Project Proposal: University of Nottingham Final Exam Scheduling

Instructor: Dr. Bonnifonte

Team members: Kien Le, Jenny Nguyen, Dionne Phan

1. Introduction and Motivation

During exam seasons, many students often complain about their packed schedules. As one student once put it, "My exam schedule is so bad. I have to take three consecutive exams from morning until dinner, and then submit a final project that is due the next morning at 10 a.m." (Kien, during his sophomore year). This problem is not unique, since almost everyone at Denison has experienced similar situations. As a result, this raises the question: Why can't the Registrar design a more balanced final exam schedule that reduces stress and gives students more time to prepare?

To explore this issue, we decided to dive into the principles and challenges behind final exam scheduling. Our initial goal was to find a dataset from Denison that represents the requirements and constraints of its exam scheduling process. However, we were unable to find such data. Instead, we found an interesting and relevant dataset: the final exam schedule from the University of Nottingham. Upon examining it, we realized that it represents a typical exam timetabling problem that aims to create a schedule satisfying all students while considering constraints such as available rooms, departmental priorities, and the limited time period for exams.

Through this project, we hope to gain a deeper understanding of how exam scheduling systems work. Is my "bad" schedule simply the result of bad luck, or is the Registrar's time scheduling model just a bad one that cannot work well within complex constraints and needs to be improved? By studying this problem, we aim to understand the logic, trade-offs, and challenges behind creating a "good" exam schedule.

2. Data and Data Structure

a. Data source

The exam timetabling dataset was obtained from the University of Nottingham and is publicly available for research purposes. It was originally introduced by Burke, Newall, and Weare (1996) in their foundational study on university examination scheduling and has since become a standard benchmark within the exam timetabling community. The dataset represents real scheduling data from Semester 1 of the 1994 - 1995 academic year at the University of Nottingham and includes detailed information on students, courses, exams, and enrolments. In total, it contains 800 exams offered by 46 departments, 7,896 students, and 33,997 enrolments, which is equivalent to an average of 4.3 exams per student and 42.5 students per exam.

A defining feature of the dataset is its room capacity constraint, whereby the total number of students assigned to exams in any given timeslot must not exceed 1,550. The principal objective is to minimize the number of students scheduled to sit two consecutive exams on the same day. In a subsequent extension, the objective is expanded to also avoid cases of consecutive overnight exams (Burke & Carter, 1998).

Because it incorporates real-world institutional data, practical scheduling constraints, and well-defined optimization objectives, the University of Nottingham dataset offers a rigorous, transparent, and reproducible benchmark for evaluating exam timetabling algorithms and models.

b. Data Structure

The University of Nottingham exam timetabling dataset is provided in text format and consists of three main files: **students**, **exams**, and **enrolments**, along with an additional file named **data** that specifies the necessary conditions for course scheduling.

1. **students**: This file contains information on student enrollment in courses. It has two columns:

- Student code: a 10-character unique identifier for each student.
- Course code: the code of the course the student is enrolled in.

2. **exams**: This file provides details about each exam and consists of four columns:

- Exam code: a unique identifier for the exam.
- Course name: the name of the course associated with the exam.
- Exam duration: the length of the exam in minutes.
- Department code: the code of the department responsible for the exam.

3. **enrolments**: This file links the **students** and **exams** datasets. It contains two columns:

- Student code: the unique identifier for a student.
- Exam code: the unique identifier for an exam in which the student is enrolled.

4. **data**: This file is not a dataset but specifies the necessary conditions and constraints for course scheduling, such as room capacities, allowable time slots, and other parameters and special constraints required for constructing the timetable.

3. Problem statements and modeling idea

The goal of this project is to develop an optimized exam timetable using the University of Nottingham exam dataset. Specifically, we aim to create a schedule that satisfies the necessary conditions and constraints set by the university, individual departments, and students to the greatest extent possible. This includes ensuring that room capacities are not exceeded, minimizing instances where students are scheduled for two consecutive exams on the same day or overnight, and respecting any additional requirements specified in the dataset. The resulting timetable should balance operational feasibility with student convenience, producing a fair and practical schedule for everyone.

a. Decision variables

Let C be the set of courses offered by the university and T be the set of available exam time slots. We define a binary decision variable $x_{i,j}$ for each course $i \in C$ and each time slot $j \in T$ as

$x_{i,j} = 1$ if the exam for the course i is scheduled in the time slot j, and $x_{i,j} = 0$ otherwise. Since the scheduling period spans two weeks, with three time slots per day from Monday to Friday and one time slot on Saturday, it results in a total of 32 time slots. Given that there are 800 courses to be scheduled, the model will have a total of 25600 binary variables of this type.

Let R be the set of available rooms. We define another binary decision variable $y_{i,k}$ for each course $i \in C$ and each room $k \in R$ as $y_{i,k} = 1$ if the exam for the course i is scheduled in the room k, and $y_{i,k} = 0$ otherwise. Given that there are 16 rooms, there will be a total of 12800 binary variables of this type.

In addition to these variables, some constraints might also require some additional variables that are not listed here.

b. Objective

Our objective will become progressively more refined as we continue building our model. We plan to split it into 3 stages. Initially, the objective is to minimize the number of students who have conflicting exams (two exams scheduled in the same time slot), as they would be unable to attend both. Ideally, our optimal solution would yield a value of 0, meaning that every student can take all their exams without conflicts. Next, we aim to improve our objective by minimizing the number of students who have two or more exams on the same day. Achieving an optimal value of 0 in this stage would imply that no student has to take multiple exams in a single day, ensuring that each student has at least one night of rest before every exam (and avoiding the kind of suffering like the student mentioned in the introduction). Finally, we also want to minimize cases of consecutive overnight exams.

c. Constraints

There are several constraints in this problem, with the two most important:

- **Exam duration constraint**: Each exam must be assigned to a time slot that is at least as long as the exam's required duration. For example, a course requiring a 3-hour exam cannot be assigned to an afternoon or evening time slot if those slots only allow for 2 hours.
- **Room capacity constraint**: Each exam must take place in a room large enough to accommodate all enrolled students, and at the same time, no more than 1,550 students can take the exam.

In addition to these, several other constraints apply only to specific subsets of courses or rooms:

- Certain courses must take place in specific rooms (e.g., art courses that require studio spaces).
- Some courses must occur during specific time slots (e.g., only morning sessions, during the first exam week, or specifically on Thursdays).
- Some courses have scheduling relationships with others: for example, two exams may need to be scheduled at the same time, or one immediately after another.

Furthermore, some courses have priority preferences, meaning they should be scheduled as early as possible. While this is not a hard constraint, it needs to be incorporated into the constraints or the objective function to guide scheduling decisions. In addition, some rooms can be combined to form larger spaces, allowing exams with more students to take place simultaneously.

4. Expected Outcome

The expected outcome of this project is a complete exam timetable that assigns all exams to specific timeslots and locations while satisfying the given constraints. To achieve this, we plan to build and refine the optimization model iteratively. Initially, the model will ensure that each student can complete all their exams without any scheduling conflicts and that the assigned rooms have sufficient capacity to accommodate the exams. Subsequently, additional constraints such as minimizing consecutive exams for students and respecting departmental or university-specific requirements will be incorporated step by step. And finally, no students will have exams on two consecutive days. This gradual approach will allow us to develop a feasible, practical, and well-structured exam schedule that meets operational requirements and maximizes student convenience.

References

Burke, E., & Carter, M. (Eds.). (1998). *Practice and Theory of Automated Timetabling II: Second International Conference, PATAT'97, Toronto, Canada, August 20 - 22, 1997, Selected Papers.* Springer.

Burke, E.K., Newall, J.P., Weare, R.F. (1996). A memetic algorithm for university exam timetabling. *In: Burke, E., Ross, P. (eds) Practice and Theory of Automated Timetabling.* PATAT 1995. Lecture Notes in Computer Science, vol 1153. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-61794-9_63

Qu, R., Burke, E. K., McCollum, B., Merlot, L.T. G., & Lee, S. Y. (2009). A Survey of Search Methodologies and Automated System Development for Examination Timetabling. *Journal of Scheduling*, *12*(1), 55-89. 10.1007/s10951-008-0077-5