Maximal Independent Set

1 Maximal Independent Set

Independent set of vertices is a set of vertices such that the set does not contain both a vertex and its neighbor. A maximum independent set is the largest such set. A maximal independent set is an independent set that can not be further increased in size.

Finding the maximum independent set is NP-complete. Can you tell me how to find a maximal independent set?

Pick an arbitrary vertex \( u \) and put it in IS
remove \( u \) and all its neighbors from consideration and remove all edges that are incident on these vertices.
repeat until all vertices are gone.

How do we parallelize it? Instead of just removing a single vertex and its neighbors on each iteration, we can try to remove set \( S \) of nodes and all the neighbors \( N(S) \).

We want to use graph contraction, and at every step we want to reduce the size of the graph by a constant factor. So in principle, we want to pick \( S \) such that \( S \cup N(S) \) is large. However, this turns out to be difficult. So instead, we will try to pick nodes \( S \) such that a large number of edges are incident on \( N(S) \) and we get to remove all those edges.

- For each vertex \( u \) label the vertex \( u \) such that \( L(u) = T \) with probability \( 1/(2d(u)) \) where \( d(u) \) is the degree of \( u \) and \( H \) otherwise. If the degree of a vertex is 0, definitely mark it as \( T \).
- For each edge \( (u, v) \) if both \( u \) and \( v \) are marked \( T \) then unmark (mark as \( H \)) the vertex with the lower degree.
- If \( L(u) = T \) then put \( u \) in MIS and remove all such \( u \) and their neighbors from \( v \) and remove all edges incident on these vertices.
- Repeat.

Each round takes \( O(m + n) \) work and \( O(\lg^2 n) \) span. Removing marked vertices and edges requires parallel prefix. If we assume that each round removes a constant fraction of the edges, then the total work is \( O(m + n) \) and the total span is \( O(\lg^3 n) \). In the next lecture, we will prove that each round removes a constant fraction of the edges.
2 Analysis of a Round

We now prove that each round removes a constant fraction of the edges. For each vertex, we define a set $X(v) = \{ u : u \in N(v) \cap d(u) \leq d(v) \}$. Therefore $X(v)$ contains a set of neighbors that have degree smaller than or equal to $v$. A vertex is good if $|X(v)| \geq d(v)/3$, that is, a lot of $v$’s neighbors have “low” degree. Otherwise a vertex is bad. An edge is good if at least one of its endpoints is good. Otherwise it is bad.

**Lemma 1** Say $v$ is a good vertex with $d(v) > 0$. Then the probability that at least one of its neighbors are marked initially is at least $1 - e^{-1/6}$.

**Proof.** The probability that at least one of $v$’s neighbors is marked is equal to 1 minus the probability that none of its neighbors are marked. Consider any neighbor $u$ of $v$. The probability that $u$ is marked is $1/2d(u)$. Therefore, the probability that it is not marked is $1 - 1/2d(u)$. Therefore, the probability that none of the neighbors is marked is

$$\prod_{u \in N(v)} (1 - 1/2d(u)) \leq \prod_{u \in X(v)} (1 - 1/2d(u))$$

For all $u \in X(v)$, $d(u) \leq d(v)$. Therefore, the above quantity is less than or equal to

$$(1 - 1/2d(v))^{|X(v)|} \leq (1 - 1/2d(v))^{d(v)/3} \leq e^{-d(v)/3 \times 2d(v)}$$

Therefore, the probability that some vertex is marked is at least $1 - e^{-1/6}$.

**Lemma 2** Let $M$ be the set of vertices that got marked and let $S$ be the set of vertices that were included in MIS. Then, given a vertex $v$ that was included in $M$, the probability that it is also in $S$ is at least $1/2$.

**Proof.** If $v$ was included in $M$, it is only removed if it had a neighbor $u$ which was also in $M$, and $d(u) > d(v)$. Given a vertex $u$, the probability that it was marked is $1/2d(u)$. The probability that some such neighbor was marked is at most

$$P \leq \sum_{u \in N(v), d(u) \geq d(v)} \frac{1}{2d(u)} \leq \sum_{u \in N(v), d(u) \geq d(v)} \frac{1}{2d(v)} \leq \sum_{u \in N(v)} \frac{1}{2d(v)} = 1/2.$$
Lemma 3 Let $V_g$ be the set of good vertices and $S$ be the set of vertices included in the MIS. The probability that a good vertex $v$ was in either $S$ or the neighbors of $S$ (that is, it disappeared in that round) is at least $1/2(1 - e^{-1/6})$.

Proof. A vertex $v$ is in the neighborhood of $S$ if there is another vertex $u$ that is neighbor of $v$ and is in $S$. Since $v$ is a good vertex, we know that some neighbor of $v$ was marked with probability $1 - e^{-1/6}$. Also, we know that the probability that a vertex is unmarked after being marked is at most $1/2$. Therefore, some neighbor of $v$ remained marked with probability at least $1/2(1 - e^{-1/6})$. Therefore, $v$ was removed with at least that probability.

Corollary 4 A good edge is removed with probability at least $1/2(1 - e^{-1/6})$.

Lemma 5 At any time, at least half the edges are good.

Proof. For each edge $(u, v)$ direct it from $u$ to $v$ if $d(u) \leq d(v)$, breaking ties arbitrarily. In the resulting graph, let $d_o(v)$ and $d_i(v)$ be the out and in-degrees. Let $V_g$ and $V_b$ be the set of good and bad vertices. For each bad vertex, $d_o(v) - d_i(v) \geq d(v)/3$. Let $E_{bb}, E_{bg}, E_{gb}$ and $E_{gg}$ be the edges from $V_b$ to $V_b$, $V_b$ to $V_g$ and so on. The sum of the degrees of bad vertices is

$$2|E_{bb}| + |E_{bg}| + |E_{gb}| = \sum_{v \in V_b} d(v)$$

$$\leq 3 \sum_{v \in V_b} (d_o(v) - d_i(v))$$

$$= 3 \sum_{v \in V_g} (d_i(v) - d_o(v))$$

$$= 3(|E_{bg}| + |E_{gg}|) - (|E_{gb}| + |E_{gg}|)$$

$$= 3(|E_{bg}| - |E_{gb}|)$$

Therefore, $2|E_{bb}| \leq 2|E_{bg}| - 4|E_{gb}|$. Therefore $|E_{bb}| \leq |E_{bg}|$. Other than the edges that go from bad to bad vertices, all other edges are good edges. Therefore, at least half the edges are good.

Theorem 6 The expected number of edges removed per round is at least $1/4(1 - e^{-1/6}) |E|$.