won't double count due to a simultaneous $A_{w,v}$ event.

What can happen is that an edge, eg. (v,x) gets removed twice, once by $A_{u,v}$ and once by $A_{w,x}$.

Each edge can be double-counted this way at most twice — one from each endpoint.

Let Y = RV giving the #edges we remove.

Claim:

$$E[Y] \geq \underbrace{\frac{1}{2} \sum_{(u,v) \in E}}_{\text{double-counting of removed edges by both endpoints}} \underbrace{\left(\Pr[A_{u,v}]\, d(v) + \Pr[A_{v,u}]\, d(u)\right)}_{\text{removed edges in both directions of } (u,v) \text{ edge.}}$$

$$= \frac{1}{2} \sum_{(u,v) \in E} \left[\frac{d(v)}{d(u)+d(v)} + \frac{d(u)}{d(u)+d(v)} \right]$$

$$= \frac{|E|}{2}$$

Overall cost:

Steps within a round easy to do in $O(m+n)$ work, $O(\log n)$ depth in BF model.
(setting priorities, removing $I \cup N(I)$, building G').

Vertices removed if no incident edges.

$O(m+n)$ work and $O(\log n)$ depth per round.

edges goes from $m \to \frac{m}{2}$ in expectation $\Rightarrow O(\log m) = O(\log n)$ rounds

$\Rightarrow$ $O(m+n)$ work and $O(\log^2 n)$ depth in expectation.

Are we done? Would be good to give tighter analysis of depth.

Let F_i = fraction removed on round i

M_i = # edges remaining on round i

i.e. $M_i = m \prod_{j=0}^{i-1} F_i$

Suppose we have $E[F_i] \le \beta$ (For Luby, $\beta = \frac{1}{2}$ in every round since we don't re-use randomness).

$$E[M_i] \le m \prod_{j=0}^{i-1} E(F_i) = m\beta^i$$

If work in a round is cM_i (assuming $|E_i| \ge |V_i|$), by LOE:

work on vertices bounded by work on the edges.

$$E[W(m)] \le \sum_{i=0}^{\infty} c\,E[M_i] \le cm \sum_{i=0}^{\infty} \beta^i = O(m)$$

Span: Use Markov

$$\Pr[X \ge a] \le \frac{E[X]}{a}$$

Consider $\Pr[M_i \ge 1]$. If $M_i < 1$, alg is done since no edges remain.

Would like to show that for some large i, this prob. is small.

$$\Pr[M_i \ge 1] \le \frac{E[M_i]}{1} \le m\beta^i$$

Let $i = k\log(m)$ for some k.

$$\Pr[M_i \ge 1] \le m\beta^{k\log(m)} = m\,m^{k\log(\beta)} = m^{1+k\log(\beta)}$$

$\beta = \frac{1}{2}$ for us. $\Rightarrow \log(\beta) = -1$

$\Pr[\text{rounds} < k\log(m)] \ge 1 - m^{1-k}$ $\Rightarrow$ $O(\log m)$ rounds whp. (also $O(\log n)$).