CSC373 – Problem Set 9

Remember to write your **full name(s)** and **student number(s)** prominently on your submission. To avoid suspicions of plagiarism: at the beginning of your submission, **clearly state any resources (people, print, electronic) outside of your group, the course notes, and the course staff, that you consulted**.

Remember that you are required to submit your problem sets as both LaTeX `.tex` source files and `.pdf` files. There is a 10% penalty on the assignment for failing to submit both the `.tex` and `.pdf`.

Ayushman Chhabra, 1004540696

Due TUESDAY, Dec 8, 2020, 22:00; required files: ps9.pdf, ps9.tex

Answer each question completely, always justifying your claims and reasoning. Your solution will be graded not only on correctness, but also on clarity. Answers that are technically correct that are hard to understand will not receive full marks. Mark values for each question are contained in the [square brackets].

You may work in groups of up to THREE to complete these questions.

[15 points]

The "Leaders of the Future" secondary school in Mississauga is organizing science competitions on subjects such as mathematics, programming and biology. The aim of the competitions is to promote scientific exchange and teamwork between students. As a result, the students are encouraged to team up in groups of five for each subject. As each competition on a certain subject will take place on a different day, a student can potentially belong to one or more teams, e.g., a student can belong to only a biology team or both a mathematics and a programming team.

The school has designed some specific rules which students need to adhere to during the competition days. Due to confidentiality issues, the school has decided to personally communicate with the students about these rules. To avoid interacting with each student, which in turn would lead to a large workload, the school administration wants to find a way of selecting as few students as possible such that every team across all the subjects has a student in the selected set.

You are given a set of students $S = \{1, 2, \dots, n\}$, and teams of students M, P, B with $|M| = m$, $|P| = p$, $|B| = b$, where for each team $T \in M \cup P \cup B$ we have that $T \subset S$ and $|T| = 5$. Note that any two teams in $M \cup P \cup B$ may overlap.

Your goal is to find a set of students $S' \subseteq S$ such that (1) each team in $M \cup P \cup B$ has has a student from S' and (2) $|S'|$ is as small as possible.

1. **(5 points)** Propose an approximation algorithm for this problem and show that it achieves an $O(\log n)$-approximation and runs in polynomial time in n, m, p and b.
   Solution
   We can use the algorithm that we talked about in lecture called GreedySetCover. Everything would be same, the only thing is that the GreedySetCover takes only one set of subsets so what we need to do we need combine M,P and B into the set 'H' and input them as one set. then the algorithm will just take the set 'G' which would maximize the number of students in 'G' who are also in other groups as well. Then just remove all of the groups which have at least 1 person who is in G. as shown in class, this can be done in $O(n^2k)$ time where n is the number of students and k is the number of given sets (k = m+p+b). it is also shown in class that this is a $(ln(n) + 1)$-approximation which is the same as $log(n)$-approximation solution which is what we are looking for as well.

2. **(10 points)** Propose an approximation algorithm for this problem and show that it achieves an 5-approximation and runs in polynomial time in n, m, p and b.

   Solution:

   Let's construct a undirected, acyclic graph G. For every student s_i in t_i teams, connect s_i to every other student $s_j \in$ team t_i where $s_i \neq s_j$.

   We will use a modified version of the greedy vertex cover algorithm covered in the lecture slides. The only difference would be that after choosing a random edge (u,v), instead of only adding the vertices u and v to set C, we would add all 5 vertices (students) who are part of the same team and then goes on to remove all of the edges which are incident to u or v.

   Proof:

   C = vertex cover

   let C* be an optimal vertex cover

   let A be the edges chosen by our algorithm

   we know that any vertex cover including the optimal solution must include at least one endpoint of each edge in A and no two edges in A share an endpoint.

   let OPT = $|C^*| \geqslant |A|$,

   Now since for each in A we add 5 vertices to C, we have:

   $|C| = |5A| \leqslant |5C^*| = 5 * OPT$

   Therefore we have an algorithm which achieves an 5-approximation.

   Runtime:

   This algorithm runs in $O(m)$ time where m is the number of edges in the graph. This is true since in the loop we at least remove one edge from the list of edges which means the worse case scenario happens when we go through the loop m times.