

Mini projects

Students are required to carry out one mini project for the course COMP 314. As the number of projects is less than the number of students, more than one student may work on the same project. However, **each student must work individually in the project**. No more than 4 students may work on the same project.

To avoid conflicts in project selection, all students are required to choose 3 projects from the provided list of projects and update the following spreadsheet:

<https://docs.google.com/spreadsheets/d/1jZXt-lqOvJOWvTN8XZjb8b74cV9CcwhXMccXibZNbdI/edit?usp=sharing>

If there are projects selected by more than 4 students, projects will be assigned to the students based on their preferences.

Students will have to submit the following:

1. A short report explaining the algorithm used, containing screenshots of the output
2. Source code

At the end of the project, students will have to demonstrate their project on the specified date.

Note that the projects may be implemented as a console program. No GUI is needed.

Important dates

Project selection deadline	Oct 16, 2020
Project submission deadline	Nov 11, 2020
Project demo	Nov 12-13, 2020

Projects

1. Write a program to find the largest cycle in a given graph. You must also implement graph data structure. Hint: Use DFS.
2. Timetable generator: A school has 5 classrooms, 6 groups of students, each group requires to attend lectures of at most 4 hours in a day. The duration of the lectures is either 1 hour or 2 hour. Write a program to generate a weekly timetable for the school. Hint: use greedy algorithms.
3. n men and n women attend a party. Each man knows exactly k women, and each woman knows exactly k men (acquaintances are mutual). Write a program to find if it is possible to arrange a dance such that each man dances with a different woman that he knows. Hint: use max flow network (bipartite matching)

4. Write a program to extract keywords from a given text using the TextRank model described in the following article:
Mihalcea, Rada, and Paul Tarau. "TextRank: Bringing order into text." Proceedings of the 2004 conference on empirical methods in natural language processing. 2004.
<https://web.eecs.umich.edu/~mihalcea/papers/mihalcea.emnlp04.pdf>
5. Write a program to solve the Knight's tour problem using backtracking. Start with a chessboard of size 3x3, then increase the size and find the size of the largest chessboard your computer can solve. Record the time taken to solve the problem for each size of the board.
6. Write a program to solve the N-queen problem using backtracking. Start with a chessboard of size 4x4, then increase the size and find the size of the largest chessboard your computer can solve. Record the time taken to solve the problem for each size of the board.
7. Write a program to solve the 0/1 Knapsack problem.
8. Write a program to find the longest common subsequence between two sequences of characters.
9. Write a program to schedule a given list of preemptable jobs (in an operating system) with a deadline. Hint: use max flow network.
10. A travel agency has a list of flights. Each flight has a flight number, origin airport, destination airport, departure time and arrival time. Write a program to assist the travel agency in determining the earliest arrival time for the destination given an origin airport and start time. Hint: use a shortest path algorithm
11. Write a program to solve Sudoku using backtracking.
12. Implement a minimum/maximum priority queue using binary heap data structure.
13. Write a program to find closely related communities in a huge social network graph. Hint: find strongly connected components.
14. Write a web crawler that collects headlines of news from bbc.com. Start from bbc.com, crawl to different categories listed on that page (such as news, sport, worklife, travel etc.). Under each category there may be subcategories such as cricket, tennis, formula 1, football etc. under sport. Collect at most 15 news headlines from each subcategory. Hint: use BFS. You might need to throttle your requests to avoid rate limiting and getting blocked.
15. Write a program to generate a random graph, and find the average shortest path length in the generated graph. Hint: check Erdős-Rényi random graph model. If d_{ij} is the shortest path distance between nodes i and j in a graph with n vertices, average shortest

$$l = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}$$

path length is

16. Write a program to find the subsets of a given set of integers such that the sum of the elements in the subset is k for any value of k . Hint: use dynamic programming.