Notes for Luby Algorithm: Parallel Maximal Independent Sets

Abdelhak Bentaleb
A0135562H

January 28, 2016

1 The Algorithm

Problem : Given a graph find a maximal independent set.

1. I = $\emptyset$, G'= G.
2. While (G' is not the empty graph) do IN PARALLEL
   (a) Choose a random set of vertices S $\in$ G' by selecting each vertex v independently with probability 1/2d(v).
   (b) For every edge (u, v) $\in$ E(G') if both endpoints are in S then remove the vertex of lower degree from S (Break ties arbitrarily). Denote this new set S'.
   (c) I = I $\cup$ S'. G' = G' (S' $\cup$ N(S')), i.e., G' is the induced subgraph on V' (S' $\cup$ N(S')) where V' is the previous vertex set.
3. output the independent set I

Correctness : We see that at each stage the set S' that is added is an independent set. Moreover since we remove, at each stage, S' $\cup$ N(S') the set I remains an independent set. Also note that all the vertices removed from G' at a particular stage are either vertices in I or neighbors of some vertex in I. So the algorithm always outputs a maximal independent set.

- A single round can be done in constant time using O($|V|^2$) processors
- The expected value of the number of rounds is in **O(log n)**. With n = $|E|$

2 Expected Running Time (Number of rounds)

In this section we prove that the algorithm take O(log n) time. Let $G_j = (V_j, E_j)$ denote the graph after stage j

Main Lemma: For some k < 1,

$$E(|E_j| \mid E_{j-1}) < k|E_{j-1}|.$$

Hence, in expectation, only O(log n) rounds will be required, where n = $|E_0|$. We say vertex v is *bad* if more than 2/3 of the neighbors of v are of higher degree than v. We say an edge is *bad* if both of its endpoints are bad, otherwise the edge is *good*.

The key claims are that at least half the edges are *good*, and each *good* edge is deleted with a constant probability. The main lemma then follows immediately.

Lemma 1 : At least half the edges are *good*.

Proof : Denote the set of bad edges by E_B. We will define $f : E_B \rightarrow \begin{pmatrix} E \\ 2 \end{pmatrix}$ so that for all $e_1 \neq e_1 \in E_B, f(e_1) \cap f(e_2) = \emptyset$. This proves $|E_B| \leq |E|/2$, and we are done.

The function f is defined as follows. For each (u, v) $\in$ E, direct it to the higher degree vertex. Break ties as in the algorithm. Now, suppose (u, v) $\in E_B$, and is directed towards v. Since v is *bad*, it has at least twice as many edges out as in. Hence we can pair two edges out of v with every edge into v.

Lemma 2: If v is *good* then $Pr(N(v) \cap S \neq \emptyset) \geq 2\alpha$, where $\alpha = 1/2 \times (1 - e^{-1/6})$.

Proof: Define L(v) := $\{w \in N(v) | d(w) \leq d(v) \}$.

By definition, $|L(v)| \geq d(v)/3$ if v is a *good* vertex.

$$\begin{aligned}
\Pr(\ N(v) \cap S \neq \emptyset\ ) &= 1 - \Pr(\ N(v) \cap S = \emptyset\ ) \\
&= 1 - \prod_{w \in N(v)} \Pr(\ w \notin S\ ) \quad \text{using full independence} \\
&\geq 1 - \prod_{w \in L(v)} \Pr(\ w \notin S\ ) \\
&= 1 - \prod_{w \in L(v)} \left(1 - \frac{1}{2d(w)}\right) \\
&\geq 1 - \prod_{w \in L(v)} \left(1 - \frac{1}{2d(v)}\right) \\
&\geq 1 - \exp(-|L(v)|/2d(v)) \\
&\geq 1 - \exp(-1/6),
\end{aligned}$$

Note, the above lemma is using full independence in its proof. And Pr: Probability.

Lemma 3: Pr (w $\in$ S' | w $\in$ S) $\leq$ 1/2.

Proof : Let H(w) = N(w) \ L(w) = {z $\in$N(w) : d(z) > d(w)}.

$$
\begin{aligned}
\Pr(\ w \notin S' \ \mid \ w \in S\ ) &= \Pr(\ H(w) \cap S \neq \emptyset \ \mid \ w \in S\ ) \\
&\leq \sum_{z \in H(w)} \Pr(\ z \in S \ \mid \ w \in S\ ) \\
&\leq \sum_{z \in H(w)} \frac{\Pr(\ z \in S, w \in S\ )}{\Pr(\ w \in S\ )} \\
&= \sum_{z \in H(w)} \frac{\Pr(\ z \in S\ )\Pr(\ w \in S\ )}{\Pr(\ w \in S\ )} \quad \text{using pairwise independence} \\
&= \sum_{z \in H(w)} \Pr(\ z \in S\ ) \\
&= \sum_{z \in H(w)} \frac{1}{2d(z)} \\
&\leq \sum_{z \in H(w)} \frac{1}{2d(v)} \\
&\leq \frac{1}{2}.
\end{aligned}
$$

Lemma 4: If v is *good* then Pr(v $\in$ N(S')) $\geq \alpha$

Proof: Let V$_G$ denote the *good* vertices. We have

$$
\begin{aligned}
\Pr(\ v \in N(S') \ \mid \ v \in V_G\ ) &= \Pr(\ N(v) \cap S' \neq \emptyset \ \mid \ v \in V_G\ ) \\
&= \Pr(\ N(v) \cap S' \neq \emptyset \ \mid \ N(v) \cap S \neq \emptyset, v \in V_G\ )\Pr(\ N(v) \cap S \neq \emptyset \ \mid \ v \in V_G\ ) \\
&\geq \Pr(\ w \in S' \ \mid \ w \in N(v) \cap S, v \in V_G\ )\Pr(\ N(v) \cap S \neq \emptyset \ \mid \ v \in V_G\ ) \\
&\geq (1/2)(2\alpha) \\
&= \alpha
\end{aligned}
$$

Corollary 4: If v is *good* then the probability that v gets deleted is at least α.

Corollary 5: If an edge e is *good* then the probability that it gets deleted is at least α.

Proof: Pr (e = (u, v) $\in E_{j-1} \setminus E_j$) $\geq$ Pr (v gets deleted).

We now return the main lemma :

Main Lemma :

$$E(|E_j| \mid E_{j-1}) < k|E_{j-1}|(1 - \alpha/2).$$

Proof:

$$
\begin{aligned}
\mathrm{E}(\ |E_j| \ \mid \ E_{j-1}\ ) &= \sum_{e \in E_{j-1}} 1 - \Pr(\ e \text{ gets deleted }) \\
&\leq |E_{j-1}| - \alpha|GOOD \text{ edges}| \\
&\leq |E_{j-1}|(1 - \alpha/2).
\end{aligned}
$$

The constant α is approximately 0.076.

Thus,

$$E(|E_j|) \leq |E_0|(1-\alpha/2)^j \leq \text{n } exp(-j\alpha/2) < 1$$

for j ¿ $2/\alpha$ log n. Therefore, the expected number of rounds required is $\leq$ 4n = O(log n).

Theorem:

The expected number of rounds in Luby algorithm for finding maximal independent set is in O(log n).

References

[1] Motwani, Rajeev, and Prabhakar Raghavan. Randomized algorithms. Chapman and Hall/CRC, 2010.

[2] Maximal independent set
`https://en.wikipedia.org/wiki/Maximal_independent_set` .

[3] Blelloch, Guy E., Jeremy T. Fineman, and Julian Shun. "Greedy sequential maximal independent set and matching are parallel on average." Proceedings of the twenty-fourth annual ACM symposium on Parallelism in algorithms and architectures. ACM, 2012.

[4] Luby, M. (1986). "A Simple Parallel Algorithm for the Maximal Independent Set Problem". SIAM Journal on Computing 15 (4): 1036.

[5] Luby's Algorithm, in: Lecture Notes for Randomized Algorithms, Last Updated by Eric Vigoda on February 2, 2006