Graduate Algorithms

Homework 6

Virginia Vassilevska, Elisabeth Crawford

April 28, 2004

Problem 1 We give an algorithm for achieving an approximation ratio of 3. Take any MST T of the graph G. Root T at some vertex v. We will collect a set of cycles S which will cover the nodes in T. For every node u in the tree we do the following. If u has at least 2 children, $(c_1, c_2, ..., c_i)$ include the edges from the cycle that starts at u, goes to c_1 then to c_3, ... c_i, c_2 and finally back to u. If u has only one child c we include the edge (u, c) and its copy (c, u). Note that these cycles contain no edge greater than twice the largest sized edge in the MST by the triangle property. Note that the cycles in S have the property that no three of them share a common point. Two cycles intersect in at most one vertex of T. Now we argue inductively that we can use these cycles to form a hamiltonian tour.

Suppose u is the intersection of a lower L and an upper U cycle (i.e. the root r of U is closer to v than u which by construction is the root of L). Suppose further that there are no cycles rooted at the boundary of L and that u is the first root that appears in a counterclockwise traversal of U starting from r (Figure 1). We know that u has at least one child c so that the tree edge (u, c) is on the boundary of L, and that there is a nontree edge (u, u') on the boundary of U incident to u and in the counterclockwise direction from u. Substitute (c, u) and (u, u') with (c, u') which is no more than the sum of those and is hence at most 3 times the maximum weight edge in T. This merges L and U into one bigger cycle C. We still have the following satisfied: To every root u''

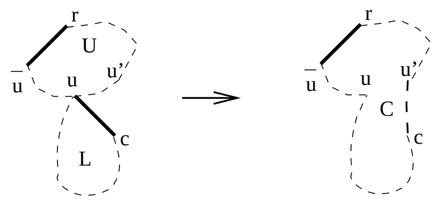

Figure 1: Here we merge cycles U and L into C maintaining our invariants. Note that it is possible that $\bar{u} = u$ or that $r = u'$ but this does not change the proof.

of a cycle rooted at C there is an incident nontree edge different from (c, u') on the boundary of C (in the counterclockwise direction from r''); the next root of a lower cycle in the counterclockwise traversal of C is the root after u in the counterclockwise traversal of U; there is a tree edge $(r, \bar{u})$ on the boundary of C. By induction we can continue this process to merge the cycles in S into a Hamiltonian cycle. We never remove edges (c, u') once we have added them and hence no edge in the Hamiltonian cycle has weight greater than 3 times the maximum weight edge in T. Now we note that the largest edge in T is always less than or equal to the largest edge in the optimal bottleneck Hamiltonian tour H: consider the maximum weight edge e in T. It connects two components T_1 and T_2 in T. Consider an edge e' in H which connects two vertices one in each component (e' exists since H is spanning). Note that if e' has weight strictly smaller than that of e then T is not a MST since substituting e' for e creates a smaller tree. Hence no edge in T is larger than the maximum edge in H. Since we only include edges in our Hamiltonian tour that are no more than 3 times the size of the largest edge in the MST we have an approximation ratio of 3 for bottleneck-TSP as required.

Problem 2

a We want the probability of returning to the origin when we take a S-step random walk in one dimension. You return to the origin in S steps if and only if you have taken $\frac{S}{2}$ steps to the right and $\frac{S}{2}$ steps to the left. The probability of a fixed sequence of steps with this property is $\frac{1}{2^{\frac{S}{2}}}\frac{1}{2^{\frac{S}{2}}} = \frac{1}{2^S}$. Yet, there are many possible such sequences of steps. In fact, the exact number of them is the number of ways to pick the positions of the steps to the right in a sequence of S steps, which is $\binom{S}{\frac{S}{2}}$. Hence the probability to return to the origin in S steps is 0 if S is odd and if S is even by Stirling's formula we have:

$$\binom{S}{\frac{S}{2}}\frac{1}{2^S} \sim \frac{\sqrt{2S\pi + \frac{\pi}{3}}S^S e^S}{(\sqrt{S\pi + \frac{\pi}{3}}(\frac{S}{2})^{(\frac{S}{2})}e^{(\frac{S}{2})})^2 2^S} = \frac{\sqrt{2S + \frac{1}{3}}}{S + \frac{1}{3}} = O(\frac{1}{\sqrt{S}})$$

b The way to get back to the origin in a 2 dimensional random walk is to have made the same number of steps up as down and the same number left as right. Let u be the number of steps going up, and r be the number of steps going right, then if you get back to the origin in S steps, $2u + 2r = S$. The probability for a particular sequence of S steps which gets you back to the origin is then $\frac{1}{4^u}\frac{1}{4^u}\frac{1}{4^r}\frac{1}{4^r} = \frac{1}{4^S}$. The number of such sequences is

$$\sum_{r=0}^{S/2}\binom{S}{r}\binom{S-r}{r}\binom{S-2r}{\frac{S-2r}{2}}$$

because, given that there are r steps to the right, there are $\binom{S}{r}$ ways to choose when to perform them in the sequence of S steps. From the remaining $S - r$

steps you can pick the r steps to the left to be in $\binom{S-r}{r}$ ways, and then there are $\binom{S-2r}{\frac{S-2r}{2}}$ ways to arrange the vertical steps. Hence the probability to return to the origin in S steps is

$$\frac{1}{4^S}\sum_{r=0}^{S/2}\frac{S!}{(r!)^2(\frac{S-2r}{2}!)^2} = \binom{S}{\frac{S}{2}}^2\frac{1}{4^S}$$

Now, we have by Stirling:

$$\binom{S}{\frac{S}{2}}^2\frac{1}{4^S} \sim (\frac{\sqrt{2S+\frac{1}{3}}}{S+\frac{1}{3}})^2 = O(\frac{1}{S}).$$

As before, if S is odd, the probability is 0. The bound above is for even S.

c Suppose we take S steps in a d dimensional random walk. If S is odd the probability of getting back to the origin in these S steps is 0. If S is even, then we can argue as before. Let k_i be the number of positive steps taken in the ith dimension. Then this should be the same as the number of negative steps in the same dimension. For a particular sequence of S steps, the probability is $\sum_{i=1}^{d}(\frac{1}{2d})^{k_i}(\frac{1}{2d})^{k_i} = (\frac{1}{2d})^S$. The number of ways to arrange the steps is

$$\sum_{\sum_{i=1}^{d} k_i = \frac{S}{2}}\prod_{i=1}^{d}\binom{S-2\sum_{j=1}^{i-1}k_j}{2k_i}\binom{2k_i}{k_i}$$

and so the probability of returning to the origin is described by the following horrific expression:

$$(\frac{1}{2d})^S\sum_{\sum_{i=1}^{d} k_i = \frac{S}{2}}\prod_{i=1}^{d}\binom{S-2\sum_{j=1}^{i-1}k_j}{2k_i}\binom{2k_i}{k_i}$$

We can simplify this by expanding the binomial coefficients into factorials:

$$\prod_{i=1}^{d}\binom{S-2\sum_{j=1}^{i-1}k_j}{2k_i} = \prod_{i=1}^{d}\binom{2\sum_{j=i}^{d}k_j}{2k_i} = \frac{S!}{\prod_{i=1}^{d}(2k_i)!}$$

So the probability of returning to the origin is described by

$$(\frac{1}{2d})^S\sum_{\sum_{i=1}^{d} k_i = \frac{S}{2}}\prod_{i=1}^{d}\binom{2k_i}{k_i}\frac{S!}{(2k_i)!} = \frac{S!}{(2d)^S}\sum_{\sum_{i=1}^{d} k_i = \frac{S}{2}}\frac{1}{\prod_{i=1}^{d}(k_i!)^2}$$

Consider $\frac{1}{\prod_{i=1}^{d}(k_i!)^2}$ for $\sum_{i=1}^{d}k_i = \frac{S}{2}$. In a product $N!\cdot(N-m)!$ there are exactly $2N-m$ factors and as $N \geq (N-m+1)$ for $m \geq 1$, we have $(N-1)!\cdot(N-m+1)! \leq N!\cdot(N-m)!$. By this logic, we note that $\frac{1}{\prod_{i=1}^{d}(k_i!)^2}$

is maximal when $k_i = \frac{S}{2d}$ for every i. So (where p is the probability of getting to the origin in S steps):

$$\frac{S!}{(2d)^S}\binom{\frac{S}{2}+d-1}{d-1}\frac{1}{(\frac{S}{2d}!)^{2d}} \geq p \geq \frac{S!}{(2d)^S}\frac{1}{(\frac{S}{2d}!)^{2d}}$$

Applying Stirling we get:

$$\boxed{O(S^{\frac{1}{2}}) \geq p \geq \Omega(S^{(\frac{1}{2}-d)})}$$

Problem 3 This problem is NP-Hard. Yet it is not obviously in NP. A real number is not a short proof of a root, since it can take an arbitrarily large amount of memory to represent a real number.

We now show a reduction from 3SAT. For each literal A in our instance of 3SAT we map A to a^2 and A' to $(1-a)^2$. We represent disjunction by multiplication and conjunction by addition. We map a variable assignment to false in 3SAT to an assignment of the corresponding variable to 1, and an assignment of true to 0. For instance $\{A\},\{A,B,C\}\{A,D'\}$ maps to $a^2+a^2\cdot b^2\cdot c^2+a^2\cdot(1-d)^2$. We say that a formula is satisfiable if the polynomial has a root and unsatisfiable otherwise.

If an instance of 3SAT has a satisfying truth assignment then clearly the polynomial has a zero. Consider each group of multiplications that represent a clause. At least one of these terms must be zero since the clause must be satisfied and true literals are mapped to 0. Hence this mapping gives a zero of the polynomial.

Suppose our multi-variate polynomial has a zero. Since all summands in the polynomial are nonnegative, the only way for the polynomial to be zero is if every product of terms comprising a summand has at least one zero term. This can occur either because a variable is zero or a variable is 1 in the case of $(1-x)^2$. Hence, in a polynomial zero, we can consider variables assigned to 0, as assignments to true and 1 variables as assignments to false in the 3SAT instance, and clearly this gives a satisfying truth assignment. Thus, solving the real zeroes of a multivariate polynomial problem would solve 3SAT and hence the problem is NP-Hard and there is no efficent deterministic or randomizing algorithm for it unless P=NP.