COMP5920M Scheduling
For revision class: May/June 2018 Exam paper

Question 1

A computer scientist has been successful in train crew scheduling research and now has attracted interest from a supermarket chain to develop algorithms for scheduling their staff.

(a) The train crew scheduling problem has some special characteristics in terms of the rail network covered, the timing of individual crew shifts and the inter-relationship between crew shifts. With reference to these special characteristics, discuss how the supermarket staff scheduling problem would differ.

[6 marks]

- Rail network covered: supermarket staffs work in a relatively confined location, they do not have complicated and time consuming travel plans during their shifts
- Timing of individual train crew shifts is very much constrained by the public timetable and the train unit schedule, there are also complex labour rules regarding work and breaks and their lengths. There are usually only a few shift patterns for all the supermarket staffs. Many staffs will be starting and ending their shifts and having their breaks at the same times.
- Train crew shifts relieve each other so that the train units they work on can continue their services uninterrupted. The number of supermarket staffs scheduled to be available during each time period is the most important consideration, the handing over between shifts is usually scheduled with groups of staffs and not individually.

(b) After some initial study of the supermarket staff scheduling problem, the computer scientist proceeds to collect data. Outline four main categories of data for train crew scheduling that the computer scientist would be familiar with, and comment on the suitability of such categorisation of data for the supermarket staff scheduling problem.

[6 marks]

The 4 categories of data are

- work to be covered by personnel to be scheduled
- Hard rules
- Soft rules
- Preferences

Yes, the above categories of data are generally applicable. However, the nature and scale of the data could be significantly different between the two scheduling problems.

(c) The project went well and the supermarket chain commissioned a pilot trial of the newly developed algorithms. The company assigned a senior manager to assist in this pilot trial project, who was able to provide any data required and answer queries from the computer scientist. At the end of the first phase of the trial, the computer scientist delivered their initial results and was jubilant that a saving of about 10% was achieved. Unfortunately, the manager rejected the results because some duties in the schedule broke the rules and some duties could not be implemented 'in practice'. It turned out that the computer scientist misinterpreted one labour rule and that the manager had not described all the local practices. It was difficult to blame the manager, to whom staff scheduling was a totally intuitive process.

Analyse and comment on the difficulty in setting up the input data correctly in an industrial pilot trial of scheduling optimisation research. Suggest to the computer scientist three actions by which the difficulty might have been avoided.

[8 marks]

- The computer scientist and the manager have very different professional backgrounds and world views that may lead to communication problems and misunderstanding, which are attributable to the multi-disciplinary nature of scheduling.
- Problem modelling capability of scheduling systems would not be perfect, real life scheduling problem instances often involve many local practices and problem scenario variations that could be overlooked
- Careful study of all available operations manuals and documentation
- Careful study of available existing and past schedules
- The researcher and the planner should go over the scheduling system and the input data set up together to the finest level of detail that is practicable

[question 1 total: 20 marks]

Question 2

(a) For the Selection Phase of the Generate and Select approach for train crew scheduling, an ILP method was discussed in lectures and the design and implementation of greedy heuristics were the focus in coursework 2. In terms of the problem model, practicality for real life problem instances, computational performance and optimality of solutions, compare and contrast the ILP method and the greedy heuristics.

[8 marks]

- Both methods seek solutions based on the set covering model
- General purpose generic ILP solvers would not be capable of solving real life train crew scheduling problem instances within practical time, good heuristics are needed for finding an all integer solution. Greedy heuristics also need very good intuition in their designs to yield solutions of reasonable quality.
- ILP can theoretically be solved to optimality if an optimal solution exists, even though it may take a long time to compute. Greedy heuristics are computationally very fast, but optimality is not guarantee and often not attainable.
- The ILP solver may not be able to return any solution at all within practical time limits. Greedy heuristics will almost always be able to return a feasible solution (unless none exists) in a relatively very short time.

(b) The cost of a train crew shift is expressed in terms of the duration (in minutes) from the time the crew is required to sign on to the time the crew signs off. With reference to the second coursework assignment, one student has designed a greedy heuristics for train crew scheduling which has yielded a solution schedule with 191 shifts at a total cost of 62494 minutes for one of the given datasets. The train company concerned expects a schedule of around 70 shifts. A normal full-time train crew shift would be around 9 hours long. The student is going to analyse their solution for clues to improve their heuristics. Discuss and suggest two areas for the student to consider in relation to the quality of the crew schedule they have obtained above.

[4 marks]

The yielded schedule is clearly very poor

- The average cost of a shift in the yielded schedule is about 5 hour 27 mins, far below 9 hours. There would be too many very short shifts in the schedule
- The heuristic might have resulted in redundant shifts in the schedule, which might not have been removed giving rise to a high total shift count.

(c) A research student is investigating the graph colouring approach for train crew scheduling. They have set up a graph in which each vertex represents a train trip in the given timetable to be scheduled. Graph edges have been inserted such that any pair of train trips (graph vertices) joined by an edge will have at least part of their journeys running at the same time. A graph colouring solution is computed and each set of trips of the same colour label are arranged into a chronological sequence. A train crew is assigned to cover from the beginning of the trip sequence until it reaches the maximum length allowed. The sequence of trips is then split for more crews to cover the remaining work. Explain four reasons why the solutions obtained would likely be infeasible.

[8 marks]

Note that this question is about solution feasibility, points concerning solution efficiency is not relevant, e.g. very short shifts may result but they are still feasible.

- there may be no suitable gaps for meal breaks
- a crew may have arrived at a certain location and the next trip to serve is departing at a distant location and there is not enough time for the crew to travel
- the trips connected by the same graph colour label may have incompatible route/traction knowledge requirements
- some auxiliary tasks needing crew cover such as (im-)mobilising the train units, (de-)coupling the train units have not been included in the model

[question 2 total: 20 marks]

Question 3

(a) Outline how an initial matrix would be constructed for Floyd’s shortest path algorithm to be applied to compute the empty running times between any pair of locations, including the terminals, intermediate stops and depots, in a bus network.

[5 marks]

- Enter all the journey times between bus stops from the route information
- Symmetrically fill in the journey times for the reverse directions, overridden by any known exceptions
- Reduce the journey times entered by an estimated rate because the journeys would be faster without passengers getting on or off.
- Enter values for known shortcuts, e.g. bypassing a housing estate, use of motorways etc.
- Finally, fill any remaining empty cells in the matrix with a very high constant value

(b) Explain the key differences between bus vehicle scheduling and train unit scheduling.

[8 marks]

- Bus movements are relatively more flexible on the roads than train units running on tracks. Therefore the scheduling of train units has to be careful in avoiding blockage between train units on the tracks.
- Each bus is scheduled individually, but multiple train units can be coupled and decoupled. High passenger demand can be satisfied by a more frequent service in bus operation. But in train operation, higher capacity can be provided through train unit coupling instead of increasing the service frequency.
- When there is imbalance of vehicle resources across a transport network, the time required to relocate a bus can be more accurately estimated than that for relocating a train unit. This is largely because train tracks clearance is required for train units to do a non-passenger carrying train trip. A train unit may also be relocated by coupling on to a timetabled passenger carrying train trip.

(c) Although the Size Limited Iterative Method (SLIM) algorithm discussed in this module is a heuristics, it has a close working relationship with an exact integer linear programming solver; describe this relationship.

[7 marks]

- SLIM is an iterative heuristics, which aims at deriving a minimal directed acyclic graph (DAG) sufficient for an optimal solution to be found by an ILP solver.
- A crude initial solution is found by heuristics.
- In each iteration, the arcs used by the current best solution (or a few of the recent best solutions) are extracted to form an “essential DAG”
- Some additional arcs are added to the essential DAG based on some search intensification strategies for the current iteration, this results in an enlarged “augmented DAG”
- The augmented DAG gives rise to opportunities for improved solutions to be found by the ILP solver, but its size is controlled so that the ILP solver would be able to find a solution quickly.
- Multiple augmented DAGs are formed by variations of the search intensification strategies applied at each iteration. Each different augmented DAG is solved by the ILP solver in parallel computation.
- The parallel computation and their results are coordinated by SLIM, and the iterations are continued until some termination criteria are reached

[question 3 total: 20 marks]

END