1 Lecture 14, January 20, 2000

Last week we presented a randomized algorithm for approximating the domatic number. In this lecture we show how to derandomize this algorithm, using the method of conditional probabilities, and get a greedy algorithm whose approximation ratio is at least as good as the expected ratio of the randomized algorithm.

In the second part of the lecture we study some algorithms in number theory. Specifically, we present the extended GCD algorithm.

1.1 Domatic number

A set $S \subset V$ of vertices in a graph $G(V, E)$ is a *dominating set* if every vertex in $V \setminus S$ is adjacent to some vertex in S. A *domatic partition* of G is a partition of its vertices into disjoint dominating sets. The *domatic number* is the maximum number of dominating sets in a domatic partition.

Let $N(v_i)$ denote the set of neighbors of vertex v_i, including v_i itself. Let $\delta = \min_i[|N(v_i)|]$. Then the domatic number can be at most δ. In the previous lecture we showed a proof based on the probabilistic method that G has a domatic partition with roughly $\delta/\ln n$ dominating sets. In this lecture we present a greedy deterministic algorithm for finding such a domatic partition. This greedy algorithm is based on a derandomization of the probabilistic method. For completeness, we sketch the probabilistic proof, though the reader is advised to review last week's lecture.

Let $l = \delta/\ln(n \ln n)$. The randomized algorithm works as follows:

1. Give each vertex a color between 1 and l at random.
2. The resulting domatic partition consists of those color classes that are dominating sets, with vertices in other color classes added to an arbitrary dominating set.

Let $A_{v,c}$ denote the bad event that $N(v)$ does not include color c. Then $Pr[A_{v,c}] \leq (1 - 1/l)^\delta \leq e^{-\ln(n \ln n)} = 1/n \ln n$. Treating the $A_{v,c}$ as 0/1 indicator variables, let $A = \sum_{v,c} A_{v,c}$. Using linearity of expectation, the expectation of A satisfies

$$E[A] = \sum_{v,c} E[A_{v,c}] = \sum_{v,c} Pr[A_{v,c}] \leq \sum_{v,c} \frac{1}{n \ln n} = \frac{l}{\ln n}.$$

We now give a greedy algorithm that colors the vertices of G with l colors, such that the number of bad events is at most $l/\ln n$. This gives a domatic partitition with at least $(1-\epsilon)\delta/\ln n$ dominating sets, where $\epsilon < (1 + \ln\ln n)/\ln n$.

In our greedy algorithm, we sort the vertices in an arbitrary order $v_1, \ldots, v_n$. We then consider the vertices in order, and assign a color to each vertex without ever needing to revoke an earlier decision. We assign colors to vertices using the *method of conditional expectations*, which works as follows.

Assume that vertices $v_1, \ldots, v_j$ have already been assigned the colors $c_1, \ldots, c_j$ (the colors need not be distinct). Let $E[A|c_1, \ldots, c_j]$ denote the expectation of A conditioned on the above. Using linearity of expectation we have

$$E[A|c_1, \ldots, c_j] = \frac{1}{l} \sum_{c_{j+1}=1}^{l} E[A|c_1, \ldots, c_j, c_{j+1}]$$

where the $1/l$ term is the probability that vertex v_{j+1} is colored c_{j+1}. It follows that for some value of c_{j+1}, $E[A|c_1, \ldots, c_j] \geq E[A|c_1, \ldots, c_j, c_{j+1}]$. We assign this color to v_{j+1}. Eventually, when v_n is assigned a color, we have by transitivity that $E[A|c_1, \ldots, c_n] \leq E[A]$, which implies that the number of bad events is at most $l/\ln n$, as desired.

What we still need to show is that we can compute the conditional expectations in polynomial time. Let $d_i^j = |N(v_i) \bigcap \{v_{j+2}, \ldots, v_n\}|$. At a certain step of the algorithm, let a_i^c be an indicator variable indicating whether none of the neighbors of v have yet received the color c. After step $j+1$, $a_i^c = 1$ if none of the vertices in $N(v_i) \bigcap \{v_1, \ldots, v_{j+1}\}$ has color c, and 0 otherwise. Then $E[A_{v,c}|c_1, \ldots, c_j, c_{j+1}] = \sum_{c=1}^{l} a_i^c (1 - 1/l)^{d_i^j}$. We color v_{j+1} by the color c_{j+1} minimizing

$$E[A|c_1, \ldots, c_j, c_{j+1}] = \sum_{v,c} E[A_{v,c}|c_1, \ldots, c_j, c_{j+1}].$$

Rather than compute explicitly the expression $\sum_{v,c} E[A_{v,c}|c_1, \ldots, c_j, c_{j+1}]$ for every value of c_{j+1}, it suffices just to find the value of c_{j+1} for which this expression is minimizes. This requires less work, due to cancelations of many of the terms involved in the difference of two such expressions.

1.2 Algorithms in number theory

These were the subjects covered in class:

- Given two integers (a, b), the extended GCD algorithm returns three integers (d, x, y) such that $d = gcd(a, b)$ and $d = ax + by$.

- A finite group $(S, \oplus)$ satifies the following four properties for every $a, b, c \in S$:

  1. Closure: $a \oplus b \in S$.
  2. Identity: $\exists e$ such that $e \oplus a = a \oplus e = a$.
  3. Associativity: $(a \oplus b) \oplus c = a \oplus (b \oplus c)$.
  4. Inverse: $\forall a \; \exists b$ such that $a \oplus b = b \oplus a = e$.

  The groups we will be interesting will also be commutative (Abelian).

- With an integer n we can associate the finite Abelian groups:

  - Addition modulo n, $(Z_n, +_n)$.
  - Multiplication modulo n, $(Z_n^*, \cdot_n)$, where Z_n^* is the set of numbers relatively prime to n. Inverses are found using the extended GCD algorithm.

- The *order* of a group is the number of elements in it. The order of subgroups divides that of the group. If $(S, \oplus)$ is a group and $S' \subset S$ is closed under $\oplus$, then $(S', \oplus)$ is a (sub)group (identity, associativity and inverse can be proved). $|Z_n^*| = \phi(n) = n\Pi_p(1 - 1/p)$, where p ranges over the prime devisors of n.

For more details, see for example the book by Cormen, Leiserson and Rivest.