A-Question

Isak Nyberg
November 4, 2022

Problem Description (Same as for the C-question)

A teacher has 5 different sets of homework in list of students N students and has another list P different pairs of students who have worked together in previous courses. The task is to assign a different homework to each student so that the number of pairs $|P|$ where both students are assigned the different homework is maximised.

Problem Reduction (Same as for the C-question)

The problem can be modelled using graph theory, create a vertex for each student and for each pair of students create an edge between their corresponding vertex. The 5 different homework sets can be thought of as 5 different colours and the goal is to colours the graph without directly adjacent vertices having the same colour. Note that this graph is not necessary a planar graph and the 4 colour theorem does not apply.

0.8-Approximation (Same as for the C-question)

Given an optimal algorithm OPT for the problem, this algorithm would not necessarily assign each student a homework such that there are no former peers who gain the same homework, but rather it would maximize the number of peers who do not get the same homework. To create a 0.8 approximation of this the new algorithm ALG would need to avoid a collisions at least 0.8 of the times when OPT would do. This form of 0.8 approximation is done in the following algorithm.

Algorithm 1 0.8 Homework Assignments

Input: N

Assign a dummy homework to every student in N
for $n \in N$ **do** // Go through the list of students
  $h \leftarrow$ least common homework for pairs to n
  assign h to n // Assign the least common homework to the current student
end for

What this very short algorithm which takes the list of students N as the first argument. It begins by assigning a dummy homework to each student and then for each student it assigns the least common real homework out of the pairs to the student. If there are multiple different homework problems that are the least common it is sufficient to pick the first one that is the least common.

Proof of Correctness

Observations

Observation 1 (Same as for the C-question)

Given an optimal solution OPT it does not actually matter what homework that is assigned to each student but rather the relation between the assignments, in this regard if all students who were assigned homework h_i switched with students with homework h_j the solution would still be optimal. This means its actually the of assignments or two students who have worked together or edges in the graph between two students that matters instead of each individual assignment for each student. Additionally let P be the number of pairs of students who have worked together in the past, OPT can at most satisfy all of them and which means $P \geq OPT$

Observation 2

Even if all students are assigned a dummy homework at the start, by the end of the algorithm each student will have been given a real homework since the algorithm iterates over the entire class of students.

Correctness

Let OPT be the assignments that yield the most pairs of students that have not worked together in the past. Let ALG be the assignments that is given in the aforementioned algorithm. Let A_i be the assigned homework at the i'th iteration of the forloop.

Lemma 1: i'th student

Lemma: When in the loop and assigned a student with a homework and adding the student n_i to A_i the student can at most share homework with $\frac{1}{5}$ its pair.

Proof: Since the student n_i is given a homework it picks the least common on from all of its pairs. Since there are in total 5 different homework the worst case would be that every student that is paired with n_i would have an even spread of the 5 homework problems which would mean that regardless of which homework is assigned to n_i it would conflict with $\frac{1}{5}$ of their pairs. By contradiction it could also be argued that if the least common homework has more than $\frac{1}{5}$ conflicts it would mean that the sum of the sum of all of conflicts for each of the 5 other homework problems would exceed 1. ($\frac{1}{5} < h_1 \geq h_2, h_3, h_4, h_5 \implies \sum_i^5 h_i > 1$) which is a contradiction. Thus for each added student n_i to A_i the expected number of pairs from n_i that do not have the same homework is $E[n_i] = \frac{4}{5}$.

Lemma 2: i'th and j'th student

Lemma: When observing a pair between two students n_i and n_j and n_j was added to M after n_i was added with the corresponding homework h_i and h_j. The expected value $E[h_i \neq h_j] = \frac{4}{5}$.

Proof: Lemma 1 states that when a student is added to M it will not share homework with $\frac{4}{5}$ of the students it has been paired with. Since n_j was added after n_i, at the time when it was added to M it would satisfy lemma 1 and neither students' homework problems have been changed since this specific pair still hold that the probability of the two students not sharing the homework is $\frac{4}{5}$.

Lemma 3: M and ALG

Lemma: When the algorithm has terminated lemma 2 still hold and and the expected number of pairs which do not share homework is $E[h_i \neq h_j] = \frac{4}{5}$.

Proof: When when the algorithm has terminated for each pair between two students n_i and n_j where n_i was assigned a homework before n_j, at the time when n_j was assigned a homework and added to M lemma 2 would hold but it will also hold after the algorithm has terminated since neither homework assignment has changed since the time n_j was added to M. Thus for each pair it hold that $E[h_i \neq h_j] = \frac{4}{5}$ after the algorithm has terminated.

Conclusion

Lemma 3 proved that the expected value for each pair to not have a homework collision is $\frac{4}{5} = 0.8P$ where P is the number of pairs of students. It is also observed that $P \geq OPT$ which means that. $E[ALG] = 0.8P \geq 0.8OPT$ or in short, $E[ALG] \geq 0.8OPT$

Complexity

Since there *for*loop is iterated once for each students and for each student every pair needs to be examined, since each pair is between two students each pair will be examined twice. This means the runtime is $\mathcal{O}(2P|N|) = \mathcal{O}(P|N|)$ where $|N|$ is the number of students and P is the number of pairs of students.