Data Structures and Algorithms

Spring 2014

Cesar Agustin Garcia Vazquez

May 1, 2014

1 Problem Set 08

1. If we restrict the problems we look at, sometimes hard problems like counting the number of independents sets are in a graph become solvable. For instance, consider a graph that is a line on n vertices. (That is, the vertices are labelled 1 to n, and there is an edge from 1 to 2, 2 to 3, etc.) How many independent sets are there on a line graph? Also, how many independent sets are there on a cycle of n vertices? (Hint: In this case, we want to express your answer in terms of a family of numbers - like "For n vertices the number of independent sets is the n-th prime." And that's not the answer.)

   Similarly, describe how you could quickly compute the number of independent sets on a complete binary tree. (Here, just explain how to compute this number.) Calculate the number of independent sets on a complete binary tree with 127 nodes.

   Answer: For a linear graph, we can see in Table 3 the independent sets and the number of independent sets for a couple of scenarios. As we increase the number of vertices, we can see that the previous independent sets remain the same.

   Table 1: Independent sets for a linear Graph

Number of Elements	Independent Sets	Number of Independent Sets
0	$\{\}$	1
1	$\{\}, \{1\}$	2
2	$\{\}, \{1\}, \{2\}$	3
3	$\{\}, \{1\}, \{2\}, \{3\}, \{1,3\}$	5

   In Table 2 we stop displaying the previous independent sets, and we only consider the independent sets that can be generated with the new vertex added. For what we can see in the tables, the Fibonacci numbers appear as the result of the number of independent sets. If we consider $F_0 = 1$, $F_1 = 1$ and $F_2 = 2$, then we might

suggest that for a graph with n vertices, we have F_{n+1}, independent sets.

Table 2: Independent sets for a linear Graph

Number of Elements	Independent Sets	Number of Independent Sets
4	$\{4\}, \{1,4\}, \{2,4\}$	8
5	$\{5\}, \{1,5\}, \{2,5\}, \{3,5\}, \{1,3,5\}$	13
$\vdots$	$\vdots$	$\vdots$
n	$\vdots$	F_{n+1}

We can see that the base cases satisfied the proposed formula. So we can assume that for n vertices, the number of independent sets is F_{n+1}. When $n+1$, we have the first n vertices generate F_{n+1} independent sets, so now we just need to find out how many independent sets are generated when a new vertex is added. This is the same as computing how many independent sets can be generated when we have $n-1$ vertices and a new is added, i.e., we are able to add the new vertex in all of the previous independent sets, except for the one that contains the element n. Again by the induction hypothesis we have that for $n-1$, there are F_n independent sets and hence for $n+1$ vertices we have $F_n + F_{n+1} = F_{n+2}$ independent sets.

For the case when we have a cycle of n vertices, we use the same technique that in the previous proof. The last element added generates the previous independent sets plus the independent sets generated by considering all of the n elements but 2, the first and the last one.

Claim: The number of independent sets in a cycle is F_n

Table 3: Independent sets for a linear Graph

Number of Elements	Independent Sets	Number of Independent Sets
0	$\{\}$	1
1	$\{\}, \{1\}$	2
2	$\{\}, \{1\}, \{2\}$	3
3	$\{\}, \{1\}, \{2\}, \{3\}$	4
4	$\{4\}, \{2,4\}, \{1,3\}$	7
5	$\{5\}, \{2,5\}, \{3,5\}, \{1,4\}$	11
6	$\{6\}, \{1,5\}, \{2,6\}, \{3,6\}, \{4,6\}, \{1,3,5\}, \{2,4,6\}$	18
7	$\{7\}, \{1,6\}, \{2,7\}, \{3,7\}, \{4,7\}, \{5,7\}, \{2,4,6\}, \{2,4,7\}, \{3,5,7\}, \{2,5,7\}, \{1,4,6\}$	29

For this condition to be valid, we need to redefine the base case to make match to what we have. Then $F_0 = 1$, $F_1 = 2$, $F_2 = 3$ and $F_3 = 4$. So we assume that for n

vertices in a cycle graph, we have $F_{n-1} + F_{n-2}$ independent sets.

As the inductive step, for $n + 1$ vertices, for the first n vertices we can generate F_n independent sets. For the extra element, we can add the extra element to all the independent sets that are generated by $n - 1$ vertices because we have to remove the sets that are generated with the first element. So we have F_{n-1} independent sets. In total for $n + 1$ vertices we have $F_n + F_{n-1} = F_{n+1}$ independent sets.

2. Consider the problem MAX-K-CUT, which is like the MAX CUT algorithm, except that we divide the vertices into k disjoint sets, and we want to maximize the number of edges between sets. Explain how to generalize both the randomized and the local search algorithms for MAX CUT to MAX-k-CUT and prove bounds on their performance.

3. We know that all of NP-Complete reduce to each other. It would be nice if this meant that an approximation for one NP-Hard problem would lead to another. But this is not the case. Consider the case of Minimum Vertex Cover, for which we have a 2-approximation; that is, we are always within a factor of 2 of the optimal solution. We know (from the NP-Completeness notes, which you may want to check) that C is a cover in a graph $G = (V, E)$ if and only if $V - C$ is an independent set in V. Explain why this does not yield an approximation algorithm that is within a constant factor of optimal for Maximum Independent Set. That is, show that for any constant c_1, there exists a graph for which even if we obtain a 2-approximation of the Minimum Vertex Cover, the corresponding independent set is not within a factor of c of the Maximum Independent Set.

   The Maximum Independent Set problem and the Maximum Clique problem are related in the following way: an independent set of a graph G is a clique in the complement of G. (The complement of G is the graph that contains exactly the edges that are not in G.) Does an approximation algorithm (that is, an algorithm within a constant factor of the optimal) for the Maximum Clique problem yield an approximation algorithm (within a constant factor of optimal) for Maximum Independent Set?

4. Suppose we have a random walk with boundaries 0 and n, starting at position i. As mentioned in class, this can model a gambling game, where we start with i dollars and quit when we lose it all or reach n dollars. Let W_t be our winnings after t games, where W_t is defined only until we hit a boundary (at which point we stop). Note that W_t is negative if we are at a position j with $j < i$; that is, we have lost money. If the probability of winning and the probability of losing 1 dollar each game are $1/2$, then with probability $1/2$, $W_{t+1} = W_t + 1$ and with probability $1/2$, $W_{t+1} = W_t - 1$. Hence

$$E[W_{t+1} = \frac{1}{2}E[W_t + 1]] + \frac{1}{2}E[W_t - 1] = E[W_t], \tag{1}$$

where we have used the linearity of expectations at the last step. Therefore when the walk reaches a boundary, the expected winnings is $E[W_0] = 0$. We can use this

to calculate the probability that we finish with 0 dollars. Let this probability be p_0. Then with probability p_0 we lost i dollars, and with probability $1 - p_0$ we gain $n - i$ dollars. Hence

$$p_0(-i) + (1 - p_0)(n - i) = 0, \tag{2}$$

from which we find $p_0 = (n - i)/n$.

Buffy encounters this game in one of her visits to a hell dimension (or, worse, Las Vegas). Unfortunately, because it is a hell dimension, the game is not fair; instead, the probability of losing a dollar each game is 2/3, and the probability of winning a dollar each game is 1/3. Each victim starts with i dollars and gets to leave the hell dimension if they reach n dollars before going bankrupt. We wish to extend the above argument to this case. Show that $E[2^{W_{t+1}}] = E[2^{W_t}]$. Use this to determine the probability of finishing with 0 dollars and the probability of finishing with n dollars when starting at position i. Then generalize the argument to handle the case where the probability of losing is $p > 1/2$? Hint: Try using $E[c^{W_t}]$ for some constant c.