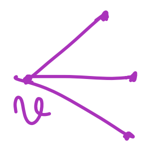
Maximal Independent Sets (MIS):

Independent set of a graph is a subset of vertices s.t for no edge both end points in that subset.

Formally, $I \subseteq V$ is an independent set if for all $v \in I$, $I \cap N(v) = \phi$.

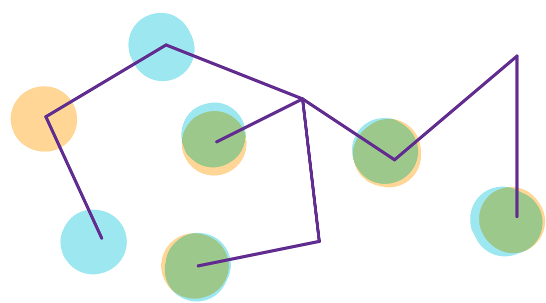
Neighbours of v.



Maximal Independent Set: An independent set I is maximal if it is not properly contained in any other independent set of G.

Note that Maximal Independent Set is not the same as Maximum Independent set.

NP-hard.



Algorithm (GreedyMIS)

Input: Graph $G = (V,E)$ where $V = \{v_1, \dots, v_n\}$.

Output: A maximal Independent Set.

1. $I \leftarrow \{\}$
2. For v in V:

   if $I \cap N(v) = \phi$:

   $I = I \cup \{v\}$
3. Return I.

while $V \neq \phi$:

Arbitrarily pick ~~v~~. Set S $\{|S| \leq k\}$

$\leq \frac{n}{k}$.

$I = \{v\} \cup I$. $\quad I \leftarrow I \cup S$

$V \leftarrow V \setminus (\{v\} \cup N(v))$.

$V \leftarrow V \setminus (S \cup N(S))$.

Question: Can this inherently sequential algorithm be parallelized?

Rough idea:

- Find a set $S \subseteq V$ s.t it is independent. (in parallel)
- Add S to I and delete $S \cup N(S)$ from G.

Repeat until I is maximal. How many iterations?

Question: How does one pick such a set S (in parallel)?

One idea is to pick a subset of vertices randomly.

Will they form an independent set?

We can modify our sampling by biasing it towards low degree vertices. Probably this will ensure that very few edges have both the end points in the random set.

Parallel algo for MIS: [Luby 1986].

Input: Graph $G = (V, E)$

Output: A maximal Independent Set $I \subseteq V$.

- $I \leftarrow \{\}$

- While V is not empty :

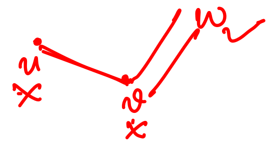
    $S \leftarrow \{\}$

    For all $v \in V$ do (in parallel)

        If $\deg(v) = 0$ then add v to I and $V = V \setminus \{v\}$.

        Else, mark v with prob $\frac{1}{2\deg(v)}$.

    For all $(u,v) \in E$ do (in parallel)



        If both u and v are marked then unmark the lower degree vertex.

    For all $v \in V$ do

        If v is marked, then add v to S.

    $I \leftarrow I \cup S$.

    Delete $S \cup N(S)$ from V and all incident edges from E.

$\boxed{\geq \frac{E}{2}}$

remaining $\leq \left(1 - \frac{1}{c}\right)^t E$

$O(\log(E))$ ✓

- Return I.

Algorithm makes progress if a good number of edges are removed from E during each iteration.

A vertex is "good" if for at least $\deg(v)/3$ neighbours have degree no more than $\deg(v)$.
An edge is good if at least one of its endpoints is good.

It is bad otherwise. $v \begin{cases} N^h(v) < \frac{2}{3}|N(v)| \\ N^l(v) \geq \frac{1}{3}|N(v)| \end{cases}$ } v is "good" else v is "bad"

Lemma: During the iteration, if a vertex w is marked then it is selected to be in S w.p $\geq \frac{1}{2}$. $S_0 \subseteq V$ S

Proof: $w \in S$ if and only if no vertex v in $N(w)$ of deg $\geq$ deg(w) is marked. $v \in N^h(w)$.

$$1 - \Pr[\exists v \in N(w) \text{ s.t } \deg(v) \geq \deg(w) \text{ and } v \text{ is marked}]$$

$$\geq 1 - \sum_{v \in N(w)} \underbrace{\Pr[v \in N(w) \text{ s.t } \deg(v) \geq \deg(w) \mid v \text{ is marked}]}_{\leq 1} \cdot \Pr[v \text{ is marked}]$$

$$\geq 1 - \sum_{v \in N^h(w)} \frac{1}{2 \cdot \deg(v)} \quad // \geq 1 - \sum_{v \in N^h(w)} \frac{1}{2 \cdot \deg(w)} - \sum_{v \in N^l(w)} \frac{1}{2 \cdot \deg(v)}$$

$$= 1 - \sum_{v \in N(w)} \frac{1}{2 \cdot \deg(w)}$$

$$\geq 1 - \sum_{v \in N^h(w)} \frac{1}{2 \cdot \deg(w)} \quad // \deg(v) \geq \deg(w) \Rightarrow \frac{1}{\deg(w)} \geq \frac{1}{\deg(v)}$$

$$\geq 1 - \frac{\deg(w)}{2 \cdot \deg(w)} \quad // |N^h(w)| \leq |N(w)|$$

Neighbours w/ higher degree.

$= 1/2.$

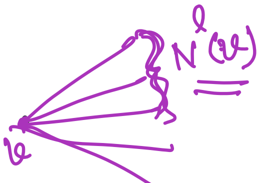
Lemma: Let $v \in V$ be a good vertex with degree $d_v > 0$. Then the prob that some vertex $w \in N(v)$ gets marked is at least $1 - \exp(-1/6)$. $\Rightarrow \frac{1 - e^{-1/6}}{2}$

Each w gets marked w/ prob $\frac{1}{2d_w}$.

Prob that these are not picked is

$$\leq \left(1 - \frac{1}{2d_v}\right)^{\frac{d_v}{3}} \cdot 1 \sim e^{-1/6}.$$

$-\frac{d_v}{3} \times \frac{1}{2d_v}$



Since v is good at least $\frac{d_v}{3}$ vertices have deg $\leq d_v$.

$$\frac{1}{2d_w} \geq \frac{1}{2d_v} \text{ for these}$$

$$\Pr \leq \left(1 - \frac{1}{\dots}\right)^{|N^l(v)|}$$

Lemma: In a graph $G=(V,E)$, the number of good edges is at least $\frac{|E|}{2}$.

Edge is good if one of its end points is good.

Proof: Recall that a vertex is bad if $\geq \frac{2}{3}$ fraction of its neighbours are of higher degree.

$$\begin{tikzcd} & N(v) \arrow[dl] \arrow[dr] & \\ N^h(v) & & N^\ell(v) \end{tikzcd}$$

Higher degree neighbours of v ($N^h(v)$) — Lower degree neighbours of v ($N^\ell(v)$)

Construct a directed graph from G:

$$(u,v)\in E \begin{cases} (u \to v) & \text{if } d_u \leq d_v, \\ (u \leftarrow v) & \text{otherwise.} \end{cases}$$

For a vertex v: $d_v^{out} ::$ out-degree of v

$d_v^{in} ::$ in-degree of v.

$V_B \subseteq V$ is the subset of all bad vertices.

$V_G = V \setminus V_B$ is the subset of all good vertices.

For a bad vertex: $\underbrace{d_v^{out}}_{\geq \frac{2d_v}{3}} - \underbrace{d_v^{in}}_{< \frac{d_v}{3}} \geq \frac{d_v}{3} = \frac{d_v^{out} + d_v^{in}}{3}$.

$$E(V_B, V_G) = \{(u \to v) \mid u \in V_B \text{ and } v \in V_G\}.$$

Similarly, we can define $E(V_B, V_B)$, $E(V_G, V_B)$ and $E(V_G, V_G)$.

Let $e(S,T)$ denote $|E(S,T)|$.

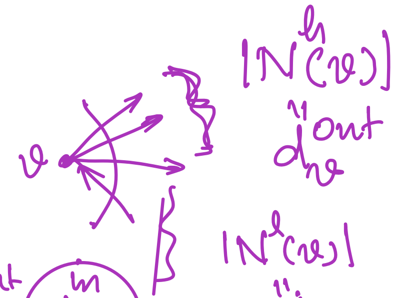
No. of good edges $= e(V_G,V_B)+e(V_G,V_G)+e(V_B,V_G)$.

✓ Sufficient to show that $e(V_B,V_B) \leq e(V_G,V_B)+e(V_B,V_G)$

Total degree of bad vertices is as follows:

$$\underline{2e(V_B,V_B)} + \underline{e(V_B,V_G)} + e(\underline{V_G,V_B})$$

$$= \sum_{v\in V_B} \left(d_v^{out} + d_v^{in}\right)$$



$$d_v = d_v^{out} + d_v^{in}$$

$$\leq 3\sum_{v\in V_B} \left(d_v^{out} - d_v^{in}\right)$$

$$= 3\left[\underbrace{\left(\cancel{e(V_B,V_B)} + e(V_B,V_G)\right)}_{=\sum_{v\in V_B} d_v^{out}} - \underbrace{\left(\cancel{e(V_B,V_B)} + e(V_G,V_B)\right)}_{=\sum_{v\in V_B} d_v^{in}}\right]$$

$$= 3\left[e(V_B,V_G) - e(V_G,V_B)\right]$$

$$\leq 3\left[e(V_B,V_G) + e(V_G,V_B)\right]$$

$$2e(V_B,V_B) + e(V_B,V_G) + e(V_G,V_B) \leq 3\left[e(V_B,V_G) + e(V_G,V_B)\right]$$

$$\Rightarrow 2e(V_B,V_B) \leq 2e(V_G,V_B) + 2e(V_G,V_B).$$

Summary: For each good vertex, its neighbour is marked w/ prob $1-e^{-1/6}$. Prob that the neighbour remains marked is $\geq \frac{1}{2}$. $\Rightarrow$ Prob that a good vertex is in SUN(S) $\geq \frac{1-e^{-1/6}}{2}$.

An edge is removed if an incident vertex is removed.

Exp no. of good edges is removed $\geqslant \alpha \cdot |\text{good } \underbrace{\text{edges}}_{}|$ $\overset{|E|/2}{}$

$$\left(\frac{1-e^{-1/6}}{2} \times \frac{1}{2}\right) E \qquad \left(\frac{1-e^{-1/6}}{2}\right)$$

Example: