Topics in Distributed Computing
Lecture 8

Lecturer: Alessandro Panconesi Scribe: Lars Arvestad

21 March 1997

8.1 Maximal Independent Sets in General Graphs

In this lecture we will continue our study of the maximal (as opposed to the *maximum*) independent set problem (henceforth MIS). We will still consider the synchronous, message-passing model but the underlying topology will be that of any graph $G = (V, E)$, not just the ring.

It is worth pointing out that the problem is absolutely trivial in a sequential setting, as the following linear time greedy algorithm shows: pick one vertex, put it into the independent set, remove its neighbors; pick another vertex, put it into the independent set, remove its neighbours, and so on.

In some sense we would like to parallelize this simple algorithm. Perhaps surprisingly, this turns out to be extremely hard to do. In fact, this is hard to do even in the PRAM model, where this problem was originally studied. This seemingly simple problem is so challenging that in order to solve it it was necessary to develop a new and very interesting method called *derandomization*, to which we return. The solution we will develop makes use of randomization; whether there exists deterministic algorithms for MIS running in poly-logarithmic many rounds is a challenging open problem.

The algorithm we discuss is known as Luby's algorithm, after his inventor [5]. Essentially the same algorithm and analysis, however, can be found in [1]. Here we follow the crisp presentation of [3].

Although Luby's algorithm was originally developed in the PRAM model, it is also a bonafide distributed algorithm. We now describe the algorithm.

The algorithm runs in stages. During each stage the algorithm computes an independent set I and removes the set $I \cup N(I)$. If the leftover graph is non empty it becomes the input to the next stage, otherwise the algorithm stops. The output of the algorithm is the union of all sets I. It is easy to see that this gives a MIS.

A stage of the algorithm is as follows.

1. Each vertex u, in parallel, enters a set S of candidates with probability $1/2d(u)$.

2. If both u and one of its neighbours are in S, discard the lower degree vertex (ties are resolved arbitrarily). The resulting set is I.

The basic problem is to show that during each stage enough progress is made. In particular we want to show that the expected number of stages is $O(\log n)$. To do this, we will show that each stage removes an expected constant fraction of all the edges of the current graph. We define

$$L(u) = \{v : v \in N(u) \wedge d(v) \leq d(u)\},$$

and call it the set of low degree neighbours of u. The next definition is pivotal.

Definition 8.1.1. *A vertex u is* good *if $|L(u)| \geq d(u)/3$.*

Intuitively, a good vertex has a lot of low degree neighbours and hence has a good chance of making it into $I \cup N(I)$.

Fact 8.1.1. *If u is good, then*

$$\sum_{v \in N(u)} \frac{1}{2d(v)} \geq \frac{1}{6}$$

Proof.

$$\sum_{v \in N(u)} \frac{1}{2d(v)} \geq \sum_{v \in L(u)} \frac{1}{2d(v)} \geq \frac{d(u)}{3} \cdot \frac{1}{2d(u)} = \frac{1}{6}.$$

□

Definition 8.1.2. *An edge $e = (u, v)$ is good if at least one of its neighbors is good. A bad edge is an edge which is not good.*

With these definitions at hand we make our plan more specific. First, we show that in any graph, at least half of the edges are good. Second, we show that the probability that a good edge is removed during a stage is a constant fraction. This means that the expected number of good edges which are removed at each stage is a constant fraction of the total and this implies that the expected number of stages is $O(\log n)$, where n is the number of vertices of the input graph. A more sophisticated analysis actually shows that the running time is $O(\log n)$ with high probability [2, 4].

Lemma 8.1.1. *Let E_G and E_B denote the sets of good and bad edges respectively. Then*

$$|E_G| \geq \frac{|E|}{2}.$$

Proof. We define a mapping $f : E_B \to E^2$ such that

$$f(e_1) \cap f(e_2) = \emptyset$$

and

$$f(e) = (g, h) \to g \neq h.$$

In other words, to each bad edge we can associate a pair of edges (bad or good it doesn't matter) in such a way that the pairs are all disjoint and each pair contains different edges. Such a mapping implies that $|E_B|$ is at most half of E.

The mapping is defined as follows. First, direct each edge of G toward the lower degree vertex, resolving ties arbitrarily. Focus now on a bad vertex. This vertex will have some edges pointing inward and some outward. By definition of bad vertex, at least one third of its neighbours have strictly higher degree. Therefore, there are at least twice as many edges pointing outward than edges pointing inward. This enables us to associate a pair of outward edges to every inward edge, each time using two brand new outward edges. Notice now that, again by definition, all inward edegs are bad. Hence, this defines a mapping $f : E_B \to E^2$.

Does f satisfies the properties we want? By construction each $f(e)$ is a pair of different edges. To see that for any two edges $e_1, e_2 \in E_B$, $f(e_1) \cap f(e_2) = \emptyset$ we consider two cases. If e_1 and e_2 point to the same vertex, then $f(e_1) \cap f(e_2) = \emptyset$ by construction. Otherwise, if e_1 and e_2 are directed towards different vertices u and v, then the only edge that could be in the intersection of $f(e_1)$ and $f(e_2)$ is (u, v). But due to the fact that the edge is directed, it can be placed in either $f(e_1)$ or $f(e_2)$, but not in both. $\square$

What remains to be shown is that, in expectation, a constant fraction of good edges is removed at each stage. To prove this we will make use of the following two facts.

Fact 8.1.2. *For any vertex u,* $\Pr(u \in I | u \in S) \geq 1/2$.

Proof. We upper bound the complementary event $\Pr(u \notin I | u \in S)$. A vertex $u \in S$ does not make it into I if and only if a neighbour whose degree is higher or equal to that of u enters S. Therefore,

$$\begin{aligned}
\Pr(u \notin I | u \in S) &= \Pr(\exists v \in N(u), v \in S, d(v) \geq d(u) | u \in S) \\
&\leq \sum_{\substack{v \in N(u) \\ d(v) \geq d(u)}} \Pr(v \in S | u \in S) \leq \sum_{\substack{v \in N(u) \\ d(v) \geq d(u)}} \Pr(v \in S) \\
&\leq \sum_{\substack{v \in N(u) \\ d(v) \geq d(u)}} \frac{1}{2d(u)} \leq d(u) \frac{1}{2d(u)} = \frac{1}{2}
\end{aligned} \tag{8.1}$$

$\square$

It is conveneient at this point to make a very important
Remark: Equation (8.1) still holds if the assumption of the independence of the unbiased coin flip is replaced with the much weaker one of *pairwise independence*. A collection of events $\mathcal{A}$ are pairwise independent if, for all distinct events $A, B \in \mathcal{A}$ we have $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

Fact 8.1.3. $\forall u \in V, \Pr(u \in I) \geq \frac{1}{4d(u)}$.

Proof. Using Fact 8.1.2,

$$\Pr(u \in I) = \Pr(u \in I | u \in S)\Pr(u \in S) \geq \frac{1}{2} \cdot \frac{1}{2d(u)} \geq \frac{1}{4d(u)}. \tag{8.2}$$

$\square$

Lemma 8.1.2. *Let $e = (u, v)$ be a good edge. Then,* $\Pr(e$ *removed*$) \geq 1/36$.

Proof. Assume without loss of generality that u is good. Then, $\Pr(e$ removed$)$ is lower bounded by $\Pr(u \in I \cup N(I))$ which we now estimate. There are two cases to consider. The first case is: $\exists v \in N(u), d(v) \leq 2$. By Fact 8.1.3,

$$\Pr(u \in I \cup N(I)) \geq \Pr(v \in I) \geq \frac{1}{4d(v)} = \frac{1}{8}.$$

The second, complementary, case is: $\forall v \in N(u), d(v) \geq 3$. This case must be examined more carefully. By fact 8.1.1,

$$\sum_{v \in N(u)} \frac{1}{2d(v)} \geq \frac{1}{6}.$$

Since each term in this sum is smaller than $1/6$, we can also find a subset A of $N(u)$ such that

$$\frac{1}{6} \leq \sum_{v \in A} \frac{1}{2d(v)} \leq \frac{1}{3}.$$

Then,

$$
\begin{aligned}
\Pr(u \in I \cup N(I)) &\geq \Pr(\exists v \in A \cap I) \\
&\geq \sum_{v \in A} \Pr(v \in I) - \sum_{v,w \in A} \Pr(v \in I, w \in I) \qquad (8.3) \\
&\geq \sum_{v \in A} \Pr(v \in I) - \sum_{v,w \in A} \Pr(v \in S, w \in S) \\
&= \sum_{v \in A} \Pr(v \in I | v \in S) \Pr(v \in S) \\
&\quad - \left(\sum_{v \in A} \Pr(v \in S) \right) \left(\sum_{w \in A} \Pr(w \in S) \right) \qquad (8.4) \\
&\geq \left(\sum_{v \in A} \Pr(v \in S) \right) \left(\frac{1}{2} - \sum_{w \in A} \Pr(w \in S) \right) \qquad (8.5) \\
&\geq \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} \qquad (8.6)
\end{aligned}
$$

where Equation (8.3) follows from the mutual-exclusion formula, Equation (8.5) follows from Fact 8.1.2, and Equation (8.6) follows from the definition of the set A. $\square$

Remark: once again, notice that the proof of Lemma 8.1.2 hinges only on pairwise independence. Specifically, this is used in Equation (8.4).

Combining the last two lemmata shows that the expected fraction of edges which are deleted per stage is $1/72$, which concludes the proof.

8.2 Detour: Small stochastic spaces

A randomized algorithm that uses $O(\log n)$ random bits actually yield a polynomial time algorithm, or better, an NC-algorithm. Simply try the $2^{O(\log n)}$ possible random strings in parallel. The random algorithm guarantees that at least one string is successful.

We will show how to convert the previous algorithm to a deterministic version. The key point is pairwise independence in order to make the above calculations to be correct. Specifically, we rely on the fact that $\Pr(v \in S | u \in S) = \Pr(v \in S)$.

8.2.1 How to generate pairwise independence

Select a prime p, $n \leq p \leq 2n$. Such a prime exists because of the density of primes. For each vertex u, choose an integer a_u such that $\frac{a_u}{p} \approx \frac{1}{2d(u)}$, and an interval A_u, $|A_u| = a_u$ in Z_p.

Let $X(u)$ be the random variable from uniformly choosing $x, y \in Z_p$ and setting

$$X(u) = \begin{cases} 1 & \text{if } xu + y \in A_u, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{8.7}$$

This means that $\Pr(X(u) = 1) = \frac{a_u}{p} \approx \frac{1}{2d(u)}$ because

$$\begin{aligned} \Pr_{x,y}(X(u) = 1) &= \Pr(\exists a \in A_u : xu + y = a) \\ &= \sum_{a \in A_u} \Pr(x = (a - y)u^{-1}) = \sum_{a \in A_u} \frac{1}{p} = \frac{a_u}{p}. \end{aligned} \tag{8.8}$$

For pairwise independence, we would like to conclude that

$$\Pr(X(u) = 1, X(v) = 1) = \Pr(X(u) = 1)\Pr(X(v) = 1) = \frac{a_u a_v}{p^2}. \tag{8.9}$$

This is also true because of the following.

$$\begin{aligned} \Pr(X(u) = 1, X(v) = 1) &= \Pr(\exists a \in A_u, b \in A_v : xu + y = a, xv + y = b) \\ &= \Pr(\exists (a, b) \in A_u \times A_v : xu + y = a, xv + y = b) \\ &= \Pr\left(\exists (a, b) : \begin{bmatrix} u & 1 \\ v & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}\right) \\ &= \sum_{(a,b) \in A_u \times A_v} \Pr\left(\begin{bmatrix} u & 1 \\ v & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}\right) \\ &= \sum_{(a,b) \in A_u \times A_v} \frac{1}{p^2} = \frac{a_u a_v}{p^2} \end{aligned} \tag{8.10}$$

Bibliography

[1] N. Alon, L. Babai, and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem J. Algorithms 7:567–583, 1986

[2] S. Chaudhuri and D. Dubhashi, Probabilistic recurrence relations revisited, Proc. of the 2nd Latin Amer. Symp. (LATIN 95), April 1995. LNCS 911, 207–219. To appear in Theoretical Comp. Sci

[3] D. Kozen, The design and analysis of algorithms, Springer Verlag, Berlin

[4] D. Karp, Probabilistic recurrence relations, Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC 1991) pp. 190-197, New Orleans, Luisiana

[5] M. Luby, A simple parallel algorithm for the maximal independent set problem Proceedings of the 17th Annual ACM Symposium on Theory of Computing (STOC 1985) pp. 1-10, Providence, Rhode Island