# An Investigation into the Efficacy of Hyper-Heuristics for Middle Grade Doctor Shift Scheduling

Dissertation

**Author**:
Daniel Ferring

**Supervisor**:
Warren Jackson

School of Computer Science

University of Nottingham

I hereby declare that this dissertation is all my own work, except as indicated in the text: D.F.

## Abstract

The NHS is in a state of turmoil, suffering from chronic understaffing and employee dissatisfaction. It is therefore of the utmost importance that staff timetabling makes the most effective use of the workforce whilst striving to satisfy their preferences. Hyper-heuristics are an AI method and have been previously applied to other timetabling problems. They provide a layer of abstraction between themselves and the problem domain, enabling the same hyper-heuristic algorithm to be used across multiple types of problem. This project aims to investigate whether existing hyper-heuristic algorithms could be applied to the problem of Middle Grade Doctor Shift Scheduling to produce superior timetables than those made by hand.

To this end, current Middle-Grade Junior Doctors and timetable makers were consulted, in order to produce an accurate formal definition of the problem that seeks to maximise shift coverage while adhering to doctor preferences. A Hyflex-compatible problem domain was encoded, and an existing hyper-heuristic algorithm was run on it. Through a study comparing the quality of timetables produced by human participants to those produced by the hyper-heuristic for the same problem instances, the project proves that the hyper-heuristic approach soundly outperforms manual methods - significantly improving shift coverage and preference adherence. Further experimentation proved the hyper-heuristic's ability to find high-quality solutions for a variety of instances, as well as its potential use as a tool for modelling a department's ability to cover shifts, laying the groundwork for future research.

## Acknowledgements

Firstly, I would like to thank my supervisor for his support throughout the project; without his persistently brilliant advice the project as it stands today would not exist, and I am incredibly grateful that he agreed to undertake this role for me.

Secondly, I would like to thank my friends, in particular Ben Galbraith and Nadav Rahimi, for putting up with more discussion about timetabling than they probably thought was possible.

Next I would like to thank all the participants for the comparative study, without whom the central aim of the project would not have been possible to evaluate.

I would also like to thank all of the NHS workers who provided invaluable input and who work each day for the benefit of others: your service is truly admirable.

Finally I would like to thank my mother, for providing the inspiration for this project.

# Contents

# Chapter 1

# Introduction

## 1.1 Background and Motivation

The NHS is facing an existential crisis; the service is under higher pressure than ever before while trying to cope with staff shortages that are only set to increase [1]. To cover demand and make up for the lack of permanent staff, NHS departments are increasingly becoming dependent on locum doctors[1] [2], often at great expense [3]. As the use of locum doctors is primarily driven by a shortage of permanent staff [2], it is clear that it is of the upmost importance that permanent staff are used as effectively as possible. One way to do this is through automated staff rostering [4], something that has been shown to reduce the administrative time spent planning rosters, minimise unused working hours, and improve the consistency of staffing levels [5].

Healthcare staff scheduling is a well-studied problem with a large emphasis on the nurse rostering problem [6]. A number of approaches have been taken to try and solve the problem, including hyper-heuristic methods [7]. Despite the widespread attention to the nurse scheduling problem, there has been less attention paid to other healthcare professionals such as Middle Grade doctors[2] [6]. Given the similarities between the problems of nurse rostering and scheduling middle grade doctors, it stands to reason that some of these techniques could be reused, if a suitable encoding of the problem were to be provided.

The primary difficulty for encoding the problem of middle grade doctor scheduling comes from the large number of constraints required for a solution to be feasible. Any schedule must adhere to both the European Working Time Directive [8] and the 2018 Junior Doctor Contract Refresh [9] which set out rules for how long a doctor can work, the number of consecutive shifts they can take, and the breaks needed after certain shifts or stretches of shifts. Additionally, each individual middle grade doctor will have training sessions they need to be able to attend, as well as scheduled annual leave, with certain shifts also requiring a specific class of middle grade doctor.

In terms of techniques to reuse, hyper-heuristics are a particularly pertinent choice: they provide a layer of abstraction between themselves and the problem domain, and should be applicable to any type of problem, if a suitable representation and set of low-level heuristics is provided [10]. Hyflex [11] is a software framework for the development of cross-domain search methodologies created by researchers at the University of Nottingham for the Cross-domain Heuristic Search Challenge (CHeSC) [12]. It provides a common interface for a number of problems with the aim of providing a benchmark for the generality of heuristic search methodologies. Encoding a problem in this way will allow any algorithms previously made with the framework to be applied to it. As part of CHeSC, a number of hyper-heuristic algorithms were produced, some of which are publicly available on GitHub and can be reused for other projects [13].

[1] See Section 2.3

[2] Ibid.

## 1.2 Aims and Objectives

The primary aim of the project is therefore to explore whether the task of junior doctor shift scheduling is something that can be automated using an existing hyper-heuristic, in a way that utilises the resources of the NHS more efficiently, while also better adhering to the needs of individual workers. To this end, the project can be split into two main objectives: (1) to provide a formal definition of the problem and (2) to investigate the efficacy of an existing hyper-heuristic algorithm for solving it.

The aims for the problem definition are as follows:

1. a) To define the problem in such a way that shift coverage is maximised, while adhering to legal requirements.
   b) To take into account the needs of the doctors, in order to attempt to improve their well-being.

Formulating the problem in this way should ensure that improvements in the efficiency of the system do not negatively impact those working in it, something that would inevitably have negative consequences in the long term.

The aims for investigating the efficacy of reusing an existing hyper-heuristic are as follows:

2. a) To prove that a hyper-heuristic approach can outperform manual methods
   b) To investigate whether the generality of hyper-heuristic methods lends itself to dealing with the idiosyncrasies of individual hospital departments
   c) To investigate whether the hyper-heuristic system could be used to model the impact of certain parameters on a department's ability to cover shifts

## 1.3 Outline of Work

### 1.3.1 Defining the Problem

It had initially been hoped that it would be possible to obtain information directly from the NHS itself, as being given instances of real-world timetables, as well as direct feedback on the problem formulation, would have greatly simplified the process of completing this project. Unfortunately direct involvement with the NHS was not possible, and the project still needed data for the problem formulation to have any hope of resembling real world scenarios. An inability to work directly with the NHS, however, did not negate the possibility of working directly with NHS workers.

Hence, the process of defining the problem entailed the following steps:

1. Surveying the junior doctors themselves, in order to obtain the information necessary to create a definition that reflects their needs (Objective 1.b)

2. Surveying and interviewing the timetable makers, in order to obtain a better understanding of timetable structure and requirements beyond those stipulated by law (Objective 1.a)

3. Providing a formal definition of the problem by defining hard constraints and an objective function based on the results of steps 1 and 2 (Objectives 1.a and 1.b)

### 1.3.2 Investigating the Efficacy of Hyper-Heuristic Methods

As mentioned in Section 1.3.1, the original hope was to obtain information directly from the NHS. Had this happened, the produced system could have been tested on real world instances. As this was not possible, it was instead necessary to write a program capable of generating realistic problem instances based on information obtained in the surveys and interviews.

Section 1.4.2 outlines how using pre-existing hyper-heuristics could reduce the overhead needed to automate timetabling. To prove that this is possible, this project provides a Hyflex-compatible problem domain and explores the performance of an existing Hyflex hyper-heuristic on the newly defined problem.

The first avenue of experimentation is a study comparing the quality of timetables produced by human participants to that of those produced by the automated system. Further experimentation has also been undertaken to ascertain the ability of the system to handle issues like understaffing, in order to understand the strengths and weaknesses of this approach. The results of this experimentation are also used to argue the applicability of the system for modelling shift coverage.

The steps taken for this part of the project are therefore:

1. The creation of an Instance Generator
2. The implementation and design of a Hyflex-compatible problem domain
3. A comparative study investigating the quality of timetables produced manually to those produced by a hyper-heuristic algorithm (Objective 2.a)
4. Experimentation with a variety of instances in order to see the impact of specific instance characteristics on the objective function values achieved (Objectives 2.b and 2.c)

## 1.4 Related Work: Heuristic Approaches to Nurse Scheduling

As discussed in Section 1.1, there is a large body of work exploring the Nurse Rostering Problem (NRP); due to the previously mentioned similarity to the problem of scheduling middle grade doctors, this literature review will primarily focus on approaches to solving NRP.

### 1.4.1 Exact Methods for NRP

There are many examples where optimisation techniques such as integer programming, branch and bound, and constraint programming have been applied to the problem. For instance Rahimian, Akartunalı, and Levine proposed a hybrid integer and constraint programming approach [14]. Their approach combines the strengths of Integer Programming for finding lower bounds and optimum solutions, with the capability of Constraint Programming for finding feasible solutions, in a manner that allows high-quality solutions to be found in a relatively short amount of time.

Hu et al. used a branch-and-price approach for the Nurse Rostering Problem with multiple units [15]. The aim was to overcome the difficulties presented to certain sub-problem formulations by the fact that most constraints in NRP are soft. They also present various acceleration strategies to improve the branch-and-price algorithm, making it more suitable for large problem instances. A comparative study was also conducted with some commercial solvers, with their approach proving to be capable of producing competitive or improving solutions.

He and Qu propose a constraint programming based column generation approach to NRP, and present a model that formulates the real-world constraints seen in several of the benchmark Nurse Rostering Problems [16]. The overall problem is decomposed into a pricing sub-problem and a master problem.

Constraint Programming is used to generate patterns with desired features of a problem solution, and Column Generation is used to select a subset of generated patterns to construct the best complete solution to the problem. Two strategies are presented to generate high-quality, diverse columns, which result in a faster convergence in linear relaxation while satisfying the integrality request of the master problem.

### 1.4.2 Justifying the Hyper-Heuristic Approach

While exact methods are applicable to the problem of nurse scheduling, there are drawbacks; the primary one being a lack of generality: these optimisation algorithms are intrinsically linked to the specific use case they were designed for, meaning that departments with even a slightly different modus operandi would require the development of an entirely new system to solve their timetabling problem, something that would incur significant overhead.

As discussed in section 1.1, it should be possible to reuse hyper-heuristics on multiple problems, given a suitable encoding of the problem in question. For healthcare scheduling, this would mean that a general purpose hyper-heuristic could be reused across departments, in tandem with a problem domain tailored to each department's specific needs. This would require significantly less overhead to implement when compared to the aforementioned optimisation techniques. As such, the remainder of this literature review will focus on the applicability of hyper-heuristics to healthcare scheduling as well as a selection of heuristic approaches containing ideas that can be incorporated into this project's problem domain.

### 1.4.3 Applicability of Hyper-Heuristics to Healthcare Rostering

Hyper-heuristics have successfully been applied to NRP. Asta, Özcan, and Curtois developed a tensor based hyper-heuristic that improved upon the best-known solutions for four real world instances [17]. Their approach consists of an online-learning hyper-heuristic that makes use of tensor analysis to find relationships between low level heuristics in order to influence their selection at various stages within the search process.

Kheiri et al. created a hyper-heuristic approach based on the hidden Markov model for a modified version of the nurse rostering problem [7]. In addition to a very in-depth definition of the problem, the paper outlines the low level heuristics used by the hyper-heuristic; while some are specific to the problem of nurse rostering, the principles of others could be repurposed for scheduling middle grade doctors.

Bilgin et al. proved the applicability of a single hyper-heuristic algorithm to multiple healthcare scheduling problems [18]. The paper outlines a general hyper-heuristic approach before testing it on two problem domains: the Nurse Rostering Problem and the Patient Admission Problem. The hyper-heuristic approach performed well on both domains, indicating that the generality of this type of optimisation approach extends to healthcare scheduling problems.

### 1.4.4 Relevant Heuristic Approaches

Dowsland and Thompson proposed a hybrid approach in which two integer programming models are used in tandem with a heuristic method [19]. Of particular novelty is their use of a knapsack integer programming model to determine whether locum doctors are required and, if so, how many are needed. Their description of the nurse scheduling problem also takes into account the requirement of specific grades for specific shifts, a requirement that this project will also seek to address.

A hybrid approach using heuristic ordering and variable neighbourhood search was formulated by Burke et al. [20]. Constraints are evaluated to be either soft or hard, with soft constraints having an associated weight and penalty function. The formulated objective being to find a solution that satisfies all hard constraints while minimising the penalty caused by violations of the soft constraints. Heuristic ordering is used to create the initial solution: shifts are evaluated and ordered based on how likely they are to cause a penalty. The shifts are then assigned in order to each available nurse, with the most troublesome being assigned first, and the penalty calculated; the nurse causing the lowest penalty is then assigned to the shift in the solution. A variable neighbourhood search is then used to find improving solutions.

Wong, Xu, and Chin explored the use of a two-stage heuristic approach [21]. As part of the process of identifying the soft constraints of the problem, a survey was conducted in order to facilitate some measure of worker preferences. While this would be an effective way to make the solutions more conducive to employee satisfaction, new surveys and recalculated weights would be required for any practical application to a different hospital: there would almost inevitably be an incongruence in the preferences of different nurses within differing departments. A shift assignment heuristic is used to develop an initial schedule that satisfies all of the hard constraints, with a sequential local search being applied consecutively to find improving solutions that satisfy more of the soft constraints.

### 1.4.5 Fairness with Regards to Preferences

An issue when trying to automate the allocation of timetables in a way that takes preferences into account is the difficulty in ensuring that the adherence to preferences is spread fairly between those impacted by the generated schedule. Ouelhadj et al. noted the lack of research into fairness for Nurse Rostering and proposed a new set of fairness measures to address this [22]. Tests were run on existing data sets with a modified objective function and an Agent-Based Cooperative Meta-heuristic; it was found that fairer timetables were produced without a reduction to their quality with respect to other constraints, leading to the conclusion that fairness should be adopted in future work.

Fairness has also been explored in other timetabling contexts, such as exam timetabling. Muklason et al. present research into preferences relating to this area, having conducted a survey of students that identified fairness as a top priority [23]. A number of approaches to measuring fairness are suggested and tested, including Jain's Fairness Index. The approach using Jain's Fairness Index was found to be able to balance fairness with producing high quality timetables. This paper has two implications for the project: firstly, it demonstrates the validity of using surveys to formulate aspects of a problem's objective function, and secondly, it proves the applicability of Jain's Fairness Index to timetabling, an approach that this project will consequently use.

### 1.4.6 Conclusion

This literature review makes it clear that there are numerous heuristic and hyper-heuristic methods that have been successfully applied to the problem of scheduling healthcare scheduling. It is also evident that hyper-heuristics can be reused effectively on multiple types of scheduling problem. From this we gain a justification for the project: by suitably defining the problem and creating a compatible problem domain, it should be possible to apply existing hyper-heuristics to the problem domain and obtain well-performing solutions. In addition to a justification, these papers provide inspiration for low-level heuristics and initialisation techniques that can be used within the problem domain. In particular, the two stage approach will be used within this project, with a heuristic initialisation method being used to obtain a feasible initial solution for the hyper-heuristic to then improve upon. Finally, Jain's Fairness Index will be used to ensure a fair distribution of doctor preferences being granted.

# Chapter 2

# Formally Defining the Problem

As the problem of Middle-Grade Junior Doctor timetabling has not previously been given much attention, this project attempts to provide a concrete definition of the problem. This chapter outlines the steps taken to gather the requisite information, as well as the final definition used for the encoding of the project's problem domain.

## 2.1 Methodology: Surveys and Interviews

In order to obtain the information necessary to properly define the problem, the following data gathering activities were conducted: (1) a survey of current Middle Grade Junior Doctors, (2) a survey of staff currently making timetables, and (3) an interview of staff involved in timetable making.

The surveys were conducted using anonymous online questionnaire forms hosted on Microsoft Forms. Participants were recruited through personal WhatsApp groups; participation was voluntary and performed in their own time. The questions themselves were somewhat open ended, to account for my lack of experience in, and knowledge of, the field, but in no way asked for personal information: it was made very clear that this was not desired.

The interviews were performed on Microsoft Teams; data was collected in the form of notes written during the interview. In this way sanitisation of the information was conducted at the point of collection, as I only recorded information relating to the project and avoided recording information that could be used to identify those that participated. Recruitment in this case came from the participants themselves: both staff members got in contact after completing the survey to say that they would be happy to talk to me directly if extra information was required.

For more information of how research ethics were applied to the data collection aspect of the project, see Section 4.4.2.

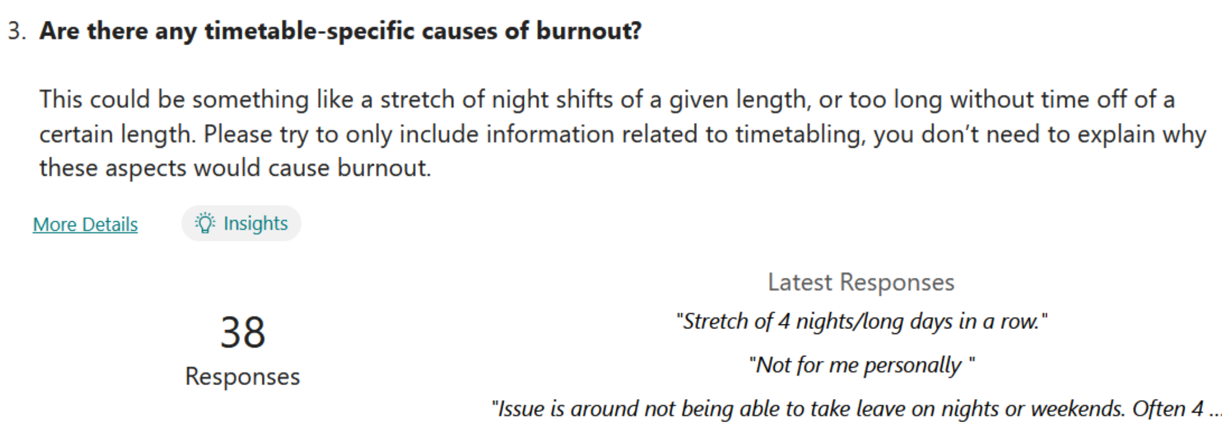



Figure 2.1: An example question posited in the Junior Doctor Survey

### 2.1.1 Junior Doctor Survey

The survey of junior doctors asked questions relating to the following aspects of timetabling:

1. Their preferences regarding working hours
2. Specific causes of burnout relating to timetabling (e.g. too many nights in a row)
3. How they choose to take leave (e.g. sporadic weekends, a long stretch etc.)
4. How training impacts their timetable
5. Any other aspects that they felt should be considered

For the survey of junior doctors, data was almost entirely qualitative. The aim with aspects 1-4 was to observe patterns in their preferences/experiences in order to inform the creation of a program capable of generating realistic instances. To this end, content analysis was performed on responses relating to these 4 areas, in order to identify the patterns, and to convert the information into more usable quantitative information. Given the very open ended nature of aspect 5, thematic analysis was employed, to try and get a grasp of general aspects of timetabling that they found important. All of the data collected from this survey was used to inform the problem definition.

### 2.1.2 Timetable Maker Survey

The survey of the timetable makers asked questions relating to the following aspects of timetabling:

1. The average planning period for doctor timetables
2. Any heuristic methods used when allocating shifts
3. The length of time needed to create a timetable manually
4. Their experience of working hour preferences
5. How they address doctor preferences
6. Any considerations, beyond limiting their use that should be made when allocating locum shifts
7. The ideal output of an automated timetabling system

The data collected from the timetable maker survey was a bit more varied in nature: aspects 1 and 3 were inherently quantitative; 4 was qualitative, and suited to content analysis; aspects 2, 5, 6, and 7 were qualitative and best suited to thematic analysis. Aspects 1 and 4 were intended to help with the problem definition. Aspect 1 was of particular importance, as it informed the length of the planning period for the problem instances used during the evaluation of the hyper-heuristic system. Aspect 2 was meant to help provide ideas for new heuristics to use in the problem domain, making use of the timetable makers' high degree of domain knowledge.

Aspect 3 was intended to help justify the project, and to provide an extra element of comparison when evaluating the performance of the automated system: the hope was that it would be possible to prove not only that the system could create better timetables, but that it could also do so in a shorter time frame. Aspect 5 was intended to find methods for ensuring an even spread of preferences being granted between doctors, while aspect 6 was intended to obtain information that could be used when formulating the objective function.

The aim of aspect 7 was to either validate or refute an assumption that I held: that the ideal output would be multiple timetables of similar quality, that the timetable maker could then choose from based off of individual qualities that they might find desirable. In hindsight, this question was not relevant to my project: it is a user requirement, and the aim of this project is a proof of concept, not an end-product to be used in actual hospitals. However, it might be useful for any future research into the ideas presented by this project, so the responses will be covered in this dissertation.

### 2.1.3 Interviews of Timetable Makers

The results of the survey of the timetable makers introduced me to a lot of aspects of the problem that I had not been previously aware of, without providing the information needed to address them. This was a result of my lack of knowledge in the area and the open ended nature of the questions: to the timetable makers it would be entirely reasonable to list part-time workers as a source of difficulty when rostering, but this left me with the issue of having to find out how part time workers operate within the NHS. As such, the interviews of the timetable makers aimed to expand on the information collected by the survey.

Questions relating to the following aspects of timetabling were asked:

1. Information regarding part-time workers.
2. Values for targets relating to the number of shifts worked in the planning period and average hours worked per week.
3. The number of doctors assigned to a standard rota.
4. The shift structure of their department.
5. Which aspects of shift coverage should be prioritised (e.g. giving extra penalisation for days entirely without coverage, in addition to the penalisation for uncovered shifts).
6. Steps that they take to improve doctor well-being.

Aspect 1 is aimed at addressing the aforementioned lack of knowledge regarding this topic. Additionally, when talking about part-time workers, the first interviewee mentioned targets for shifts worked and average hours worked, which again were aspects that I had been unaware of. I also realised that I had no real understanding of the shift structure of an average day, something that would be required in order to make realistic problem instances; aspects 3 and 4 sought to address this. Aspect 5 was covered because I was unable to get the desired information from the survey (Section 2.2.2), and I hoped that I would be able to better articulate the point of the question through conversation. Aspect 6 was included to attempt to understand what choices could be made in the problem definition to help improve doctor well-being.

## 2.2 Results of the Data Collection

### 2.2.1 Junior Doctor Survey

The survey of junior doctors was highly successful, garnering 38 responses. It provided valuable insight into the current status quo for timetable creation within the NHS, ways that the process could be improved, and the variety of preferences individuals can hold. The sanitised results of the survey can be seen in full in Appendix A.

#### Preferences

The survey of the junior doctors found a huge variation between individuals in terms of preferences. While this was expected, the survey also provided great insight into the types of preferences individuals might have. Broadly speaking, there were three main categories of preferences:

1. Preferences relating to the number of days worked in a row
2. Preferences relating to the number of nights worked in a row
3. Preferences for specific times of specific days not to be worked

For the purposes of modelling preferences for the project definition, the choice was made to enable a doctor to have a preference for each of the three categories; the system is then penalised for violating them. Requests that fall into category 3 are translated into requests to not work specific shifts, something that has been done before for formulations of the Nurse Rostering Problem [14].

**Burnout**

Interestingly, some of the answers for burnout correlated with the respondent in question's answers for preferences. This indicates that a potential remedy to the NHS' issues with burnout [24] would be to adhere to doctor preferences where possible, thereby justifying the inclusion of preferences in the problem definition. Disparity in shift staffing was also cited as a cause of burnout, highlighting the need for a system capable of better spreading out limited resources.

Also mentioned was the potential for assigned night shifts to prevent leave from being taken, and the subsequent difficulty in swapping night shifts, due to the extra restrictions surrounding their allocation. This point brought to light an incorrect assumption that I had held: that leave would be agreed before assigning shifts for the rota. The responses, however, indicated that the opposite was true: doctors are expected to fit their leave around the rotas.

**Leave**

In addition to an insight into how leave is currently handled, the surveys also provided a way to identify the general ways in which leave is taken. Through content analysis, the following "types" of leave were identified:

- Large Chunks (2-3 weeks)
- Medium Chunks (3-12 days)
- Small Chunks (2 days)
- Single Days
- Weekends

Additionally, after identifying these categories, it was possible to assign each response to one of them. Thereby a probability for a doctor taking each "type" of leave was obtained and used in the instance generator, in order to realistically generate leave requests for each doctor in the instance.

**Training**

Through the survey, the following "types" of training schedule were identified:

- One day per week
- A random day each week
- Two days per month
- One day per month
- A fixed afternoon every two weeks

- A fixed afternoon every week

Again, the survey led to a revelation that surprised me: that middle grade doctors were not guaranteed to be given rotas that allowed them to attend all of their scheduled training. Multiple doctors mentioned cases where they could not attend their training, or stated that it came second to the needs of the rota. Furthermore, the issue was also brought up in the final question of the survey, which asked whether there were any other aspects of timetabling that they felt were important. From this, training was identified as a major issue to be addressed by this project, especially given that neither question asked whether the doctors had difficulty in attending training: this is clearly an issue for them.

#### Impact on the Problem Definition

In response to the survey results, the following choices regarding the problem definition were made:

1. Leave and Training would be fixed as input data for the problem instances: the system would not be allowed to assign shifts that clashed with a given doctor's leave and training, and would have to work around them.
2. Preferences would be incorporated into the problem definition, at an individual level.
3. The objective function would harshly penalise the existence of days or shifts entirely without cover, in order to ensure a more even spread of coverage.

### 2.2.2 Timetable Maker Survey

The survey of timetable makers was also successful, although it uncovered more questions than the junior doctor survey had. 11 responses were received, which is very good, considering that there will be far fewer timetable makers than junior doctors. Full, sanitised responses can be seen in Appendix B.

The aspects that needed further investigation after the survey are as follows:

1. There was a large variety in the length of planning period given, so no conclusion about which would be best to use for the generated instances could be drawn.
2. In response to the question about preferences "Less Than Full Time Doctors" were referenced a lot, the consequences of this are discussed in Section 2.1.3.
3. It was still unclear how the locum shifts should be utilised.[1]

However, there were still some concrete conclusions drawn from the survey:

1. **It appears that most timetables are checked by HR staff**
   An automated system could therefore be easily incorporated into the current workflow: the output of any automated system would need to be checked, but this is already the case for manual timetables.
2. **Doctor preferences did not seem to be a priority**
   4 respondents even gave the impression that preferences were not even considered during the rostering process.
3. **Most respondents preferred the idea of an output of multiple timetables of a similar quality**
4. **Some timetable makers just use a pre-made rota from HR**
   This helps explain why so many of the respondents to the junior doctor survey complained about difficulties with leave or training. Additionally, the use of a template did not appear to result in a significant reduction in the time taken to produce a timetable: this is clearly not an ideal approach.

[1] The question mentioned the possibility of locums having preferences as an example of something that respondents might want to talk about. Many respondents then seemed to interpret the question as being solely about locum preferences, which was not the case.

### 2.2.3 Interviews of Timetable Makers

Interviews were conducted with two consultants from two different departments who were both involved with their department's timetabling. The questions for each interview varied: due to their conversational nature, responses to questions often entailed follow-up questions. This section will primarily focus on the overall findings, rather than the structure of each interview, but the full notes for both interviews can be found in Appendix C.

#### Doctor Targets

This topic was brought up in response to a question about the minimum hours a doctor could work; from the legal requirements it was clear that there was a maximum number of hours that could be worked but no further information was given. The response was that there were targets for the average number of hours worked per week, and that there were targets for numbers of specific types of shift worked: an objective for a successful timetable is to get as close to these targets as possible. There were two additional things to note: (1) that it was preferable to be under the targets than over them, and (2) that the targets regarding shift types were more "flexible" than those for hours, specifically that it was acceptable to go over the targets for shift types, but that going over the targets for hours worked was not an option.

#### Part Time Workers

Part time workers are contracted on the basis of a percentage, relative to the work performed by full time staff. This means that part time workers still have targets relating to hours and shifts worked, but the values are a percentage of the standard targets, as stipulated in their individual contract. i.e. a doctor on an 80% contract would work 80% of the standard targets. Beyond that, the interviewee made it very clear that the number of part time workers was increasing: the problem description would therefore have to be capable of accommodating part time workers.

#### Use of Locum Doctors

The interviewees made it clear that the use of locum doctors occurs with a large degree of regularity, and that the aim when assigning them is to make sure that the shifts were as spread out as much as possible. This is for two reasons: (1) (assuming full coverage is possible with the use of locums) it would be better to have a mix of doctors experienced with how the department is run alongside the temporary workers, and (2) there is no guarantee that a locum shift will be filled: were this to happen and there were no permanent staff scheduled, there would be no coverage for the shift at all.

#### Other Findings

- 8 week rotas are quite standard and could be used for the project's problem instances[a]
- There is an incredible degree of variation for how different departments organise themselves[b]
- One interviewee had taken the choice to provide 48 hours of rest after stretches of nights of any length, rather than the legally required 46; this improved doctor well-being without significantly impacting their ability to cover shifts
- Both interviewees gave me the general shift structure for their departments, which allowed for them to be modelled in the instances for this project

[a] Using an 8 week rota over a longer one made it easier to recruit volunteers for the comparative study

[b] For instance, shifts often require doctors of a certain experience level, or "grade". However the classification of grades is entirely dependent on the department; this potential for variation between departments was again something that the problem definition had to account for.

## 2.3 Problem Definition

### 2.3.1 Key Terms

- *Planning Period*: The time frame covered by the rota
- *Middle Grade Doctor*: A doctor who has completed their foundation training and is in the process of training for a specialist area. This entails working shifts in departments alongside attending training sessions.
- *Grade*: Middle Grade doctors are often split into "grades" according to their level of experience. The classification of these grades varies on a departmental basis.[a]
- *Doctor Targets*: doctors have targets for the average hours worked per week, the total number of day shifts worked during the planning period, and the total number of night shifts worked during the planning period[b]
- *Locum Doctor*: A doctor who temporarily fills a rota gap within a hospital, clinic, or practice.
- *Shift*: A period of work that needs to be suitably covered. The number of Middle Grade doctors required per shift can vary, as can the required grade of each.
- *Assignment*: For simplicity, the staffing requirements of a single shift are broken down into multiple "Assignments", one for each required doctor. Each assignment can either have a required grade, or can accept any assignee.[c]
- *Weekend Work*: Any work that falls between 00:01 on Saturday and 23:59 on Sunday.
- *Training Sessions*: Formal educational sessions undertaken by Middle Grade doctors as part of their progression; the timing of these sessions is predefined.
- *Annual Leave*: Doctors are entitled to a set amount of leave each year, the days taken as leave are defined before the rota is made.
- *Preferences*: Doctors are allowed to express preferences for (1) their preferred length for a stretch of days worked, (2) their preferred length for a stretch of nights worked, and (3) specific shifts that they would prefer not to work.

[a] For simplicity, this project only considers the grades "junior" and "senior", however, the problem domain is capable of handling any classification of or number of grades, so long as they are provided in the problem instance.

[b] Targets for part time workers are a proportion of the standard targets for the department

[c] For instance, a shift requiring two doctors, with at least one being senior, would have two associated assignments: one accepting only senior doctors, and the other being open to any grade of doctor.

### 2.3.2 Objective Function

**Min:**

$$P_D \sum_{d \in D} C_d + P_S \sum_{s \in S} C_s + P_A \sum_{a \in A} C_a + \sum_{m \in M} \text{deviation}(m) + \text{fairness} + \text{violations} \quad (2.1)$$

Table 2.1: Notation used in Equation 2.1

| Symbol | Definition |
|---|---|
| fairness | Result of Equation 2.7 |
| violations | Result of Equation 2.8 |

**Constituent Objectives**

1. **Minimise the number of days without coverage**

$$\textbf{Min: } P_D \sum_{d \in D} C_d \tag{2.2}$$

Table 2.2: Notation used in Equation 2.2

| Symbol | Definition |
|---|---|
| $D$ | The set of Days for a given instance |
| $P_D$ | The penalty for a day lacking any coverage |
| $C_d$ | $\begin{cases} 1, & \text{if day } d \in D \text{ lacks any coverage} \\ 0, & \text{otherwise} \end{cases}$ |

2. **Minimise the number of shifts without coverage**

$$\textbf{Min: } P_S \sum_{s \in S} C_s \tag{2.3}$$

Table 2.3: Notation used in Equation 2.3

| Symbol | Definition |
|---|---|
| $S$ | The set of Shifts for a given instance |
| $P_S$ | The penalty for a shift lacking any coverage |
| $C_s$ | $\begin{cases} 1, & \text{if shift } s \in S \text{ lacks any coverage} \\ 0, & \text{otherwise} \end{cases}$ |

3. **Minimise the number of unassigned assignments**

$$\textbf{Min: } P_A \sum_{a \in A} C_a \tag{2.4}$$

Table 2.4: Notation used in Equation 2.4

| Symbol | Definition |
|---|---|
| $A$ | The set of Assignments for a given instance |
| $P_A$ | The penalty for an unassigned assignment |
| $C_a$ | $\begin{cases} 1, & \text{if assignment } a \in A \text{ lacks an assignee} \\ 0, & \text{otherwise} \end{cases}$ |

4. **Minimise the deviation from doctor targets**

$$\textbf{Min: } \sum_{m \in M} \text{deviation}(m) \tag{2.5}$$

**where:**

$$\begin{aligned} deviation(m) = {} & P_{HU} \cdot (TH_m - HW_m) \cdot U_H \;+\; P_{HO} \cdot (HW_m - TH_m) \cdot O_H \\ & + \; P_{DU} \cdot (TD_m - DW_m) \cdot U_D \;+\; P_{DO} \cdot (DW_m - TD_m) \cdot O_D \\ & + \; P_{NU} \cdot (TN_m - NW_m) \cdot U_N \;+\; P_{NO} \cdot (NW_m - TN_m) \cdot O_N \end{aligned} \tag{2.6}$$

Table 2.5: Notation used in Equations 2.5 and 2.6

| Symbol | Definition |
|---|---|
| $M$ | The set of Middle Grade Junior Doctors for a given instance |
| $TH_m$ | The target for the average number of hours to be worked per week by doctor $m \in M$ |
| $HW_m$ | The average hours worked per week by doctor $m \in M$ |
| $P_{HU}$ | The penalty for each hour below the target number |
| $U_H$ | $\begin{cases} 1, & \text{if } TH_m - HW_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |
| $P_{HO}$ | The penalty for each hour over the target number |
| $U_O$ | $\begin{cases} 1, & \text{if } HW_m - TH_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |
| $TD_m$ | The target number of day shifts to be worked by doctor $m \in M$ |
| $DW_m$ | The number of day shifts worked by doctor $m \in M$ |
| $P_{DU}$ | The penalty for each day shift under the target specified |
| $U_D$ | $\begin{cases} 1, & \text{if } TD_m - DW_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |
| $P_{DO}$ | The penalty for each day shift over the target specified |
| $O_D$ | $\begin{cases} 1, & \text{if } DW_m - TD_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |
| $TN_m$ | The target number of night shifts to be worked by doctor $m \in M$ |
| $NW_m$ | The number of night shifts worked by doctor $m \in M$ |
| $P_{NU}$ | The penalty for each night shift under the target specified |
| $U_N$ | $\begin{cases} 1, & \text{if } TN_m - NW_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |
| $P_{NO}$ | The penalty for each night shift over the target specified |
| $O_N$ | $\begin{cases} 1, & \text{if } NW_m - TN_m > 0 \\ 0, & \text{otherwise} \end{cases}$ |

5. **Minimise the disparity in preferences being met**

$$\textbf{Min: } P_U \cdot (3 - (f(VD) + f(VN) + f(VS))) \tag{2.7}$$

Table 2.6: Notation used in Equation 2.7

| Symbol | Definition |
|---|---|
| $P_U$ | The penalty for an unfair distribution of preference granting |
| $f()$ | A function returning the value for Jain's Fairness Index $\in [1, 0]$ (Appendix D) |
| $VD$ | A set containing the number of times each doctor's preference regarding the number of days worked in a row has been violated[a] |
| $VN$ | A set containing the number of times each doctor's preference regarding the number nights worked in a row has been violated[a] |
| $VS$ | A set containing the number of times each doctor's preferences for shifts to avoid working have been violated[a] |

[a]Doctors only have a corresponding element if they have a preference in this area: the value for a doctor without a preference will always be zero

6. **Minimise the violations of individual doctor preferences**

$$\textbf{Min:}\ P_V \cdot (\sum_{d' \in VD} d' + \sum_{n' \in VN} n' + \sum_{s' \in VS} s') \tag{2.8}$$

Table 2.7: New Notation used in Equation 2.8

| Symbol | Definition |
|---|---|
| $P_v$ | The penalty for each violation of a doctor preference |

### 2.3.3 Hard Constraints

Some of these constraints are taken directly from the 2018 Junior Doctor Contract Refresh and European Working Time Directive: any timetable that does not meet these constraints cannot legally be used.

1. Doctors must be given at least 11 hours of rest between shifts.
2. Doctors must be given at least 24 hours of unbroken rest each week, or a stretch of 48 hours within 2 weeks.
3. Doctors can work no more than 4 night shifts in a row.
4. Doctors must be given 48 hours of rest after a night shift, or a stretch of night shifts.
5. Doctors can work no more than four long shifts (more than 10 hours long) in a row, and must be given 48 hours of rest after the fourth shift.
6. Doctors can work no more than 7 days in a row, and must be given 48 hours of rest after the last shift in that row.
7. Doctors must be given 2 weekends off for each weekend worked (1:2 ratio).
8. Assignments must be given to doctors of a suitable grade.
9. Doctors cannot be assigned shifts that clash with training sessions.
10. Doctors cannot be assigned shifts that clash with their annual leave.

# Chapter 3

# Investigating the Efficacy of a Hyper-Heuristic based Approach

Having defined the problem, the task still remained to provide a proof-of-concept for the applicability of an existing hyper-heuristic to find solutions for it. This chapter outlines the design and implementation of a Hyflex-compatible problem domain, as well as the subsequent evaluation of the performance of the GIHH hyper-heuristic [25] on it.

## 3.1 Design

The use of Hyflex necessitated an object-oriented design that implemented the methods specified by the Problem Domain interface. The impact of this is self-explanatory and can be seen in the codebase for the project, which is hosted on GitHub.

### 3.1.1 Feasibility

Given the importance of adhering to the hard constraints of the problem (a timetable that doesn't satisfy them would be illegal, and therefore useless), the decision was taken to limit the search space to feasible solutions, only allowing assignments that are legal. In this way the output of the system is guaranteed to be usable, and the algorithm is prevented from exploring regions of the search space that are without promise. Furthermore, this approach simplified the encoding of the objective function, as any given solution could be guaranteed to be feasible.

Feasibility is calculated dynamically by the system through delta evaluation, reducing the overall computation required for it. Inspiration was taken from the graph colouring problem, wherein a shift holds references to other shifts that will need to be reassessed once it is allocated to, or deallocated from, a doctor. Further checks are done for rows of days worked and for rows of night shifts and long shifts. These checks also use delta evaluation: the revaluation of feasibility at each stage is only done for shifts relevant to the row in question. For the implementation of this, see the GitHub repository.

### 3.1.2 Initial Solution Generation

The initial solution generation takes a heuristic approach to assigning shifts, by allocating shifts in order of their perceived difficulty. As leave and training are taken to be hard constraints by the problem definition, some shifts will have fewer feasible doctors from the start: these are assigned first. Next night shifts, and then day shifts are assigned until all of the doctor targets for shifts worked have been met. Pseudocode is provided in Algorithm 1. The inspiration for this approach came from work done by Burke et al. [20], and the ordering of difficulty for each type of shift was informed by the results of the timetable maker survey. It also prioritises having a single assignee for each difficult shift over fully covering each, as an even spread of coverage overall had been identified as being more desirable than full coverage for individual shifts.

### 3.1.3 Heuristics

The 11 heuristics provided by the problem domain are either taken from the Hyflex module for personnel scheduling [26], or are slight variations of those taken from there. While it had been hoped that heuristics could be gleaned from the survey of timetable makers, their responses were either not appropriate for low-level heuristics, or were more suited for use in the initial solution generation.

**Mutation Heuristics**

- *Heuristic 0*: randomly deallocates an IOM-dependent[a] number of assignments that were previously allocated to a randomly selected doctor
- *Heuristic 1*: randomly allocates an IOM-dependent number of previously unallocated assignments to a randomly selected doctor
- *Heuristic 2*: randomly deallocates an IOM-dependent number of allocated assignments (independent of the doctor that they are assigned to)
- *Heuristic 3*: randomly allocates a feasible doctor to an IOM-dependent number of previously unallocated assignments

**Crossover Heuristics**

- *Heuristic 4*: creates a solution from common assignments of both input solutions

**Ruin and Recreate**

- *Heuristic 5*: randomly deallocates a number of assignments previously allocated to a randomly chosen doctor and allocates that doctor to the same number of randomly selected unallocated assignments
- *Heuristic 6*: randomly deallocates a number of allocated assignments (independent of the doctor assigned to them) and reallocates them to a randomly chosen, feasible doctor

**Hill-Climbing**

- *Heuristic 7*: performs first-improvement hill climbing, but by assigning a random feasible doctor to each assignment
- *Heuristic 8*: performs steepest-descent hill climbing, but by assigning a random feasible doctor to all available unallocated assignments
- *Heuristic 9*: performs first-improvement hill climbing, exhaustively trying each available feasible doctor for each unallocated assignment until it finds an improving solution
- *Heuristic 10*: performs steepest-descent hill climbing, exhaustively trying each feasible doctor for each unallocated assignment

[a] Intensity of Mutation

## 3.2 Implementation

While Hyflex was written in Java, the choice was made to implement the problem domain using Kotlin. The reasoning behind this was simple: Kotlin compiles to the JVM, providing comparable performance to, and interoperability with, Java, while employing a more concise and clear syntax, something that simplified the development and maintenance of the codebase. The codebase also includes a number of unit tests which test that relevant functions operate as they should; these, and the source code, can be seen in the aforementioned GitHub repository.

Problem instances are implemented as text files, with a very simple structure: the grades for the department are specified at the top of the file, along with other general information such as the planning period length and doctor targets. This is followed by the information regarding each doctor: their leave, training shifts, preferences, etc. The rest of the file is then dedicated to the information associated with each shift of the planning period.

### 3.2.1 Instance Generator

As previously mentioned, it was necessary to create a program capable of generating realistic instances in order to be able to test the hyper-heuristic based approach. The generator is able to create rotas for the two departments identified through the timetable maker interviews (Appendix C); the number of doctors, the proportion of them that are on leave or part time, and the number of training groups can all be specified when generating an instance.

The information gathered from the junior doctor survey (Appendix A) was used to inform the generation of doctors for each instance. Each doctor can be on leave during the planning period, and can have a preference in three different areas. As discussed in Section 2.2.1, "types" of leave and the probability of a doctor taking each type were identified through the survey. The generator uses this information to realistically assign leave to doctors within the problem instances. Once the "type" of leave is selected, the generator chooses a random day within the planning period for the leave to start from. The preferences expressed by each respondent to the survey have been encoded as a triple containing their preferences for each area; the generator then selects a random triple from the survey results for the generated doctor to have.

Also taken from the doctor survey are the different training schedules that certain departments use. By default, there is a probability for each training schedule being selected that was identified from the survey. Once the type of training schedule for the generated department has been selected, the doctors are assigned to training groups. These groups are then assigned specific times for their training to take place in; the shifts that would violate these training times are then associated with each doctor in the generated instance file for use in the Problem Domain.

#### Departments Modelled

By chance, the two timetable makers who volunteered to be interviewed worked in very different departments. This was of great benefit to the project as it meant that it would be possible to investigate how well the hyper-heuristic could adapt to the differences in the structure and organisation of each department.

Department 2 follows a "standard medical" rota, with an expectation that there will be less work to do at night. Department 1, by contrast, follows a "full shift" rota, where the workload is expected to be consistent throughout a 24 hour period. The worker for Department 2 provided the standard number of doctors available to them, which they said tended to be sufficient to cover the rota. Department 1, on the other hand, suffered from perennial understaffing, so the baseline instances used in the experimentation start from this already low water mark in terms of doctor numbers.

## 3.3 Evaluation

Having implemented the problem domain, it was then necessary to investigate the ways in which an existing hyper-heuristic could be used in tandem with it. The first way was the most obvious: creating timetables, the aim of the project from the very beginning. To this end, three experiments were conducted: (1) an investigation into the time limit that should be used for the other experiments, (2) a study comparing timetables made by human participants to those produced by the automated system, and (3) experimentation with various problem instance characteristics to see their effect on the system's ability to find high quality solutions.

It was while considering the details of (3) that a new idea was found: namely that, if the problem definition is accurate, the hyper-heuristic system could be used as a tool for modelling a department's ability to cover shifts. This would be a very powerful tool for the timetable makers. For instance, a worker wishing to find the optimal training schedule for their department could experiment with different schedules, observing their impact on the objective function value and the usability of the solutions found. The same would apply for deciding whether or not to grant a doctor leave, whether the department could afford for doctors to start working part time, or how many doctors are needed to adequately cover shifts. This section of the evaluation will also use the results of the experimentation conducted in (3), in order to assess the potential benefits or limitations of such an approach to modelling.

### 3.3.1 Experimental Setup

All experimentation was done using a Windows PC with an AMD Ryzen 5 1600 CPU at 3.2GHz and 16 GB of DDR4 RAM at 3000MHz.

#### Objective Function Parameters

The problem definition deliberately leaves the weighting of penalisations for the objective functions as parameters to be defined by those solving the problem; the idea being that the parameters could be set according to each department's needs, given the large degree of variation in their individual modus operandi. In order to evaluate the produced system, however, concrete values for each parameter had to be chosen. Table 3.1 shows these values.

The values were chosen to try and prioritise shift coverage above all else. The targets for day shifts were given the lowest possible priority as I was told by one of the interviewed workers in a follow-up email that these were not particularly important, and that night shift targets and average hour targets should have a much greater impact on the objective function. The values relating to preferences were chosen after some experimentation with the aim of getting the hyper-heuristic to pay some attention to the preferences, without overpowering the objectives relating to shift coverage, night shifts worked, and the average hours worked.

Table 3.1: Parameter Values used in the Objective Function

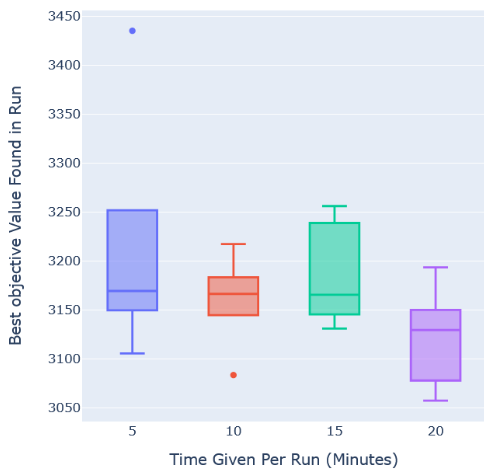
| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $P_D$ | 30 | $P_S$ | 20 |
| $P_A$ | 15 | $P_H U$ | 5 |
| $P_H O$ | 10 | $P_D U$ | 2 |
| $P_D O$ | 2 | $P_N U$ | 10 |
| $P_N O$ | 12 | $P_U$ | 500 |
| $P_V$ | 4 | | |

### 3.3.2 Time Limit Experimentation

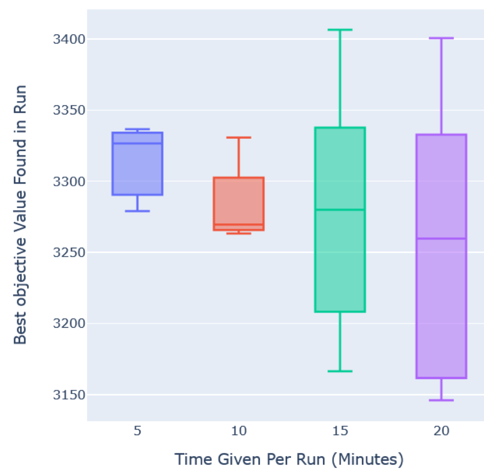
#### Methodology

A natural starting point for an evaluation of the hyper-heuristic's performance was an exploration of the time needed to find a high-quality solution: a major factor in wanting to automate the process of timetable creation was the time spent by NHS when making them manually. Furthermore, experimenting with different time limits gave a justification for one to be chosen for the others for use in the rest of the project's experimentation. There were two criteria that needed to be met for a potential time limit: (1) it had to be short enough that the experimentation could finish with reasonable time for an interpretation of the results to be made and (2) it should not cause an excessive decrease in the quality of solutions found.

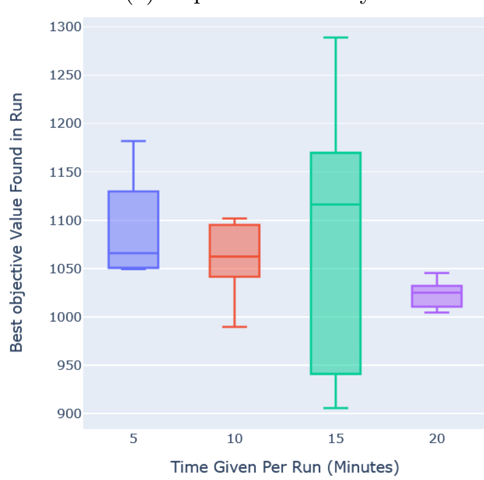
To this end, the hyper-heuristic was run on all 4 instances created for the comparative study (see Section 3.3.3 for details of their characteristics), 5 times for each time limit. The tested time limits were 5, 10, 15, and 20 minutes.



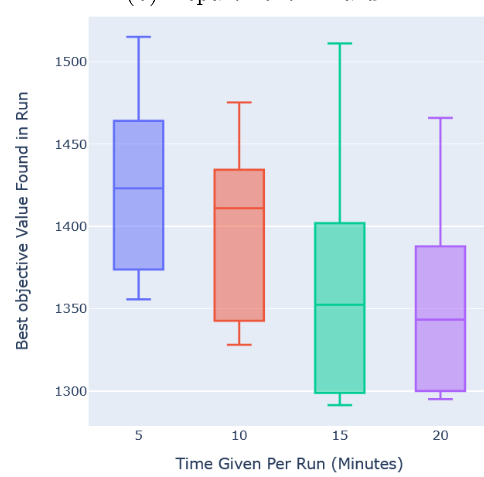

(a) Department 1 Easy




(b) Department 1 Hard




(c) Department 2 Easy




(d) Department 2 Hard

Figure 3.1: Results of the Time Limit Experimentation

#### Analysis of results

The results can be seen in Figure 3.1; for the most part, there was not a significant improvement beyond 15 minutes of execution time. In the interest of balancing the need to find good solutions with being able to complete the experimentation in time for the project deadline, 15 minutes was therefore chosen as the time limit for all future experimentation. Of interest is the fact that the range in quality of the solutions found tended to be quite large regardless of the time limit. I believe that this indicates a proclivity for the search to get stuck in local optima. I can see three potential causes of this:

1. The choice to limit the search space limits the exploratory potential of the search
2. There might not be enough variety in the heuristics provided by the problem domain - there is only one crossover heuristic, for instance
3. It is a limitation of the hyper-heuristic used; the Hyflex competition took place in 2011, and the techniques for heuristic selection and move acceptance have definitely improved since then: it is entirely plausible that a more modern hyper-heuristic would perform better on the same problem instances

As the fastest time identified in the survey was 3-4 hours, the hyper-heuristic system is clearly significantly faster than manual approaches. What remained to be seen was whether it could outperform manual methods, which was the aim of the next experiment.

### 3.3.3 Comparative Study with Manual Methods

#### Methodology

This study aimed to prove that the hyper-heuristic system was capable of creating better timetables than could be created by hand. It had been hoped that it would be possible to get participants with experience in healthcare timetabling, but, apart from one volunteer (See Conclusions), this was not possible. To make up for this a GUI was provided that allowed them to interact with the encoded problem domain. This meant that the feasible doctors for each shift were automatically calculated for the participants, and that it would be impossible for them to create infeasible timetables.

The interface also provided a view of the data associated with each doctor: their grade, preferences, targets, etc. This meant that the problem presented to the participants was simpler than the one dealt with by NHS staff as these details, and the feasibility of shifts, would ordinarily have to be kept track of manually; the idea being that this could mitigate the impact of their lack of timetabling experience. Participants were also given a guide that explained how to use the interface, their objectives when creating a timetable, and the rules surrounding their creation (Appendix F).

Four instances were created for this study, two for each department, with one instance representing a normal scenario for the department, and the other presenting more of a worst case. This normal instance had a standard number of doctors for its department, with only 20% of doctors taking leave during the planning period and only 20% being 80% part-time workers[1]. The worst case instance had two fewer doctors than the normal one, as well as twice the number of part-time workers and doctors on leave. The participants were split into two groups, with each group receiving the normal instance for one department and the more challenging one for the other.

The rationale behind this approach was that it was going to be difficult to find volunteers willing to take part in the experiment if too much was asked of them (making just two timetables was already a relatively large task). The hope was that measuring and comparing performances for the best and worst cases of each department, should provide some insight into how the hyper-heuristic might perform, relative to

[1] a survey respondent said that this was the most common type of part-time worker

manual methods, across a spread of instances of varying difficulty. At the very least it would highlight the particular strengths of the automated approach when compared to making the timetables manually.

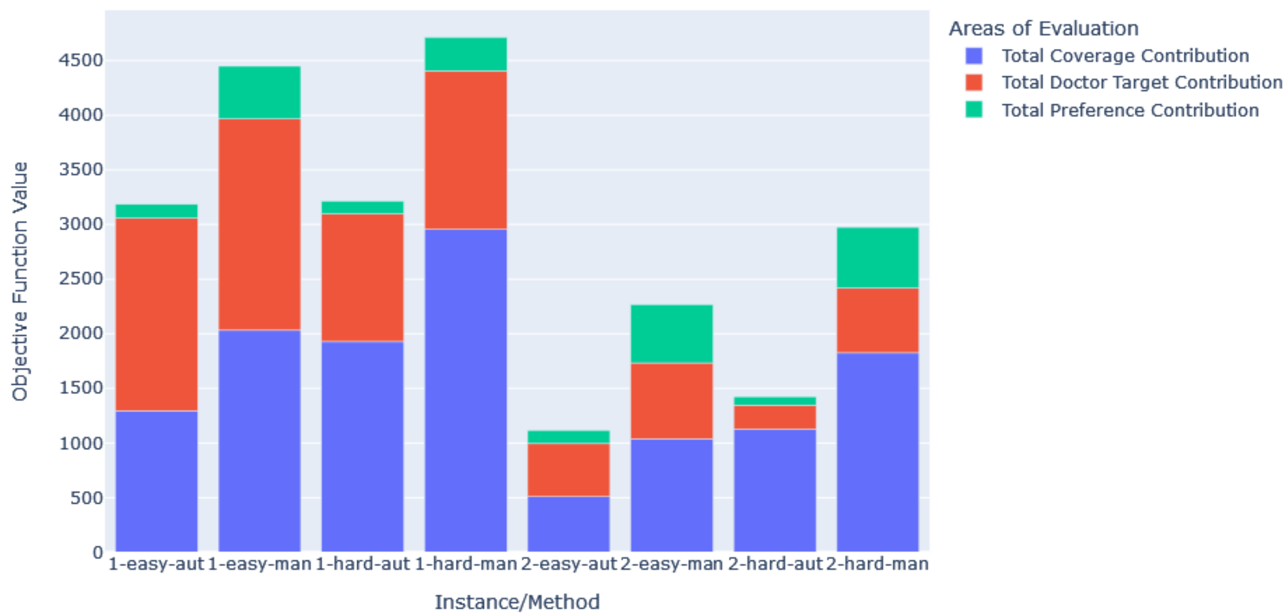



Figure 3.2: Results of the Comparative Study

### Analysis of Results

Figure 3.2 shows the mean objective function value for each instance for the automated and manual approaches. It also depicts the contribution of the three main aspects of evaluation for a created timetable: penalisation for lacking shift coverage, penalisation for not meeting doctor targets, and penalisation for not meeting doctor preferences.

The aim of the study was to prove that the hyper-heuristic could outperform humans at creating timetables for the same instances, something clearly visible in the results of this study: the hyper-heuristic outperforms the participants on each instance, soundly achieving better shift coverage and adherence to the doctor preferences. Notably, the hyper-heuristic is consistently better across both the "easy" and "hard" instances, indicating that it should outperform manual methods in most scenarios.

As noted in Section 3.2.1, the two departments are very different in terms of resources and structure; the fact that the hyper-heuristic outperforms the human participants for both shows that it is capable of adapting to inter-departmental differences in organisation. It should be noted that the timetables made for Department 1 have worse scores than those produced for Department 2, but this is not an indicator of the system being unable to cope with Department 1's structure: the instances for Department 1 are inherently harder - even the "easy" instance is understaffed.

Despite the success with shift coverage and adherence to preferences, the contribution of doctor targets to the objective function value was strange. For instance, in Department 1, despite easily outperforming the participants in the other areas, the hyper-heuristic achieved a similar contribution for doctor targets. This warranted further investigation.

**Issues Surrounding Doctor Targets**

The issues with doctor targets also resulted in strange values for the overall objective function value. Looking at the automated results for Department 1, we can see that the overall objective function is essentially the same for both instances, despite the hard instance having significantly worse coverage. This presented a major issue, as shift coverage was meant to be the objective with the most priority - having a timetable with inferior coverage being evaluated as equivalent to one with better coverage was unacceptable.

The cause for this was that the doctor targets were having a disproportionate contribution to the objective function value. Some aspect of the hard instance then made it possible for the hyper-heuristic to get closer to the doctor targets, reducing their contribution to the overall score and effectively acting as a buffer for the increase in the shift coverage contribution. Upon closer inspection, it became clear that the values for the doctor targets were contradictory for Department 1: it would be impossible to adhere to the target for hours worked while adhering to the targets for the shift numbers.

This made sense, as the values for the shift targets came from the NHS worker who also provided the shift structure of Department 2 and had described these values as "classic targets" for a department. I took this to mean that they were the classic targets for a department in general, but I have come to realise that they could have meant for a department of their particular specialty. Alternatively, it could be that these values are standard, but just cannot be applied to Department 1 due to its non-standard shift structure (Section 3.2.1).

Whatever the case, it was clear that I would need to get the correct information for Department 1, so I got in touch with the worker that provided its structure. In addition to the issues surrounding the shift targets, a significant contribution to the score was coming from the doctors consistently being below their targets for the average hours worked per week, so I also asked about this.

The worker agreed that the targets were not applicable to their department, and gave me some amended values that they felt were more appropriate. In response to the question about hours worked, they asked whether my calculation for the hours took training and leave into account. The problem domain did adjust for leave, but I had never been explicitly told that training was taken into account when calculating the targets - I had always taken these values to refer only to shift work.

Unfortunately, this was found out too close to the project's deadline for me to be able to redo all of the experimentation with amended instances, so I had to prioritise redoing a certain aspect of the experimentation. For what was done, and how, see Section 3.3.4. In spite of these issues, the comparative study met its original primary aim, so conclusions could still be drawn from its results.

**Conclusions**

Despite the issues surrounding doctor targets, the hyper-heuristic easily outperformed manual methods. For the targets that were modelled correctly it was able to effectively minimise their impact on the objective function score: shift coverage and preference adherence were consistently better achieved and, for department 2, the shift targets were also adhered to, indicating that the hyper-heuristic is capable of balancing the multiple objectives of the problem, provided that it is fed accurate information about the department.

Furthermore, the adherence to preferences and the fairness in adherence did not appear to diminish the hyper-heuristic's ability to find higher quality timetables: it achieved a much higher degree of shift coverage than the participants while incurring only a fraction of the penalty for preference violations. This aligns with the findings relating to preferences fairness explored in Section 1.4, indicating that this is indeed a major advantage of an automated approach and should certainly explored further.

There are limitations to these results, however: the inaccuracy in the instances does mean that the hyper-heuristic would have to be tested on real-world instances in order to properly assess its applicability to the problem. But this study does prove that it can meet the objectives for a similarly complex approximation of the real-world problem, which does indicate that it would be worth investigating the system's performance on real-world data.

Another big limitation comes from the participants of the study: they had no experience in making timetables prior to the study. The steps taken to limit this were outlined the Methodology. It was also possible to find a single NHS worker willing to take part in the studies. They created a timetable for each of the four instances, Table 3.2 shows their results. These results are either slightly better or similar to those of the participants which does indicate some validity in the conclusions of this study. The hyper-heuristic greatly outperforms the timetable maker in terms of shift coverage, which is the primary goal for the automated system. However, these results are not statistically significant, so more willing participants with timetabling experience would be needed in order to corroborate this.

Table 3.2: Results of an NHS Timetable Maker for the experimental Instances

| Instance | Objective Function | Coverage | Doctor Targets | Preferences |
|---|---|---|---|---|
| Department 1 Easy | 4595.98 | 2230.00 | 1883.95 | 478.03 |
| Department 1 Hard | 4655.92 | 3445.00 | 938.51 | 270.41 |
| Department 2 Easy | 1683.50 | 670.00 | 581.64 | 451.86 |
| Department 2 Hard | 2418.95 | 1760.00 | 340.48 | 294.47 |

The final limitation stems from the manner of evaluation, namely that it is dependent on parameters in the objective function that have been set with only an approximate idea of what type of solution would be desirable. This means that specific aspects could be weighted too greatly or too minimally, resulting in a numerical evaluation of the timetables that is not reflective of the real world. The remedy to this problem would be to run the hyper-heuristic on real-world instances and to allow a member of staff to assess the quality of the produced timetables, but this is unfortunately beyond the scope of what is possible with this project.

### 3.3.4 Instance Experimentation

#### Methodology

It is an established practice to explore how automated systems for a timetabling problem perform across a variety of instances [7, 14, 17, 27]. Doing this allows for the evaluation of a system's ability to cope with the different characteristics problem instances might have. For this project, the objectives of this aspect of the evaluation were twofold: (1) to explore how well the system could cope with a variety of instances and to isolate which specific characteristics it struggled with, and (2) to investigate whether or not the system could be used to model a department's shift coverage.

From preliminary testing, there were four main characteristics that could impact the system's ability to cover shifts:

1. The proportion of doctors on leave during the planning period
2. The proportion of doctors working part time
3. The degree of understaffing
4. The training schedule used

In order to explore the effect of varying these characteristics, another set of instances were generated. For every instance other than those exploring the training schedules, a fixed training schedule for the "2 days per month" approach was used.

First, a baseline instance was created for each department, which had zero doctors working part-time or on leave during the planning period. These baseline instances were then altered to create the others, with only the characteristic in question being altered. For example, in order to explore the impact of leave, four instances were created with 20%, 40%, 60%, and 80% of the available doctors taking leave respectively. In order to isolate the change in objective function value to the proportion of doctors on leave, once a doctor had been assigned shifts to take as leave, these shifts became fixed across the instances. i.e. If a doctor was on leave in the 20% instance, they would be on leave in all the other instances with the same shifts taken off.


Figure 3.3: The Average Results for Each Baseline Instance

Four instances, one for each of the percentages used for the leave-exploring instances, were created in a similar manner to explore the impact of the proportion of part-time workers on the objective function value; for simplicity, all part-time workers were taken to be 80% part time. Also created, were instances where doctors could be part-time, on leave, or both, with the same set of percentages as before.

Understaffing was also explored, with instances for each department in the absence of the following staff (when compared to the baseline instance):

- 1 Junior Middle-Grade
- 1 Senior Middle-Grade
- 1 Junior and 1 Senior Middle-Grade
- 2 Junior Middle-Grades
- 2 Senior Middle-Grades

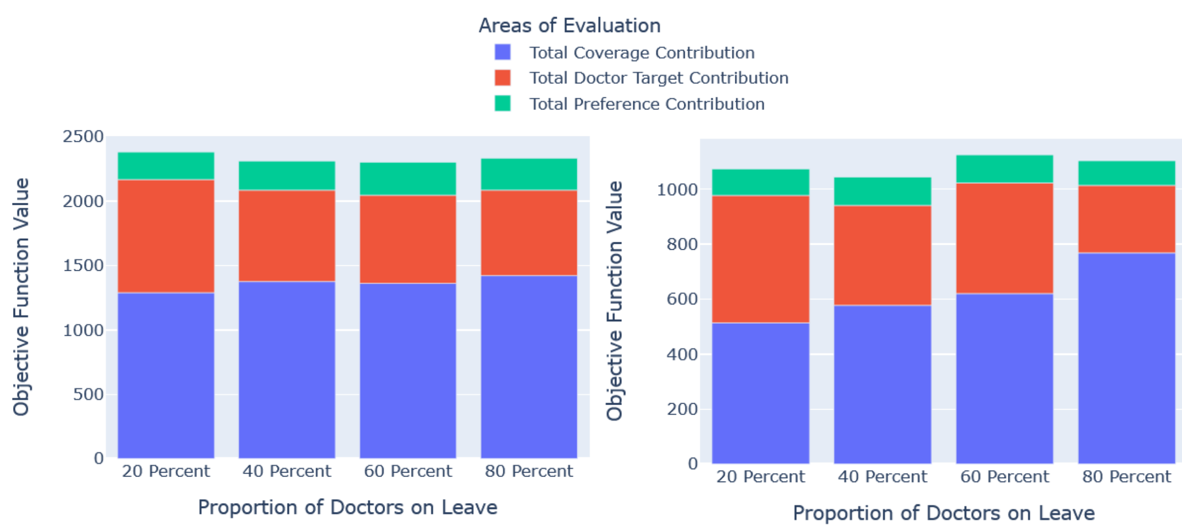
For the investigation of training schedules, a baseline instance was created for each department with 20% of the doctors taking leave during the planning period and 20% of the workforce being 80% part-time. Examples for each of the schedules identified in Section 2.2.1 were generated and applied to the doctors of the baseline instance.

By running the hyper-heuristic on each of these instances, and comparing its results for each, it would be possible to isolate their impact on the objective function. This would provide two outcomes. First, the system's robustness to changes in these characteristics could be observed, thus exposing its particular strengths and weaknesses. Secondly, it would provide an avenue by which to assess the system's ability to model a department, as these experiments mirror the tasks that could be carried out when doing so. To this end, the hyper-heuristic was run 5 times for 15 minutes on each of the instances. The mean objective function value for the best solution found during these runs was recorded, as well as a breakdown of the individual contribution of each main area of evaluation (shift coverage, adherence to doctor targets, and adherence to preferences).

**Rerunning the Experiment for Department 1**

As noted in Section 3.3.3, there were some issues with the "doctor targets" aspect of the objective function: the instances for Department 1 had incorrect targets, and the measure for average hours worked per week did not take training into account. This was found when analysing the results of the comparative study, which had been run in parallel to this one, so all the experimental results of this study were impacted by these errors. While it was relatively simple to expand the mechanism that accounted for leave to also account for training, I unfortunately didn't have time to rerun all of the experiments, but I felt that it made sense to rerun this study with corrections for Department 1.

The rationale behind redoing part of this study in particular was simple: one of the aims of it was to investigate the applicability of the hyper-heuristic as a tool for modelling, which could not be done at all if the model being evaluated had glaring issues. Having results from versions of the system with and without the inaccuracies also allowed for a comparison between them, and an analysis of how this oversight affected the solutions found. Department 1 was chosen over Department 2 as it would give an opportunity to see the solutions produced for it when given the correct shift targets.




(a) Department 1

(b) Department 2

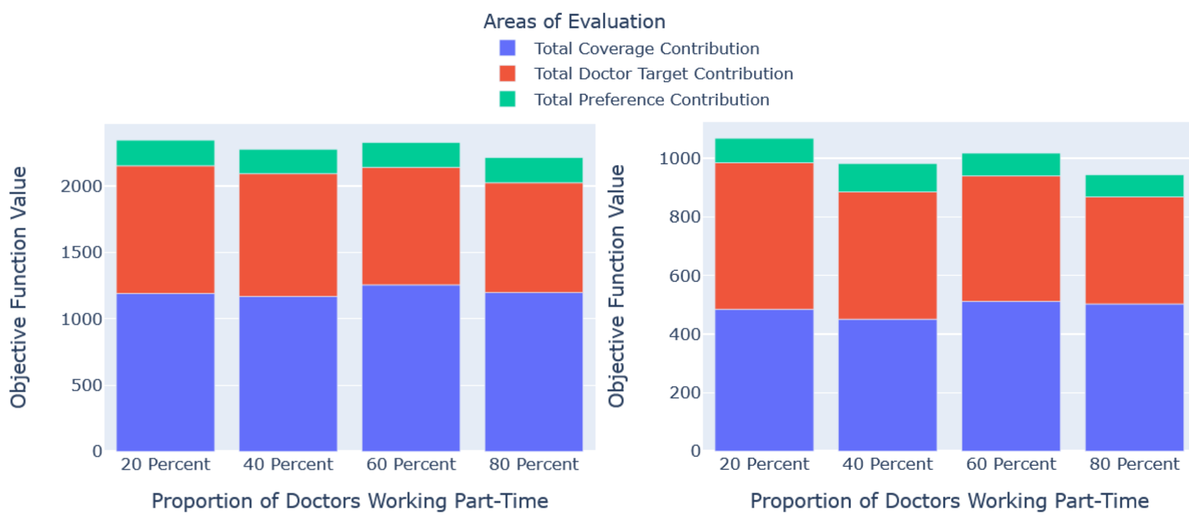
Figure 3.4: Results for Leave Experimentation

**Analysis of Results**

Shown in Figure 3.3 are the average results for each of the baseline instances, as well as a breakdown of where the score has come from. The objective score is much higher for Department 1, but this makes sense: it is understaffed and therefore has worse coverage, and has more doctors, resulting in greater opportunities for deviation from their preferences and targets. Also of interest is the fact that the score for Department 2 was the same across all 5 runs, with the same breakdown in areas of evaluation, indicating that some kind of optimum has been found (it could just be a local one).

Figure 3.4 shows the average results for doctors taking leave. For department 1, the system shows a great degree of robustness, with only slight changes to the coverage and preference contributions. The same cannot be said for department 2, where the coverage is inversely proportional to the percentage of doctors on leave; adherence to preferences remains relatively constant, however. For both departments the overall objective function remains similar, which is problematic given that the coverage has worsened.

For Department 2, this could be pinned on the fact that the system didn't take training into account when calculating the average hours, meaning that doctors would be classed as being below their target hours, despite working as much as they feasibly could given the rota rules. Leave, however, was taken into account at this stage of the project, so each doctor that went on leave would have their target for hours worked reduced, and would thereby be closer to their target hours, minimising the penalty incurred. The same cannot be said for Department 1, however, as training was taken into account; this indicates that either the weightings for the objective function need to be put even further in favour of shift coverage, or that the values for the doctor targets are still incorrect.

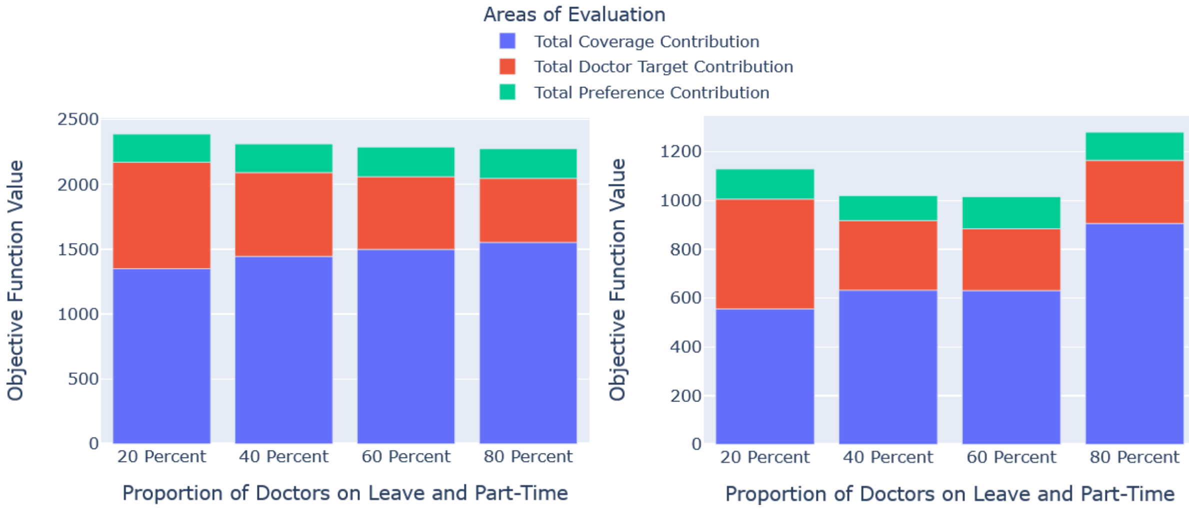



(a) Department 1

(b) Department 2

Figure 3.5: Results for Part-Time Experimentation

Figure 3.5 shows the average results for part-time workers. In this case doctors working part time seem to have minimal impact on either department. There are caveats to this, however. Firstly the doctors are only 80%, which will have less of an impact than part-time workers who work fewer hours. Furthermore, this modelling of part time workers only impacts their target hours; from the timetable maker survey, it seems that part-time workers can request not to work specific days, which is not implemented in my problem instances.

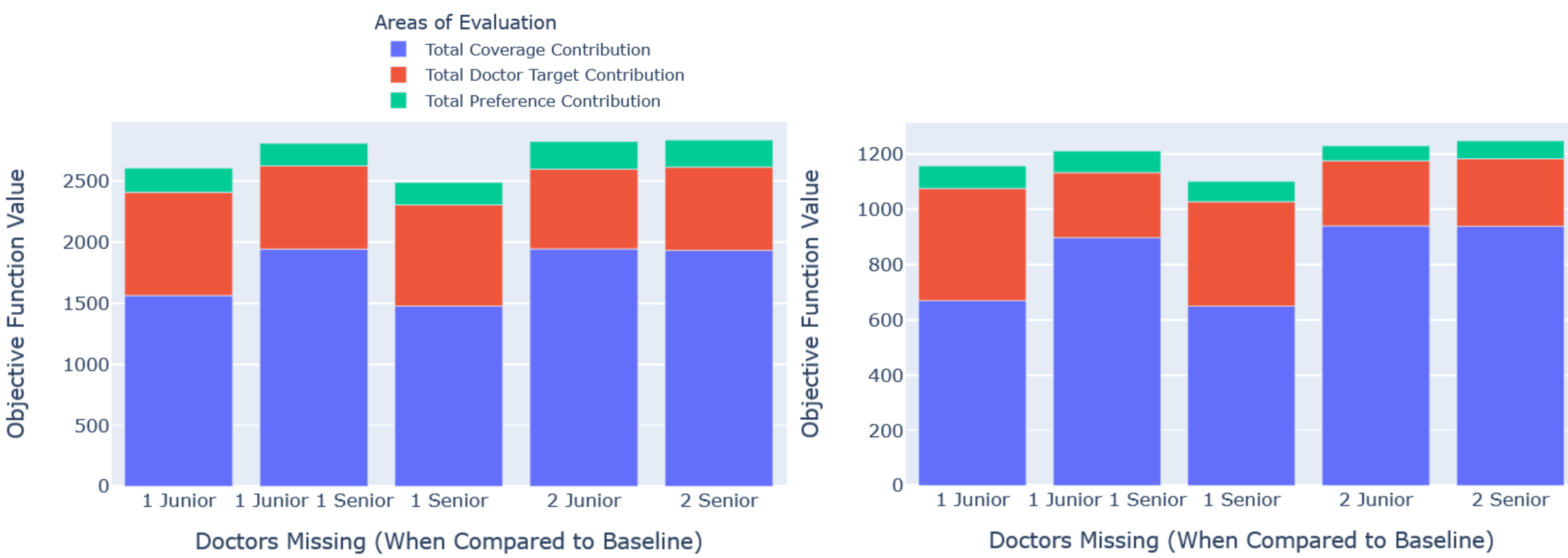



(a) Department 1

(b) Department 2

Figure 3.6: Results for Leave and Part-Time Experimentation

Figure 3.6 shows the average results for a mix of doctors on leave and working part-time. Again there is minimal impact on Department 1 while Department 2 sees a significant drop-off in coverage at 80%. The cause for such a sudden increase could be in the instance itself: while doctor leave was consistent across the instances once being taken by a doctor, it was still randomly generated once the doctor to take leave had been selected. The two new doctors that were selected to take leave in the 80% instance (they were not on leave in the 60% instance) both took a significant number of shifts off in the planning period, which will have made it a lot harder to cover shifts.

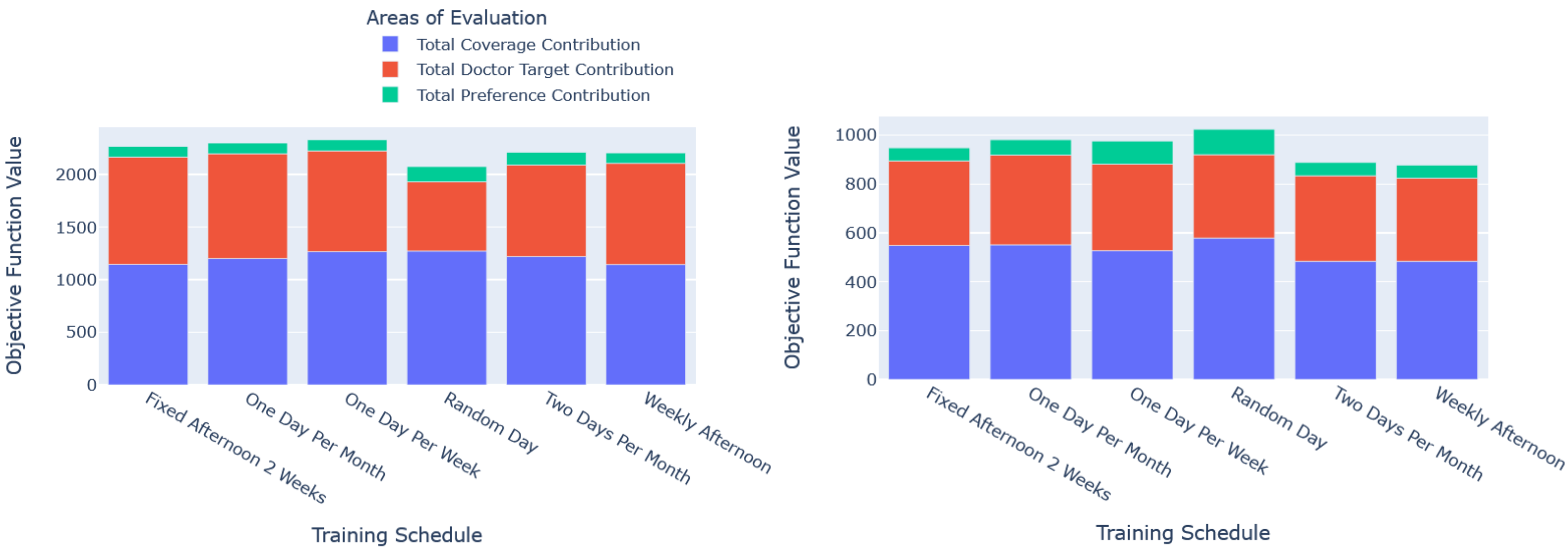



(a) Department 1

(b) Department 2

Figure 3.7: Results for Understaffing Experimentation

Figure 3.7 shows the average results for the various states of (further for Department 1) understaffing. The results are remarkably similar for both departments, with the main deciding factor in a department being able to provide coverage being the number of doctors rather than their grade. However, part of this will be due to the fact that the system only penalises a lack of coverage for a shift, and doesn't take the grade of the missing doctor into account. I would imagine that some departments would put a greater emphasis on shifts requiring a specific grade. This would be trivial to add to the objective function and could form the basis for further work exploring this area.




(a) Department 1

(b) Department 2

Figure 3.8: Results for Training Schedule Experimentation

Figure 3.8 shows the average results for each of the training schedules when applied to each department. While some training schedules results in worse coverage or less of an ability to adhere to preferences, the results are quite similar for each schedule. This makes sense as these are all real schedules identified from the survey, and it is unlikely that a significantly worse approach would be used by a department.

### Conclusions Regarding Robustness

Overall the results show a good resilience to the various characteristics that might hinder its ability to find good solutions. Due to the likely inaccuracy of the values for the doctor targets, it is more useful to look at the results relating to coverage and preferences. In both of these cases, the hyper-heuristic is very consistent, showing minimal change in their contribution for the different training schedules, and for varying degrees of part-time workers. In the case of leave, it is robust up until the very extreme scenario where 80% of doctors are on leave during the planning period, a scenario that would be unlikely to occur in real life.

The only characteristic that the hyper-heuristic struggles with is understaffing. This was already evident from the disparity in the objective function values of the best solutions found for each department's instances in the previous experimentation: solutions for instances of Department 2, which was fully staffed, were always evaluated as being better than those found for Department 1. The results for the experimentation show that the number of doctors per instance is the only characteristic that impacted the objective in a linearly proportional manner.

This was to be expected, as the impacts of understaffing can only be mitigated: no matter how well the doctors are spread out across shifts, there will not be enough staff to cover all of them. This means that the hyper-heuristic does not have any particular glaring weaknesses in the areas of the problem that can be reliably assessed. Because of this, these results, combined with the comparative study, make a strong case for the efficacy of hyper-heuristics for finding solutions to this problem.

#### Conclusions Regarding the Potential for Use in Departmental Modelling

I feel that the results of this study lend credence to the idea that hyper-heuristics could be used to model shift coverage in addition to making timetables. For instance with the jump between the 60% and 80% leave instances, a clear image is provided for a use case: doctors would be able to submit their leave requests in advance, and the timetable makers could use the system's ability, or inability, to cover shifts as a determining factor in whether or not to grant their requests. Furthermore, it would be possible to simulate potential training schedules, or a doctor's request to work less than full time; there are a lot of use-cases.

The limitation to the evaluation of this potential application is the fact that real data is not being used, and the fact that the objective function does not necessarily take into account all the factors that need to be considered when evaluating a timetable. Additionally, the issues with the doctor values mean that the overall objective function value - in the project's current state - could not be used, and a breakdown of the individual factors of contribution to the score would have to be evaluated. Further experimentation with real data and an objective function with weightings agreed upon by NHS professionals would therefore be required in order to truly evaluate the applicability of hyper-heuristics to this use-case. Nonetheless, this project does show that it should be possible, and this aspect of the work is definitely one that could be taken further.

# Chapter 4

# Summary and Reflections

## 4.1 Contributions

The project provides a formulation of a problem mostly untouched by research into healthcare scheduling. It also provides a problem domain on which any hyper-heuristic could be run, provided it adheres to the specifications of the Hyflex framework. Both the problem domain and definition are designed to be open enough to accommodate many types of department, in order to exploit the generality of the hyper-heuristic approach to optimisation. The project proves the efficacy of hyper-heuristics for creating timetables, despite the highly constrained nature of the problem. The project also provides further evidence for the fact that preferences can and should be included in problem definitions for healthcare scheduling, as doing so in no way significantly impedes the quality of solutions found. Furthermore, the project outlines the potential for hyper-heuristics to be used as a tool for modelling, which is a concept that I had not seen before.

In this way, the project provides proof-of-concept for the potential applications of hyper-heuristics to the problem of Middle Grade Doctor shift scheduling. This lays the foundation for a variety of research opportunities which will be discussed in Section 4.2.

## 4.2 Future Directions

The first and most obvious future direction is to take the system as is and provide it with some compatible real-world instances. In this way the veracity of the conclusions of this project could be truly assessed. Another continuation of the project would be to make different objective functions to represent the unique priorities each department might have - e.g. a greater penalisation if a certain grade of doctor is not covering shift. This would further test the capability of the hyper-heuristic approach to cope with inter-departmental differences and, if it could cope with these additional objectives, would be strengthen the argument for using hyper-heuristics over other methods.

Another idea would be to try and encode a more comprehensive collection of low-level heuristics, to see if that would improve the quality of solutions found. Additionally, it would be interesting to explore how other hyper-heuristics perform on the problem - at the very least, the problem provides an additional test for the generality of any developed hyper-heuristic algorithms.

The ability of a hyper-heuristic approach to cope with a more complex modelling of preferences could also be explored. The current implementation in the problem domain is relatively basic, only allowing for a preference in three basic categories. There is certainly room for this to be expanded to allow a more flexible system, capable of accounting for the individuality of each doctor. The question would then be whether a hyper-heuristic is capable of adhering so such a complex degree of constraint or whether another approach would need to be considered.

I also feel that the project provides a solid argument for the use of existing hyper-heuristics to reduce the overhead required to automate timetabling. The pre-existing hyper-heuristic was able to vastly outperform manual methods without the need for fine-tuning. It would therefore be useful to explore whether the same could be done for other areas of healthcare timetabling, or even timetabling in other contexts. Doing so could help highlight the applicability of hyper-heuristics to a plethora of real-world problems, which would help move them from being a topic of theoretical interest to having a tangible impact.

Finally, further exploration of the applicability of a hyper-heuristic system as a tool for modelling would be a very worthwhile avenue. The project's ability to assess this was hampered by the issues surrounding instance creation but it did show that changes in the objective function could be linked to specific instance characteristics, which would be useful for modelling. The steps needed to truly test this are outlined in Section 3.3.4; doing so would enable the creation of a truly transformative tool for timetable makers.

## 4.3 Project Management

I would class the management of this project as acceptable, in that I delivered what I had agreed to by the submission deadline, but behind the roadmap that I had set for myself. This was due to personal circumstances that hindered my work in the first term of the year, and the issues described in Section 3.3.3, which resulted in me having to redo a significant portion of the evaluation of the system. The fact that the project was delivered in time, despite these difficulties, does indicate that enough time was given at the start of the project to account for any such unforeseen problems.

Unfortunately, not being ahead of schedule meant that I was unable to do some of the further experimentation I had initially hoped to do, if I had the time. In particular, I had wanted to carry out two of the ideas currently listed in Section 4.2 if I had the time. Namely, the experimentation with different objective functions, and the evaluation of the performance of different hyper-heuristics on the problem. Of the two I think the bigger shame is not getting to do the objective function experimentation, as it is a very natural continuation of what has already been done.

The original Work Plan provided in the Interim Report can be seen in Appendix Section G.1. The core idea behind the plan was to split the project into two halves, one in which data gathering would be performed simultaneously to tasks that were not dependent on the data to be collected - e.g. the encoding of the hard constraints for the problem domain. Following the setbacks of the first term, a revised plan was made, which pushed all of the remaining tasks back by one week. This can be seen in Appendix Section G.2.

In this plan, and the original, there was a linear progression in tasks being completed one at a time in succession. This was not what eventually happened. Due to the difficulty in getting all of the information necessary to provide a suitable problem definition, the process evolved into a slightly iterative one. Wherein I would find areas of uncertainty, ask one or both of the workers who had agreed to help me for clarification, and then update my problem domain and definition accordingly. It would probably have been ideal to operate in this way from the start, but at that point I didn't know that I would have access to people who were so willing to help.

Additionally, some work ended up being completed in parallel; due to the extra time needed to complete the problem definition and consequently the problem domain encoding, the updated low-level heuristics had to be encoded at the same time as preparations were being made for the comparative user study. Parts of the dissertation itself were written in parallel to the experimentation phase. While this definitely had to be done due to the time constraints, it is probably what I would have ended up doing anyway, given that there wasn't much to do while waiting for all of the timetables to be created. That being said, it would have been nice to have more time to refine the wording and structure of the dissertation.

Overall, I think that I adapted adequately to the problems that came up during the process of completing this project. It was my first time managing something of this scale by myself, and I feel that I have learnt a lot from it, particularly in terms of being flexible and prioritising the most important objectives when under time pressure.

## 4.4 Laws, Social, Ethical, and Professional Issues

### 4.4.1 Intellectual Property

The only intellectual property created by this project is the codebase for the problem domain and the experiment interface; the hyper-heuristics themselves and the Hyflex interfaces used with the problem domain are not my work and have been taken from an open source GitHub repository. As the problem domain itself is entirely my own work, I own the rights to the intellectual property produced by this project. I have no plans to commercially exploit it, nor do I intend to start some open source project based on it. It will, however, be made publicly available on a GitHub repository under the Apache 2 license, so that others can make use of it if they so wish.

### 4.4.2 Research Ethics

Three aspects of this project have involved human participants: the surveys of Junior Doctors and timetable makers, the interviews with NHS workers, and the study comparing the quality of timetables produced by the hyper-heuristics to those produced by human volunteers.

The surveys were anonymous, but there was still a risk that participants would provide personal information that could be used to identify them in their answers. Two steps were taken to mitigate this risk. The first was to carefully formulate the questions in such a way that made it clear that personal information was not required; in cases where there was a chance of incorrect interpretation, a request not to include personal information was explicitly made.

The second step was to sanitise the responses to the surveys after they had been collected, removing any information that was either not relevant to the question or could in some way be used to identify the participant. For instance, in the junior doctor survey, some participants gave reasons for wanting specific shifts of specific days off. This was sanitised by removing the information regarding why the time off was requested, and altering the request from a specific day to something more general. For instance, a request to not work on Wednesday mornings was altered to "wants a morning of a weekday off". In this way any potential for identification was removed, while still retaining the information required for instance generation.

The interviews were again conducted anonymously: the aim was not to understand the roles of any specific worker, rather, it was to understand the structure of some real world departments, as well as to obtain a better grasp of the requirements of timetables for junior doctors beyond those stipulated by law. As such, there were no questions relating to the workers themselves or their roles. Furthermore, there were no questions specific to their departmental specialty or the region they operated in. Were someone to look at the data collected, they might be able to recognise the general medical field of the department from the layout of its shifts, but they would definitely not be able to identify which hospital it would specifically belong to. There is therefore minimal risk of identification through the information collected by the interviews.

The study comparing automated and manual timetable creation methods poses little to no risk for the participants involved. This is because the only data collected is the timetables that they produced, which do not in any way contain personal data. Furthermore, results are not stored under their names, instead they are labelled by generic names such as "participant x". In this way no data collected by the study can in any be linked back to its participants.

### 4.4.3 Legislation

The first piece of legislation to consider is the Data Protection Act (2018). Given that the project does not collect personal data (see the section on Research Ethics), and steps have been taken to remove such data where it has inadvertently been collected, there is no need to address the issue further than to say that these steps have been taken.

Of more relevance to my project are the two pieces of legislation that form a fundamental component of the problem definition: the 2018 Junior Doctor Contract Refresh, and the European Working Time Directive.

The requirements of the European Working Time Directive that are relevant to this project are as follows:

1. a) The average hours worked per week can be no higher than 48 hours
   b) Doctors must be given 11 hours of rest per day
   c) Doctors must have a day off each week
   d) Doctors must be given 5.6 weeks of paid leave each year

The requirements of the 2018 Junior Doctor Contract Refresh that are relevant to this project are as follows (some repeat):

2. a) The average hours worked per week can be no higher than 48 hours
   b) There can be a maximum of 72 hours worked in any consecutive period of 168 hours
   c) A minimum of 46 hours of rest is required after any number of night shifts
   d) A maximum of 4 night shifts can be worked in a row, and require a minimum period of 46 hours of rest after the last shift
   e) A maximum of 4 ”long shifts” (longer than 10 hours) can be worked in a row, and there must be 48 hours of rest after the last one
   f) A maximum of 7 days can be worked in a row, 48 hours of rest are required after the final shift of the row
   g) Doctors must be given a minimum of 11 hours of continuous rest between rostered shifts
   h) A maximum of 1 in 3 weekends can be worked

These aspects of the legislation are addressed by the hard constraints of the problem definition (Section 2.3); Table 4.1 shows which constraints address which requirements alongside a justification where required.

Table 4.1: Mapping Hard Constraints to Requirements from Legislation

| Requirement(s) | Constraint(s) | Justification |
|---|---|---|
| 1.a, 1.c, 2.a, 2.b | All | According to the NHS staff that I spoke to, it should not be possible to violate these requirements if the other timetabling rules are followed - something that as been borne out by the solutions produced by the hyper-heuristic system. |
| 1.b, 2.g | 1 | - |
| 1.d | 10 | - |
| 2.c, 2.d | 3, 4 | See Section 2.2.3 for a justification of using 48 hours over the stipulated minimum of 46. |
| 2.e | 5 | - |
| 2.f | 6 | - |
| 2.h | 7 | - |

### 4.4.4 Broader Considerations

#### Junior Doctors

The Junior doctors stand to benefit greatly from the research of the project. The first way in which they benefit is through the fact that the automated approach results in better shift coverage and minimises the disparity in shift staffing, something that was identified as a cause of burnout during the survey process. The second is that the system attempts to take into account their preferences: many answers for the causes of burnout were reflected in their preferences. For instance, some doctors would cite a certain number of night shifts in a row as a preference and would then state that being given a greater number in a row was a cause of significant burnout. Furthermore, based on the responses to the survey of timetable makers, many departments do not even attempt to accommodate preferences; an automated system that does so would therefore be an improvement.

The automated system would also improve the capacity for doctors to take leave when they want to. Based on the survey of junior doctors and the interviews with NHS workers, departments often create a timetable with fixed night shifts and then allow doctors to request leave around these fixed points. These night shifts are hard to swap as any swaps made cannot violate the rules for rest regarding night shifts, so doctors often end up in a situation where they would have to swap multiple other shifts in order to move the single night shift: this is clearly not ideal. The proposed system allows for leave requests to be included in the problem instance, with the hyper-heuristic attempting to work around them. This means that the hyper-heuristic could be used as a tool to intelligently decide whether or not to approve leave requests, and should result in more requests being met. As this was identified as a cause of frustration in the surveys, it follows that doing so would improve the well-being and satisfaction of affected doctors.

#### Timetable Makers

The NHS workers currently manually producing the timetables also stand to gain greatly from this project's research. More often than not, the creation of timetables only forms a small part of their overall responsibilities, yet it is a very time-consuming task. Being able to partly delegate the task to an automated system would greatly reduce their workload, allowing them to focus on other tasks; in the case of consultant doctors, for instance, this would free up more of their time for treating patients.

Furthermore, the project does not aim to make them redundant: the produced timetables would still need to be checked to ensure that they don't break laws, and there would almost certainly still be room for small improvements based on their specific domain knowledge. The proposal for the automated system is therefore to provide a very strong foundation on which they can build, with the hope of saving them time and improving the quality of their output.

As with the junior doctors, timetable makers were consulted through a survey. Their responses helped gain an understanding of the challenges they currently face. As such, the project aims to address these challenges by reducing the time that would need to be spent on timetabling, and simplifying the process of adhering to the complicated legal constraints. They also helped influence the output of the system: based on their feedback, it was decided that an output of multiple timetables comparable in quality would be better than just the highest quality one found.

#### Locum Doctors

This project could potentially be seen to have a negative impact on locum doctors, however, I will argue here why this is not the case. The first is that an improved use of resources will help the NHS stay afloat at a time when it is struggling; although they are not permanent staff, the NHS is still their employer and source of income. Due to understaffing, the system is also very unlikely to eliminate the need for locum doctors, it will only ensure that they are used when they absolutely must be; it has been shown that overreliance on locum doctors can be bad for the locums themselves [2]. Furthermore, the system also aims to spread out the locum shifts, based on feedback from the timetable makers. This should ensure a more steady flow of shifts and therefore income for the locum doctors, rather than haphazard periods of great activity followed by periods without work. In this way, while the project aims to limit their use, it will actually benefit them in many ways. It should also be noted that some of the respondents to the junior doctor survey were locum workers, so their views have had an impact on the project.

### 4.4.5 Project Risks

The risks the project poses are minimal. As all that it provides is a problem domain encoding for the problem of junior doctor timetabling, it is difficult to envisage situations in which it could be used for illicit means. What would be more likely to have an impact are the potential unintended consequences of automated timetabling. For example, in trying to incorporate doctor preferences there was a plausible risk of unfairness in preference granting, where the system would be able to minimise the objective function by granting one doctor's requests at the expense of another. In order to avoid this, the choice was made to treat the granting of preferences as a resource to be split evenly between doctors using Jain's Fairness Index (Appendix D).

The other risk stems from the potential for the hyper-heuristic system to be used for modelling shift coverage. This could have negative consequences were the model to be inaccurate. As such, this project does not claim to have produced a system that should be used for modelling in an actual hospital. Rather, it exists as a proof of concept, that such a system could exist, and that a hyper-heuristic approach is capable of handling the high number of constraints involved with the problem. Were such a system to be actually implemented, it would need much greater, direct, cooperation with the NHS at a departmental level. Even following that, the system would still have to be subject to the same rigorous checks performed on the manually made schedules.

### 4.4.6 UN Development Goals

This project aligns with the third UN Sustainable Development Goal: "Ensure healthy lives and promote well-being for all at all ages"[28]. In particular, it has the potential to contribute to Target 3.8. This is because the project aims to improve shift coverage within health departments and to better use the resources available to the health system, both of which should improve access to quality health care. Furthermore, a key tenet throughout the project is to improve the well-being of doctors, a group that is increasingly overworked and overstressed.

## 4.5 Reflections

In conclusion, the project provides a definition for a new problem, the optimisation of which can now be explored; it also provides a proof-of-concept for the applicability of hyper-heuristics to the problem, with an encoded problem domain that should be generalised enough to cater to the needs of a variety of departments with minimal adaptation. Because of this, I feel comfortable labelling the project as a success: it lays the groundwork for further exploration of the area, which was my aim from the start.

In an ideal world I would have been able to work directly with the NHS, and then would have been able to make a real, usable product as opposed to a mere proof of concept. In this regard it could be argued that my project was slightly ambitious for an undergraduate project - at least from the perspective of data gathering. Despite this, I in no way regret undertaking this project: regardless of its moments of difficulty, I found it immensely enjoyable, and the lessons I have learned from undertaking it are ones that I will take with me to any future project I work on, no matter its context.

# Bibliography


[1] Denis Campbell. “NHS staff shortages in England could exceed 570,000 by 2036, leaked document warns”. In: *The Guardian* (2023). URL: https://www.theguardian.com/society/2023/mar/26/nhs-england-staff-shortages-could-exceed-570000-by-2036-study-finds (visited on 11/26/2023).

[2] Gemma Stringer et al. “The use of locum doctors in the NHS – results of a national survey of NHS Trusts in England”. In: *BMC Health Services Research* 23.1 (Aug. 23, 2023), p. 889. ISSN: 1472-6963. DOI: 10.1186/s12913-023-09830-9. URL: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-023-09830-9 (visited on 11/26/2023).

[3] Katie Forster. “One in 30 locum doctors earn £1200 for ten-hour NHS shift, reveal new figures”. In: *The Independent* (2017). URL: https://www.independent.co.uk/news/health/locum-doctors-nhs-120-an-hour-liaison-figures-staff-shortage-hospitals-a7713331.html (visited on 11/26/2023).

[4] NHS. *The NHS Long Term Plan*. Jan. 2019. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf.

[5] NHS England and NHS Improvement. *E-rostering the clinical workforce*. Sept. 2020. URL: https://www.england.nhs.uk/wp-content/uploads/2020/09/e-rostering-guidance.pdf.

[6] Zahraa A. Abdalkareem et al. “Healthcare scheduling in optimization context: a review”. In: *Health and Technology* 11.3 (May 1, 2021), pp. 445–469. ISSN: 2190-7196. DOI: 10.1007/s12553-021-00547-5. URL: https://doi.org/10.1007/s12553-021-00547-5 (visited on 11/26/2023).

[7] Ahmed Kheiri et al. “A hyper-heuristic approach based upon a hidden Markov model for the multi-stage nurse rostering problem”. In: *Computers & Operations Research* 130 (June 1, 2021), p. 105221. ISSN: 0305-0548. DOI: 10.1016/j.cor.2021.105221. URL: https://www.sciencedirect.com/science/article/pii/S0305054821000137 (visited on 11/26/2023).

[8] *Doctors and the European Working Time Directive*. British Medical Association. Feb. 8, 2022. URL: https://www.bma.org.uk/pay-and-contracts/working-hours/european-working-time-directive-ewtd/doctors-and-the-european-working-time-directive (visited on 11/26/2023).

[9] *Rota Rules at a glance*. 2018. URL: https://www.nhsemployers.org/system/files/media/Rota-rules-at-a-glance_0.pdf.

[10] Edmund K Burke et al. “Hyper-heuristics: a survey of the state of the art”. In: *Journal of the Operational Research Society* 64.12 (Dec. 1, 2013). Publisher: Taylor & Francis _eprint: https://doi.org/10.1057/jors.2013.71, pp. 1695–1724. ISSN: 0160-5682. DOI: 10.1057/jors.2013.71. URL: https://doi.org/10.1057/jors.2013.71 (visited on 11/26/2023).

[11] Edmund Burke et al. *HyFlex: A Benchmark Framework for Cross-domain Heuristic Search*. July 27, 2011. DOI: 10.48550/arXiv.1107.5462. arXiv: 1107.5462[cs]. URL: http://arxiv.org/abs/1107.5462 (visited on 11/27/2023).

[12] Edmund Burke et al. *The Cross-Domain Heuristic Search Challenge - An International Research Competition*. Vol. 6683. Journal Abbreviation: LNCS Pages: 634 Publication Title: LNCS. Jan. 17, 2011. 631 pp. ISBN: 978-3-642-25565-6. DOI: 10.1007/978-3-642-25566-3_49.

[13] *HyFlex*. original-date: 2021-03-07T17:43:41Z. Aug. 18, 2023. URL: https://github.com/seage/hyflex (visited on 12/10/2023).

[14] Erfan Rahimian, Kerem Akartunalı, and John Levine. "A hybrid integer and constraint programming approach to solve nurse rostering problems". In: *Computers & Operations Research* 82 (June 1, 2017), pp. 83–94. ISSN: 0305-0548. DOI: 10.1016/j.cor.2017.01.016. URL: https://www.sciencedirect.com/science/article/pii/S0305054817300163 (visited on 12/03/2023).

[15] Wanzhe Hu et al. *A branch-and-price approach for the nurse rostering problem with multiple units*. Nov. 9, 2023. arXiv: 2311.05438[math]. URL: http://arxiv.org/abs/2311.05438 (visited on 04/18/2024).

[16] Fang He and Rong Qu. "A constraint programming based column generation approach to nurse rostering problems". In: *Computers & Operations Research* 39.12 (Dec. 1, 2012), pp. 3331–3343. ISSN: 0305-0548. DOI: 10.1016/j.cor.2012.04.018. URL: https://www.sciencedirect.com/science/article/pii/S0305054812000986 (visited on 04/18/2024).

[17] Shahriar Asta, Ender Özcan, and Tim Curtois. "A tensor based hyper-heuristic for nurse rostering". In: *Knowledge-Based Systems* 98 (Apr. 15, 2016), pp. 185–199. ISSN: 0950-7051. DOI: 10.1016/j.knosys.2016.01.031. URL: https://www.sciencedirect.com/science/article/pii/S0950705116000514 (visited on 12/08/2023).

[18] Burak Bilgin et al. "One hyper-heuristic approach to two timetabling problems in health care". In: *Journal of Heuristics* 18.3 (June 1, 2012), pp. 401–434. ISSN: 1572-9397. DOI: 10.1007/s10732-011-9192-0. URL: https://doi.org/10.1007/s10732-011-9192-0 (visited on 12/03/2023).

[19] K. A. Dowsland and J. M. Thompson. "Solving a Nurse Scheduling Problem with Knapsacks, Networks and Tabu Search". In: *The Journal of the Operational Research Society* 51.7 (2000). Publisher: Palgrave Macmillan Journals, pp. 825–833. ISSN: 0160-5682. DOI: 10.2307/253963. URL: https://www.jstor.org/stable/253963 (visited on 12/03/2023).

[20] Edmund K. Burke et al. "A hybrid heuristic ordering and variable neighbourhood search for the nurse rostering problem". In: *European Journal of Operational Research* 188.2 (July 16, 2008), pp. 330–341. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2007.04.030. URL: https://www.sciencedirect.com/science/article/pii/S0377221707004390 (visited on 12/03/2023).

[21] T. C. Wong, M. Xu, and K. S. Chin. "A two-stage heuristic approach for nurse scheduling problem: A case study in an emergency department". In: *Computers & Operations Research* 51 (Nov. 1, 2014), pp. 99–110. ISSN: 0305-0548. DOI: 10.1016/j.cor.2014.05.018. URL: https://www.sciencedirect.com/science/article/pii/S0305054814001476 (visited on 12/03/2023).

[22] Djamila Ouelhadj et al. "Fairness in Nurse Rostering". In: (). URL: https://pure.port.ac.uk/ws/portalfiles/portal/123556/FairnessinNurseRostering.pdf.

[23] Ahmad Muklason et al. "Fairness in examination timetabling: Student preferences and extended formulations". In: *Applied Soft Computing* 55 (June 1, 2017), pp. 302–318. ISSN: 1568-4946. DOI: 10.1016/j.asoc.2017.01.026. URL: https://www.sciencedirect.com/science/article/pii/S1568494617300418 (visited on 04/18/2024).

[24] Health {and} Social Care Committee House of Commons. "Workforce burnout and resilience in the NHS and social care". In: (June 8, 2021). URL: https://committees.parliament.uk/publications/6158/documents/68766/default/.

[25] *GIHH*. URL: https://github.com/seage/hyflex/tree/master/hyflex-hyperheuristics/hyflex-hh-adaphh.

[26] Tim Curtois et al. "A HyFlex Module for the Personnel Scheduling Problem". In: (). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a4b0fecf854f4602c9380f394e5a3a3a8ebb04de.

[27] Sara Ceschia et al. "The Second International Nurse Rostering Competition". In: *Annals of Operations Research* 274.1 (Mar. 1, 2019), pp. 171–186. ISSN: 1572-9338. DOI: 10.1007/s10479-018-2816-0. URL: https://doi.org/10.1007/s10479-018-2816-0 (visited on 04/18/2024).

[28] United Nations. *Goal 3 — Department of Economic and Social Affairs*. URL: https://sdgs.un.org/goals/goal3 (visited on 04/07/2024).

# Appendices

# Appendix A

# Junior Doctor Survey

**Question 1: What are your preferences for shift patterns?**
E.g. do you prefer working a stretch of nights, or are there specific days that you prefer to work? This could be due to things like caring arrangements, religious observances etc., but please try to avoid divulging any personal information - there is no need to explain why you have these preferences.

Table A.1: Responses to Question 1 (Part 1)

| Respondent | Response |
|---|---|
| 1 | I prefer two or three nights in a row |
| 2 | Stretch of shifts followed by rest days |
| 3 | Not more than 2 days in a row |
| 4 | I prefer not to work too many weekends, I also prefer not to work on [Weekday] nights In regards to nights shift I'd rather do a stretch of 4 in a row than 1 on its own [Redacted] |
| 5 | Prefer to work stretch of nights rather than one off |
| 6 | Prefer 2 nights. Specific days yes, [Redacted] |
| 7 | No preference |
| 8 | Stretch of shifts and single shift patterns |
| 9 | Prefer to work longer hours and avoid working over multiple days. E.g.would happily work 12 HR shifts daily for a few days if it meant I got extra one day off work each week. |
| 10 | Prefer to do nights in 2/3s. |
| 11 | 2 LDs in a row and 2- 3 nights |
| 12 | No sundays |
| 13 | I work [Redacted] less than full time, I like to have 2 half days off a week. I prefer to do no more than 3 on call days in a row and ideally well spaced apart on the rota |
| 14 | Prefer to work stretch of shifts - 4 day shifts in a row but 3 night shifts (4 night shifts is too many in a row for me) |
| 15 | Prefer stretches of similar shifts together |
| 16 | I hate doing 4 days or 4 nights in a row, it's too much. As the [Redacted] I prefer to do one or two days a week rather than a block of long days (e.g. 4 days in a row) |
| 17 | I prefer working 3-4 long shifts in a row and having more days off than working 5 shifts a week of 9am to 5pm |
| 18 | Would prefer work 3 nights maximum at a stretch. Don't like 12 hour shifts. |
| 19 | I prefer stretches of shifts generally but with long rests to follow. |

Table A.2: Responses to Question 1 (Part 2)

| Respondent | Response |
|---|---|
| 20 | Stretches of on calls providing there is adequate rest time/ compensation for out of hours working |
| 21 | I rather an ‘on-call block’, I rather most my oncalls within a short period of time, this allows planning other parts of the year for specialty training, holidays etc |
| 22 | Find stretches of nights very difficult but single nights also disrupt the routine because you are constantly switching. |
| 23 | Prefer a single day on-call per week with the caveat that this is easily moveable to other days should I need to book annual leave. I hate night shifts and would not do them at all if possible. |
| 24 | I just prefer consistency day to day, then it can change week to week |
| 25 | *I like to have [Weekday] off. *3 nights in a row as a max and avoid single nights *8 hours maximum for days and evenings, 10 hours for nights *at least half hour overlap between shifts to allow handover *Having the opportunity to do admin/ study work from home once a week *1 in 4 weekends *Evening shifts should finish at 10pm, not any later *I would be happy with 2-week day shifts, one week of evenings and one week of 3 nights in a row |
| 26 | I prefer nights to be grouped. I think there should be weeks without long days/on calls in them in order to maintain SDT that is 9-5. |
| 27 | I would prefer rolling and predictable shift patterns, with potentially longer hours overnight/weekends, but fewer overnights/weekend shifts in the end! |
| 28 | 2-3 nights, prefer certain days or nights on call either for family or to minimise impact on health and wellbeing |
| 29 | Preferably 7 am till 4pm day shifts 2pm till 10 pm twilight 10pm till 8am night |
| 30 | I dont like stretch of nights, although i prefer 2 nights and 3 off and thn 2 nights. |
| 31 | I prefer to work a maximum of 3 long shifts in a row (e.g nights or long days) and these should be followed by normal days, rather than more on call shifts. |
| 32 | No particular preference but will rather not do nights( Unfortunately I don’t have a choice) |
| 33 | I like working a stretch of nights as it makes swapping an entire set easier. Swapping individual shifts is difficult as we would need to take into consideration post on call time off. |
| 34 | Specific days/shift patterns that I prefer to work |
| 35 | Prefer nights grouped in small clusters rather than one night shift every week |
| 36 | Prefer Maximum of 3 nights |
| 37 | Volunteering for [Redacted] with meetings on [Weekday] evenings; Timetable all over the place but try to occasionally swap out of [Weekday] evenings; No particular preference for shift patterns |
| 38 | Prefer to work days rather than nights. 2 long days then a day off in the middle, followed by another set of 2 long days and the remaining 2 days off. |

**Are there any timetable-specific causes of burnout?**
This could be something like a stretch of night shifts of a given length, or too long without time off of a certain length. Please try to only include information related to timetabling, you don't need to explain why these aspects would cause burnout.

Table A.3: Responses to Question 2 (Part 1)

| Respondent | Response |
|---|---|
| 1 | Five shifts in a row is tiring |
| 2 | Rotas where full time work requires more than 40 hours a week - this doesn't leave enough rest days and too many weeks with 6 working days |
| 3 | 4 nights in a row |
| 4 | I think for me it is more the acquity of the shifts. If they are under staffed and I'm not getting an adequate break and running from sick patient to sick patient and not feeling I'm being taught anything I'm more likely to burn out. When training is cancelled and it feels like we're just fire fighting I feel much more vulnerable to burn out. |
| 5 | Not long enough off after nights Too long stretches of shift without adequte break |
| 6 | More than two nights is harder to recover from. I used to work a rota where we would work six days in a row and that led to burnout |
| 7 | Stretch of night , nights should not be more than two in a row. |
| 8 | Yes, regularly working weekends and nights |
| 9 | Not enough break between shifts. E.g. finishing night shift on Monday morning and then having to go into work Tuesday am or wed am |
| 10 | Working too may weekends in a row. Weeks which are the maximum of the rota eg 72 hours |
| 11 | Long working weeks and nights frequently spaced with long days in between |
| 12 | 4 long days back to back |
| 13 | Not having enough doctors for a particular shift, working a lot of on call shifts in close proximity to each other |
| 14 | Repeated swapping between days and nights in a period |
| 15 | No main issue is not being able to take prolonged stretches of annual leave due to eg random night shifts which can't be swapped |
| 16 | Short number of rest days post night shifts is really difficult. |
| 17 | I think burnout happens when one has to do multiple shifts a week and commute long distances. |
| 18 | 12 hour shifts leads less productivity after 10 hours of shift and obviously more stress |
| 19 | any night shift run that is more than 3 causes burn out to me, any shift more than 8 hours. |
| 20 | Long stretches of on call shifts/ prolonged working (e.g. unpredictable long days in theatre) without adequate rest time, no rota'd time to complete other training requirements eg portfolio, audit, research and admin time- so that these are done on evenings and weekends or rest time. |
| 21 | They should be crossover of shifts- the handover leopard means often working 13 hours + and with commute time not enough uninterrupted rest. |
| 22 | To much variation in starting times across the same week. Worst is starting with late shifts then moving on to early |
| 23 | From experience I don't enjoy 24h/48h non-resident on calls unless there are guaranteed zero days (usually there are not for non-resident on calls) |
| 24 | Flipflopping from nights to mornings is a jiller |

Table A.4: Responses to Question 2 (Part 2)

| Respondent | Response |
|---|---|
| 25 | *12-hour shifts *4 nights *changing shift patterns within the same week *Not knowing the rotas till last minute |
| 26 | Too many on calls in a row without sufficient days of rest in between. Following a busy on call that finished at 2130, finishing late getting home late then having to get up for normal shift following day not feeling rested. |
| 27 | Driving home after a night shift through the morning rush hour is a risk that should be avoided. I would encourage all sites to provide accommodations on-site for the night worker to be able to sleep in before driving home. |
| 28 | Single night shifts, rapid turnover from nights to days, more than 2 long days in a row, 72 hour weeks, 4 nights in a row |
| 29 | Day shifts should be followed by twilight and then nights. Going back to days after nights causes burnout. Instead there should be twilights after night and then back to days. |
| 30 | Stretch of nights |
| 31 | Long days or night shifts (12.5hrs) cause burnout quickly, I would say that more than 3 long shifts followed by inadequate rest periods (for example, 1 zero day) and then returning to more on call shifts causes burnout. Inflexibility of rotas is a huge issue, I need to be sure that I can attend weddings and training courses/conferences etc, if I give enough advance notice. |
| 32 | Long stretch of nights. |
| 33 | It's frustrating when we have to organise swaps with multiple people to get certain periods of leave approved. Would be great if there was a AI system that could do this. |
| 34 | No |
| 35 | Too many shifts over 8 hours in a row without breaks Cycling from nights into earlies then straight back into late shifts Random shift patterns with no progression of start time, or jumping back and forth between start times Rotas with too many different shift patterns Having days off spread apart so only having one day off at a time Too many shift timings that can't have annual leave taken on, limiting the days you can have AL |
| 36 | Issue is around not being able to take leave on nights or weekends. Often 4 nights in a row. Switching between different twilight shifts in a week eg 1-10, then 5pm-2am etc |
| 37 | Not for me personally |
| 38 | Stretch of 4 nights/long days in a row. |

**Question 3: How do you tend to take your annual leave?**
For instance, would you take it in a few big chunks, or would you spread it out across multiple weekends?

Table A.5: Responses to Question 3 (Part 1)

| Respondent | Response |
|---|---|
| 1 | I prefer taking in a stretch for three weeks which includes my regular off |
| 2 | Spaced out, timed with shifts to maximise time off |
| 3 | One big chunk with others scattered |

Table A.6: Responses to Question 3 (Part 2)

| Respondent | Response |
|---|---|
| 4 | In large chunks. Y family aren't closeby so I would rather have it in chunks so I can go and see them. Currently I only have to be told my rota a few week weeks in advance. For example I rotate to a new hospital at [Redacted] and I recieved the rota from [Month] till [Month] last week. This [Season] and [Season] I've got quite a few close friends and family weddings and family big birthdays and it's been difficult being unsure if I will actually be able to attend them. |
| 5 | Try to maximize big chunks by taking leave when already only working 1-2 days per week then spread out |
| 6 | Usually medium sized chunks of a week or 3-4 days. You can't take annual leave on weekends in my job |
| 7 | Spreaf it out across multiple weekends |
| 8 | Big chunk so that I can travel to see my family back home. Sometimes it is not possible due to nights and weekends |
| 9 | Large Chunks due to travel abroad |
| 10 | One big chunk and then long weekends through out the block |
| 11 | Chunks for holidays and odd days for the rest |
| 12 | Depends on the month |
| 13 | Spread out in little chunks- a few days here and there |
| 14 | Few big chunks, using post night shift days to extend to ensure maximum days off , if any remaining days will take them during busy parts of rota |
| 15 | Preference is a few blocks so having eg 10-14 day stretches off, with a few days saved to make long weekends/for specific events etc |
| 16 | Both as I am given annual leave yearly as a [Redacted] compared to a [Redacted]. |
| 17 | big chunks at a time |
| 18 | Quite mixed, would prefer at least 3 week to go home and come back and few in between so get break from hospital. |
| 19 | big blocks of annual leave |
| 20 | Variable- mostly big chunks however taking smaller breaks to allow attendence at family events, religious holidays, sometimes to attend courses and conferences |
| 21 | Depends on the year- generally one big chunk for a holiday and the rest spread out for short breaks. |
| 22 | Mixture of the two |
| 23 | Usually I take a big chunk to take a trip away for 1-3 weeks then spread the rest out taking 1-2 days at a time. |
| 24 | Couple bigger chunks then extended weekends where I can |
| 25 | *Mostly big chunks, occasionally spread to allow some rest *worst of all is built-in annual leaves |
| 26 | I tend to take a mix of a long holiday of 2-3 weeks then scatter the rest over long weekends. |
| 27 | I would rather have multiple smaller chunks of AL than a long, continuous AL period. |
| 28 | Some big chunks with the rest used to break up bad stretches on the rota |
| 29 | Big chunks once every 6 months, what ever is left over, can be split whenever |
| 30 | in big chunks |

Table A.7: Responses to Question 3 (Part 3)

| Respondent | Response |
|---|---|
| 31 | Staffing is usually poor which limits the days available to take leave. For this reason, AL is usually taken on random days here and there. I would prefer to be able to take longer periods off work, currently we have to use our yearly leave split equally between our rotations (for example, 9 days every 4 months), some years, I would prefer to take a longer holiday once a year. Sometimes I like to take AL in between on call shifts or after a particularly long set of on call shifts to allow adequate rest. |
| 32 | Few chunks usually 7 to 10 days |
| 33 | I take usually a big chunk and then a few ones seperately |
| 34 | Few big chunks |
| 35 | A mixture of a few days here and there with a longer block of AL for at least 1 week off or 2 |
| 36 | Spread out to maximise time with [Redacted] |
| 37 | I try to take chunks of time for a big holiday; if not possible, I use the days off in between shifts to fly to [Continent] and anywhere else I can squeeze in; e.g. went to [Country] for 24 hours, [Country] for 48 hours - went to airport from work after a shift, flew back and started work either a few hours after or the day after return flight |
| 38 | In chunks and spread out in between stretches of night shifts. |

**Question 4: How does training impact your rota schedule?**
How many hours of training would you have per week (or planning period if the rota is allocated differently – please specify the timeframe), and is there a time of day at which such sessions normally take place?

Table A.8: Responses to Question 4 (Part 1)

| Respondent | Response |
|---|---|
| 1 | 2 hours educational development time |
| 2 | I haven't made teaching since August. Getting procedures done out of hours is more difficult, but leave can only be taken in normal hours and therefore normal shifts are significantly reduced |
| 3 | 1 day off per month for self study time and 1 day teaching day |
| 4 | We currently have teaching every other [Weekday] for around 4 hours. However if I am rotaed to be off but do want to attend teaching then I do. Then you loose out on an off day which you can try and claim the time back from. But the logistics of doing that with medical staffing is sometimes more hassle than it's worth. |
| 5 | Frankly training comes 2nd to rota |
| 6 | Four hours per week for full time doctors, fixed day each week |
| 7 | N/a |
| 8 | 8 hours every 2 weeks |
| 9 | I have to cover on-calls in [Redacted]. This has very little training element compared to usual normal working day. |
| 10 | Edt seems to be allocated at random would be nice to choose it myself around planned activities. |
| 11 | I'm in a training job, and I get [Weekday] afternoon weekly teaching when I'm not on nights or on call |
| 12 | Wednesday 2-4pm |

Table A.9: Responses to Question 4 (Part 2)

| Respondent | Response |
|---|---|
| 13 | Our training days tend to be spread out on different days of the week with variable notice |
| 14 | [Weekday] afternoons 1.30 to 4.30 is teaching for us |
| 15 | We have [Weekday] every week for non clinical time/teaching |
| 16 | One whole day a month on a [Weekday] is my protected teaching time for my speciality. I also have half an afternoon every two weeks for speciality training day. If I am on call I can't attend. |
| 17 | approximately 3 hours per week, and most of it is in the afternoon 1pm-2pm ([Weekday], [Weekday], [Weekday]); We also have a regional training day per month and that is a fulll day between 9am to 4pm |
| 18 | On paper it says 48 hours per week. Normally takes place on [Weekday] either in morning or noon. |
| 19 | morning and evening shifts before midnight are the best opportunity for training |
| 20 | training is fitted in around the on call schedule- usually on normal working/ elective days (8am-5pm) |
| 21 | Training does not effect rota schedule. if rotered on call them training opportunity is lost. |
| 22 | 8 |
| 23 | Currently 1 day per week is teaching. If I am on call I attend teaching then go to my on call afterwards. |
| 24 | Really varies, [Weekday]s are set for teaching but outside that it could be nothing |
| 25 | * 8 hours per week |
| 26 | 20 per cent for portfolio commitments/higher degrees/Self directed learning/clinic attendance. Preferably afternoon with flexibility for working from home if possible. |
| 27 | [Weekday] are weekly training day with the morning local sessions and regional afternoon ones. |
| 28 | 17 |
| 29 | Day shifts must be purely designed for training. All rest can be for service provision |
| 30 | 8 |
| 31 | We have mandatory teaching one afternoon per week, if we are on call or nights we cannot attend. We have no allocated time for mandatory clinics. |
| 32 | When on call it is difficult to participate in training activities. Training sessions are usually in the afternoon |
| 33 | Training is best provided during normal days when we are not on the on-call rota (in most cases that means not holding a bleep) - this frees us to seek out training opportunities. |
| 34 | [Weekday] afternoons - Once per month + training day once per month. |
| 35 | Half day of EDT per week, half day of teaching per week, another half day of teaching once a month (both teaching usually [Weekday]s), department specific teaching on top of this |
| 36 | 8 hours week, not including [Weekday] afternoon regional teaching each week |
| 37 | Occasionally have a monthly regional teaching session; I have to take annual leave for teaching sessions and courses |
| 38 | Reduced amount of time/opportunities to spend in outpatient clinics. |

**Question 5: How many shifts do you usually work per week?**
(or planning period – please specify the timeframe)

Table A.10: Responses to Question 5

| Respondent | Response |
|---|---|
| 1 | 4-5 |
| 2 | 5 |
| 3 | 4 |
| 4 | 5 |
| 5 | 5 |
| 6 | 3-4 |
| 7 | Four days per week |
| 8 | 4-5 shifts |
| 9 | 5 nwd but some weeks it can be 4 nights in a row |
| 10 | 5 |
| 11 | 4-6 |
| 12 | 40 |
| 13 | 4 |
| 14 | Average 4 shifts + 1 SPA day per week |
| 15 | 5 |
| 16 | 5 days a week |
| 17 | 4-5 shifts |
| 18 | About 4-5 excluding our teaching days. |
| 19 | 4 shifts |
| 20 | 1 in 16 on call rota |
| 21 | 5-7 |
| 22 | 5 |
| 23 | Currently 1 in 6 on call rota with day weekends. No night shifts |
| 24 | 4 |
| 25 | *4-5 |
| 26 | 5 |
| 27 | Usually 4 shifts/week, sometimes 5 shifts a week when working over the weekend |
| 28 | 3-4 |
| 29 | 5 |
| 30 | 4 |
| 31 | 87 shifts over 4 months, I am LTFT at [Redacted] |
| 32 | Usually 40 to 60 hours |
| 33 | 5 |
| 34 | 1-2/week on average |
| 35 | 5 |
| 36 | 3-4 |
| 37 | 4-6 shifts a week |
| 38 | 4 to 5 |

**Question 6: How long is your average shift?**

Table A.11: Responses to Question 6

| Respondent | Response |
|---|---|
| 1 | 10.5 hrs day and 11 hrs night |
| 2 | 9 hours |
| 3 | 10 hours |
| 4 | 9 to 12 hours |
| 5 | 10 |
| 6 | 9 hours |
| 7 | 8-10 hours |
| 8 | 8 hours |
| 9 | 8-9hrs |
| 10 | 9 hours |
| 11 | 8-5:30 normal day. 8-8 long day. 7:30-8:30 night |
| 12 | 8 |
| 13 | 8 hours |
| 14 | 9 hours |
| 15 | 9 hours |
| 16 | 8 hours |
| 17 | 8-12 hours |
| 18 | 9 hours |
| 19 | 8-10 hours |
| 20 | Normal working days are 8am-5pm, on calls are 12.5hrs |
| 21 | 8.5 short day 12.5 long day |
| 22 | 9 hours but have done 10, 11 & 12 in other places |
| 23 | 8h normal day 12h on call |
| 24 | 9hours |
| 25 | *8 hours for days and evenings, 10 for nights |
| 26 | 9hours |
| 27 | 9 hours |
| 28 | 9.5h |
| 29 | 8 hours |
| 30 | 9 to 10 hours |
| 31 | NWD is 8hrs, Long day or night shift is 12.5hrs, multiple different shift types ranging between these two, with different start and finish times-I can provide more info on shift types if useful |
| 32 | 8 to 12.5 hours. |
| 33 | 9 hours |
| 34 | 12.5 hours |
| 35 | 9 hours |
| 36 | 9 houes |
| 37 | 9 hours |
| 38 | 12 hours |

**Question 7: Is there anything else that you think should be considered when rotas are being created?**
Please try to refrain from providing information that is very specific to you; this question is only asking for overarching considerations that might have been overlooked.

Note: Question was Optional

Table A.12: Responses to Question 7 (Part 1)

| Respondent | Response |
|---|---|
| 1 | Not to have nights during deanery teaching and looking for swaps |
| 2 | Number of rest days required and type of work being undertaken |
| 4 | No particularly I do think it is a hard thing to get right. I thinknthe constant rotations of doctors in training make it particularly hard. |
| 5 | Self rostering |
| 6 | Most drs are willing to be flexible so good communication with each dr and collaboration |
| 7 | No more than two to three nights in a row |
| 8 | Asking in advance if someone has annual leave preference that would be great such as exam or important event coming up so that you don't have to arrange swaps last minute. |
| 9 | Have some slack in the system as invariably people will call in sick, be off for unforseen circumstances |
| 10 | Looking at the overall pattern of the rota. For example going from nights straight into long days. Or also looking at where annual leave can be easily taken due to off days etc |
| 11 | The [Redacted] department has recently started using DB rotas which is an AI system allowing you to build in leave and not on call shifts which has greatly improved my work life balance |
| 12 | Odd shifts like 11am till 12noon do not give you enough time to do anything in the morning |
| 13 | Because of less than full time training becoming more common there are vacant shifts automatically created alongside the unexpected vacant shifts from sickness. |
| 14 | The lack of flexibility to swap weekend/night shifts. Difficult to swap with other people as the days off afterwards are also required |
| 15 | Important dates e.g. weddings |
| 16 | Break times. Sometimes we don't get to have lunch or prayer breaks in a day and I feel those should be protected times for employees. Even an hour in a 12-hr shift can be quite helpful. |
| 17 | I would prefer same number of people working at any given day or night. Suddenly night and weekend becomes less staffed, given how NHS has evolved over few years I think we should get same number of team members for nights and weekends. |
| 18 | Different training requirements at different levels of seniority- e.g. more elective work required towards tge end of training |
| 19 | In training there are two rotas, specialty training and internal medicine/ oncall training. There is a disconnect between rota teams, thus more training opportunity is lost or it relies on the person to make difficult swaps. |
| 20 | Is it possible to have different shift lengths? some colleagues prefer to work more days but shorter shifts but some prefer longer so they work less days |

Table A.13: Responses to Question 7 (Part 1)

| Respondent | Response |
|---|---|
| 21 | The easiness of being able to swap into different slots if we need to take leave at a specific time - if we give sufficient notice it should not be our responsibility to arrange cover even if we are on call. |
| 22 | To give trainees an extra day after a set of nights or a weekend to reset properly. The first day off is always disjointed and horrible then by the time you have your next day, you're often back in work. It's so jarring, never ever feel like you've had a rest |
| 23 | Childcare flexibility |
| 24 | Ideally, the pattern should ensure the continuity of care, by arranging a shift time start in the last hour of another shift end. That would allow for fewer patient handovers between clinicians and allow the leaving clinician to pick up any potential loose ends and go home on time. |
| 25 | Many bespoke rotas don't take account of wellbeing e.g. you get the time off you asked for but have no control of the run of shifts given to you, it is then too late to use leave to break up bad runs of shifts as you've used it for your preference requests. |
| 26 | Half of the rota should be applied after asking preferences, rest core shifts can be given as per the depts requirements |
| 28 | Having rotas available as far in advance as possible, I see no reason why a full years rota couldn't be supplied. This will allow adequate time for planning and more potential time for swaps to be made if necessary. Important events should be taken into account (e.g exams, weddings, religious holidays etc), if advance notice is given, a person should not be scheduled on call. Mandatory training requirements eg clinics should be scheduled onto the rota. Mandatory training (eg stat/mand) should have allocated time. [Redacted] should also have admin/academic time scheduled on the rota, there are lots of non-[Redacted] activities that are part of our job plan that we have to do on our own time. |
| 29 | Individual preferences and pre requested leaves |
| 30 | If leaves are provided before the rota is created, it would be great if this is incorporated into the rota so that we don't have to find swaps later. I am now moving to a new hospital where they've taken into consideration my leave dates and put that into the rota. This means that the rota is not a rotating one but it caters to my leaves which I value more. |
| 31 | Staffing/more manpower to run the rota without gap and allow flexibility for trainees to work - Has always been recognised as a huge issue and challenging to resolve. |
| 32 | Much earlier rotas so life can be planned in advance (not just 6 weeks) Being contacted by rota teams well in advance of rotas being distributed to ask for any annual leave dates that are accommodated in the rota, too many times have early requests for annual leave been denied and then you're placed on night shifts the week you asked for off and struggle to swap the shifts with colleagues |

# Appendix B

# Timetable Maker Survey

**Question 1: What is the length of an average planing period for a middle grade doctor rota?**
i.e. How many weeks do you cover with each rota?

Table B.1: Responses to Question 1

| Respondent | Response |
|---|---|
| 1 | 20 week rota |
| 2 | we have 12 weeks rolling rota so the doctor is aware of their ongoing rota pattern |
| 3 | 6 weeks |
| 4 | 6 |
| 5 | 4-24 |
| 6 | 13 or 26 weeks |
| 7 | 8-12 weeks |
| 8 | 3 months |
| 9 | 26 weeks |
| 10 | 2 - 3 weeks |
| 11 | 1 hour |

**Question 2: How long on average does it take you to produce a timetable by hand?**
If there are any tools you use to speed up the process, please describe them. In order to avoid the risk of identifying yourself, if these tools are specific to your place of work, please just refer to them as "in-house tools".

Table B.2: Responses to Question 2 (Part 1)

| Respondent | Response |
|---|---|
| 1 | in house tools |
| 2 | as this is a rolling rota, it is not time consuming to produce |
| 3 | 6 hours for the first draft, multiple revisions usually needed afterwards |
| 4 | 8 |
| 5 | rota system with in built rota rules |
| 6 | We use rota software DBRotas |
| 7 | I use excel spreadsheets x 2. it is quick to copy and paste but takes considerably longer to populate |
| 8 | 3-4 hrs in house tools |
| 9 | 10-15 hours depending on the number of less than full time colleagues |

Table B.3: Responses to Question 2 (Part 2)

| Respondent | Response |
|---|---|
| 10 | 1-2 weeks as the rolling pattern is already there. If for any reason the pattern changes this would take about 1 month to sort. This can sometimes differ as with our area each registrar has to follow their consultants job plan so I do 16x each rota as opposed to 1 overall rota covering everyone. The only pattern they follow is the on call pattern and around that it is individual |
| 11 | 8 hours |

**Question 3: Are there any preferences with regards to shift patterns that have been very common amongst middle grade doctors?**
i.e. are there preferences for specific days of the week or times of day? Please try to avoid including any information regarding the middle grade doctors in question, we don't require a justification of the preferences.

Table B.4: Responses to Question 3

| Respondent | Response |
|---|---|
| 1 | lots of less than full time Drs - wanting different days off within the week/Rota |
| 2 | our doctors are not given choices as such - unless they are employed part -time in which case they advise which weekdays they are unable to work - the majority of our doctor work the 12 week base rota |
| 3 | Stretches of the same type of shifts Shifts getting progressively later in the day Some like to do all nights in a single stretch, others a maximum of 2 nights in a row |
| 4 | Some doctors work less than full time and therefore have specific days off each week (usually 80% with one day off or 60% with two days off). Most will choose to have either Monday or Friday off. Most doctors prefer day shifts to lates or nights. |
| 5 | Mondays or Fridays off |
| 6 | No one wants to work a long day Friday! The majority of LTFT doctors chose Friday as their non working day |
| 7 | Not within my specialty |
| 8 | yes |
| 9 | Generally less than full time colleagues prefer mondays or fridays as non-working days |
| 10 | Within our areas a lot of middle grades require quite specialist areas. e.g. they need to do x amount of theatres in a certain area for their portfolio or x amount of clinic time. |
| 11 | yes |

**Question 4: How does training impact the rota schedule for middle grade doctors?**
How many hours of training would there be per week (or planning period if the rota is allocated differently – please specify the timeframe) and is there a time of day at which such sessions take place?

Table B.5: Responses to Question 4

| Respondent | Response |
|---|---|
| 1 | there are variable hours for training for middle grades mostly for clinics so will vary between speciality |
| 2 | the trainees discuss with me as rota coordinator and we decide, rota allowing, when they can take their education training hours - for example ST4 - ST6 have Wednesday off as teaching /EDT. ST3 are allowed 4 hours per week, so usually have a day off alternate weeks; CT1 allowed 2 hours per week so take a day off per month. FY2 have 2 hours teaching plus 2 hours EDT every Tuesday |
| 3 | 12 hours per week for senior middle grades - this means they can only do nights Thurs-Sun 4 hours per week for junior middle grades scheduled as a whole day a fortnight plus an additional day per month |
| 4 | Four hours a week as fixed protected teaching plus four hours per week of personal development time (to go to clinics or do audits or work on their portfolios) |
| 5 | differs for different grades and specialties - most would rota it as a normal working day then allocate them training as required. |
| 6 | Three to four days a week on average |
| 7 | Can usually accommodate training without impact on rota |
| 8 | no |
| 9 | No specified amount |
| 10 | Our trainee middle grades go to teaching each friday afternoon, it was an issue until we rejigged the rota and now it is no bother as we just make sure our trust doctors are rostered for the afternoon where they should have been. |
| 11 | yes |

**Question 5: Do you have any rule-of-thumb techniques that help you when constructing a timetable?**
(The technical term for this is 'Heuristic', for further explanation, see: https://www.verywellmind.com/what-is-a-heuristic-2795235)

Here are some examples of the type of thing this question is looking for:

- Allocating certain shifts first because you know that they'll be difficult
- Allocating all the shifts for one individual before moving onto the next
- Allocating the maximum number of nights in a row at once

If possible, please categorise these rules based on whether they speed up the process of creating the timetables or if they improve aspects such as fairness.

Table B.6: Responses to Question 5 (Part 1)

| Respondent | Response |
|---|---|
| 1 | we follow a generic Rota that is sent to us from HR- So everyone follows their predecessor for the Rota |
| 2 | as we have a set 12 week rolling rota, the above does not really apply |

Table B.7: Responses to Question 5 (Part 2)

| Respondent | Response |
|---|---|
| 3 | For trainees rota templates are used to ensure fairness, these are entered first a single person at a time. Then shifts are removed for annual and study leave. Then for each day doctors are reallocated to spread them out through the days concentrating on nights as these are most important to fill. Might offer to exchange two or three day shifts for a single night shift, ensuring at least one senior middle grade on the night. Then add in permanent locums starting with those who want to work nights, lates and weekends, trying to fit their preferred pattern of working. |
| 4 | Rotas are usually agreed with HR to ensure they are fair - eg 60% LTFT trainees do 60% of each type of shift and oncalls. Inputting one person's rota at a time is quickest |
| 5 | ensuring nights / weekend shifts are balanced out within the rules, and a period of a 2 or 3 weeks where normal working days are rotated to allow a time for drs to take leave. When building an ltft rota, allocating the nights and weekends first as they are more difficult to place. |
| 6 | We use automated rota software. |
| 7 | *core activity to be covered *prioritisation of activity *annual leave/sickness gaps |
| 8 | divide shifts equally among all junior doctors |
| 9 | I use a master rota. Where possible I put 2 ltft colleagues into the same slot with opposite working days, divide the shifts in half and then add additional shifts to make up the numbers |
| 10 | We follow a pattern for the on calls so this is already set. I make a blank rota with this pattern on for all 16 doctors for 6 months each. I then go doctor by doctor, matching to their individual consultant job plan. At the end I go back to make sure all shifts needed are covered. |
| 11 | the middle grade rota is usually tailor made to each individual |

**Question 6: Is the process for creating a timetable iterative?**
i.e. would you make multiple versions of a timetable before publishing it, making alterations at each stage based on feedback?

Table B.8: Responses to Question 6

| Respondent | Response |
|---|---|
| 1 | yes |
| 2 | no |
| 3 | yes |
| 4 | Yes, after inputting the rotas swaps and adjustments are made to try to cover as many of the shifts as possible |
| 5 | yes |
| 6 | yes |
| 7 | No |
| 8 | yes |
| 9 | yes |
| 10 | I made a checklist on excel that I tick off for each person, once all boxes are ticked i publish |
| 11 | yes |

**Question 7: What method do you use to assess whether a timetable meets all of its legal requirements?**

Table B.9: Responses to Question 7

| Respondent | Response |
|---|---|
| 1 | they are agreed by HR- so once they get to me they are compliant with working hours etc |
| 2 | Medical Resourcing assess the rota to ensure it is compliant |
| 3 | Checking manually that there are 48 hours off after every stretch of nights and that shifts have adequate rest between them - sometimes missed when moving from one week to the next as on separate sheets of paper |
| 4 | Initial rotas are checked by HR. The official software used by the trust will flag up illegal shifts but inputting the timetable onto this software is done last as it is slow and not very easy to use. |
| 5 | the rota system has the drs rota rules inbuilt and it highlights if a rota has failed for any of these rules being broken such as too many hours or not enough rest between shifts. |
| 6 | The software does a live compliance check so if we move shifts around that make that pattern of working non compliant it flags as red immediately |
| 7 | N/A |
| 8 | usually through HR |
| 9 | Use the good rostering guidance, then medical resourcing will check it is compliant with software |
| 10 | I know the rules like the back of my hand having done the role for 8 years, however I do check each rotation on the BMA to make sure nothing has changed. I also cross check with our HR if i need too. |
| 11 | being aware of the national and BMA rules well |

**Question 8: How do you assess whether a timetable meets the preferences of the middle grade doctors?**

Table B.10: Responses to Question 8

| Respondent | Response |
|---|---|
| 1 | all preferences for days off are requested and normally agreed |
| 2 | not usually an issue as it is a set rota |
| 3 | Send them the draft rota and ask them to confirm they are happy |
| 4 | They are asked to discuss their initial rota with their supervisor, otherwise they just ask for swaps if needed |
| 5 | if a rota meets the rules and department requirements, it wouldn't first be agreed with the drs unless it was mid rotation, before being rolled out as you cannot please everyone. |
| 6 | We did a user satisfaction survey |
| 7 | Meeting the needs of the service is more important than accommodating preferences but I do allow this where possible |
| 8 | discuss preferences and training requirements beforehand |
| 9 | Not really possible |
| 10 | This is a conversation had before they start and then we accommodate where we can |
| 11 | by discussing personally with the doctors |

**Question 9: How far in advance are timetables produced with respect to the planning period?**
E.g. a week before, a few days etc.

Table B.11: Responses to Question 9

| Respondent | Response |
|---|---|
| 1 | 6 weeks before |
| 2 | the base rota is set for 6 months at a time |
| 3 | Ideally meant to tell the doctors 6 weeks in advance. In practice more often 7-10 days for the final version |
| 4 | 2-4 weeks |
| 5 | should be 8 weeks notice but can often be less |
| 6 | We aim to start the process 12 weeks before the rota period starts as first step is collecting user preferences, we aim to publish 9 weeks before that rota period starts |
| 7 | 8-12 weeks |
| 8 | 4-6 weeks before start of rotation |
| 9 | Aim for 8 weeks before |
| 10 | 6 months at a time for middle grades |
| 11 | 4 months in advance |

**Question 10 : Can timetables be changed after they have been published? Under what circumstances would this be done?**

Table B.12: Responses to Question 10

| Respondent | Response |
|---|---|
| 1 | Can be changed but only in exceptional circumstances and with 6 weeks notice given to the Drs |
| 2 | obviously the rota is updated as and when annual leave/study is requested |
| 3 | Yes, usually for sickness or urgent leave. Sometimes by mutual agreement when doctors swap shifts for their own convenience. Occasionally for study leave if a place becomes available on a course at short notice. |
| 4 | Yes - due to sickness or study leave |
| 5 | live rotas can't be amended on the system, it needs copying and amending before replacing the live rota. If the department had a vacancy they may want to change the rota to take off a slot. |
| 6 | Yes we are very flexible and allow swaps and changes as long as it's fair and compliant |
| 7 | Yes - via myself who supports the service |
| 8 | yes. requests to make changes from middle grade doctors |
| 9 | swaps amongst colleagues but would all need to be agreed |
| 10 | Yes, the rota pattern itself cannot be changed within a rotation unless we give 6 weeks notice to the doctors but if individual normal working days needed amending I would do this and resend out |
| 11 | we do swaps due to our training commitments and leave |

**Question 11: If you answered yes to question 11, do you have specific rule of thumb-methods (see question 6) that you use in response to the circumstances that necessitated the change?**

Table B.13: Responses to Question 11

| Respondent | Response |
|---|---|
| 1 | yes - it will only be changed if the need is agreed for cover of the rota |
| 2 | the 0800-1800 and 2200-0800 shifts are the most important to fill |
| 3 | Yes, move doctors across the day again to try and balance what cover is left |
| 4 | Try to cover the most essential shifts such as lates and nights by moving doctors around |
| 5 | the main check is whether the rota is compliant, after that we have to meet the notice period of 6 weeks minimum, if changing a rota |
| 6 | We try to accomodate all requests but if cover was too thin we would say no and explain that was the reason why. |
| 7 | If can accommodate all activity then changes will be allowed |
| 8 | no |
| 9 | no |
| 10 | We work on needs of the doctor whilst also making sure everything is safely covered |
| 11 | no |

**Question 12: What considerations, besides limiting their use, do you make when allocating locum shifts?**
For instance, are there locums you hire on a regular basis whose preferences you try to accommodate?

Table B.14: Responses to Question 12

| Respondent | Response |
|---|---|
| 1 | We have locums but they are not normally on the on call rote and will only fill vacancies |
| 2 | we do use regular locums but they are offered the vacant shifts - their preferences are not accommodated, although we may adjust the hours slightly when necessary |
| 3 | We use regular locums we know are clinically safe, particularly if senior enough to be in charge of nights. I ask them their preferred pattern of working or one or two of them tell me which days they are available and I slot them in. I find by doing so they are happy to stay and work in our department even if they could be paid more somewhere else, and are more likely to help out in a crisis situation. Word gets around and other good locums join as well, we get to know their strengths and weaknesses and how to use them best as part of the team. |
| 4 | Trying to avoid having a day where only locums are present. If a locum prefers to do particular shifts such as nights will offer the trainee on the nights if they want to swap out of them to do a long day shift instead. |
| 5 | n/a |
| 6 | Locum shifts are generated from the list of shifts unfilled we don't consider locum preferences when generating the rota. |
| 7 | N/A for this speciality |
| 8 | NA |
| 9 | No regular locums. generally very difficult to fill. |
| 10 | Locums tend to cover short term gaps, but if we have long term locums that we know well and are loyal to us then we adapt for them if we need to |
| 11 | we dont accommodate locums preferences. But we do employ locums as needed |

**Question 13: What would be the ideal output of an automated timetabling system for your use case?**
i.e. Would you like only the highest-quality timetable that was generated or would it be more useful to have multiple timetables of a similar quality to choose from?

Table B.15: Responses to Question 13

| Respondent | Response |
|---|---|
| 1 | Yes it would |
| 2 | multiple timetables of similar quality to choose from |
| 3 | One highest quality timetable |
| 4 | A single best quality timetable with a week to view at a time that can be printed off and displayed in the clinical areas |
| 5 | useful to have multiple timetables, including those for drs that cannot work nights but can do longer shifts, and those that work part time in different frequencies eg with 1 day off a week or 2 days off a week. The ability for departments to view them but not change them would be more efficient for all as well. |
| 6 | The system provides different options but I would prefer just the single best one to be honest. |
| 7 | Ease of use and understanding |
| 8 | few options of timetables with similar quality |
| 9 | high quality |
| 10 | At present i wouldn't trust an automated timetable as our area is so technical so I wouldn't be able to comment until I saw something |
| 11 | i dont think automated time tables work for any NHS organisation |

# Appendix C

# Interviews of Timetable Makers

Two interviews were conducted as part of this project; both participants volunteered after completing the survey. The first got in touch because their department had recently adopted an automated system, the second because they were willing to provide any additional information that would be required.

## C.1 Interview 1

**Question 1: What are the rules surrounding the minimum hours worked by a given doctor? How does this work for part time workers?**

*Full time rotas aim for an average of 40-47 hours per week. There are also targets for the number of day shifts (20 for an 8 week rota) and night shifts (7 for an 8 week rota) worked during the planning period. The targets for shifts are relatively flexible whereas the average hours worked per week should not go over the target. The average hours per week are adjusted according to the leave taken by the doctor in the planning period.*

*Part time workers are given targets that are a percentage of the standard targets for a full time contract - the percentage is determined on an individual basis. The number of part time workers in the NHS is increasing; some examples of departments and their proportion of part time workers include:*

- *Anaesthetics: 40-50%*
- *Emergency: 20 %*
- *Medicine: approaching 30%*
- *Surgery: less than 10%*

**Question 2: How many junior doctors are there usually per department and, as this will likely vary, could you give a range?**

*It varies greatly: you can have 43-44 in a department, while surgical departments could have as few as five, and certain departments like gynecology might have around 25. Rotas, however, are made for subsets of the department, common sizes for rotas are:*

- *8 people per rota*
- *10 people per rota*
- *12 people per rota*

**Question 3: What is the structure of shifts for an average day?**

*This will vary for every specialty, the structure for their department is as follows:*

- *08:00-17:30 (2 people)*
- *08:00-20:00 (1 person)*

- *19:30-08:30 (1 person)*

**Question 4: Do shifts have specific staffing requirements? How would this vary between departments**

*This will be different for each department; "grades" are not standardised between departments, each department could have its own way of grouping the doctors.*

**Question 5: Has the automated system successfully improved the efficiency of using salaried staff?**

*Yes, definitely: on average it has resulted in a reduced cost of £2,000 per rota*

**Question 6: How does the system handle understaffing, or how should my system deal with situations where locum shifts are necessary?**

*It would be best to spread out locum shifts where possible; partly to avoid a scenario where only temporary workers are present but also because it is easier to get coverage for a steady stream of locum shifts than for the same number of shifts in a shorter period.*

**Question 7: What is the input given to the system?**

*The shift routine of the department, any requested leave, and targets for day shifts and night shifts. This is important as the doctor's pay is partially dependent on the number of each type of shift worked.*

## C.2 Interview 2

**Question 1: How common is your use of locum doctors? Are there any specific considerations that should be made when allocating locum shifts?**

*Locum use is necessary every day, particularly at weekends - this is due to the 1 in 3 weekend stipulation of the junior doctor contract. It is best to spread out the shifts as this makes them easier to cover; the nightmare scenario is having only a new locum doctor working a shift - particularly for night shifts. It has gotten to the point where some locums are used on such a regular basis that their preferences are now taken into account.*

**Question 2: What kind of rota length would be acceptable for testing the automated system?**
(It was mentioned that it would be necessary to recruit volunteers for the comparative study)

*6-9 weeks would be fine; 8 weeks is relatively standard for a rota.*

**Question 3: How are doctor "grades" managed in your department?**
*This is different for every department, in their department they have two grades: "junior" and "senior".*

**Question 4: What is the standard shift structure for your department?**

- *08:00-16:00 (2 people)*
- *14:00-22:00 (1 person)*
- *17:00-01:00 (1 person)*

- *22:00-08:30 (2 people*

**Question 5: Are there specific choices that you made with regards to timetabling for your department that improved doctor well-being?**

*48 hours after night shifts rather than the stipulated 46; helps doctor recovery and seems to have improved well-being; did not impact ability to cover shifts. Also tries to give 48 hours after a weekend - often not possible.*

**Question 6: Do you attempt to account for doctor preferences at all?**

*Where possible, unfortunately not always doable - has resulted in department being requested by trainees stating preferences for rotations, the attempt to take individual preferences into account is cited as the reason for this.*

# Appendix D

# On Doctor Preferences and Fairness

As mentioned in section 2.2.1, 3 main types of preference were identified. However, the question was not asking for preferences in each of these areas, rather, it was a general question about what their preferences were. This means that doctors likely only mentioned the things that mattered to them most; had they been presented with each of the three areas there would likely have been more examples of the types of preference doctors can have for each area.

This led to a conundrum when generating instances: it was potentially unrealistic for a doctor to only have a preference in one of the categories, but there was not enough data to be able to assign a set of three preferences to each generated doctor; a probability for each collected example of a preference for each identified type could be calculated, but, due to the inherently individualistic nature of preferences, there would be no way to guarantee that each of the three preferences would be held by the same person.

As such, the decision was made to simply take the given preferences for each respondent as a set - the absence of a mention of one of the areas was taken to indicate a lack of a preference for that area. Each generated doctor was then assigned a random set of preferences identified from the survey. The hope was that this would provide a proof-of-concept for a hyper-heuristic system handling preferences without creating unrealistic instances. At the very least it would provide proof that the system is capable of adhering to the doctors' most passionately held preferences.

The process of adhering to the identified preferences, however, is relatively complex; there are numerous pitfalls: weight the preferences too weakly and the system won't adhere to them at all, too strongly and it will do so at the expense of the primary objectives, and without proper consideration a situation could be reached where preferences are granted to one doctor at the expense of another.

Identifying the correct weighting of the individual preference violations was simply a matter of trying a few values and observing the results produced, choosing the values that resulted in behaviour closest to that desired. What required more consideration was the minimisation of the disparity of preferences being granted. The difficulty was in choosing an approach that wouldn't inherently punish a doctor for holding preferences in more or fewer of the identified types.

Eventually, it was decided that the three types of preference would be considered separately. That is, the granting of a preference with regards to night shifts would have no impact on the granting of preferences for day shifts - the distribution of preferences being granted was only considered within the confines of each type of preference, not the number of preferences granted overall. This approach does have downsides, but it doesn't inherently negatively impact one group of people: if, for example, a doctor had no preference regarding the allocation of night shifts, it shouldn't concern them if people are having their preferences regarding night shifts adhered to.

As the problem of middle grade doctor shift scheduling has been modelled as a minimisation problem, it was decided that the best way to model preferences being granted was to try and minimise the number of violations for each preference. What was then required was a way to ensure that preferences were being adhered to evenly between each doctor, i.e. that there was not a great disparity in the number of violations for each doctor's preferences.

Jain's fairness Index (JFI) was identified as a way to measure this. Originally intended for use in networking, JFI (Equation D.1) aims to provide a measure of how evenly a resource is distributed, returning 1 for a completely fair distribution, and 0 for an entirely unfair distribution [1]. See section 2.3 for how the function is used within the objective function. It ensures that the violation of a given preference type is spread evenly between doctors with preferences of that type.

$$f(x_1, ...x_n) = \frac{(\sum_{i=1}^{n} x_i)^2}{n \sum_{i=1}^{n} x_i^2} \tag{D.1}$$

# References


[1] R. Jain, D. Chiu, and W. Hawe. *A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems*. Sept. 24, 1998. DOI: 10.48550/arXiv.cs/9809099. arXiv: cs/9809099. URL: http://arxiv.org/abs/cs/9809099 (visited on 04/08/2024).

# Appendix E

# Initial Solution Generation Pseudocode

**Algorithm 1** Pseudocode for the initial Solution Generation

```
doctorsNeedingNightShifts ← Doctors with a target greater than 0 for night shifts
doctorsNeedingDayShifts ← Doctors with a target greater than 0 for day shifts
sortedShifts ← List of shifts sorted in descending order for their number of infeasible doctors

for all s ∈ sortedShifts do
    a ← getRandomAssignment(s)
    feasibleDoctors ← getFeasibleDoctors(a)

    if isEmpty(feasibleDoctors) then
        continue
    else
        doctor ← getRandomDoctor(feasibleDoctors)
        allocateAssignment(doctor, a)

        if extraDayShiftsNeeded(doctor) = 0 then
            remove(doctor, doctorsNeedingDayShifts)
        else if extraNightShiftsNeeded(doctor) = 0 then
            remove(doctor, doctorsNeedingNightShifts)
        end if
    end if
end for

remainingDayShifts ← day shifts without an assignee
remainingNightShifts ← night shifts without an assignee

while isNotEmpty(remainingNightShifts) AND isNotEmpty(doctorsNeedingNightShifts) do
    s ← getRandomShift(remainingNightShifts)
    a ← getRandomAssignment(s)
    feasibleDoctors ← getFeasibleDoctors(a) INTERSECT doctorsNeedingNightShifts
    if isEmpty(feasibleDoctors) then
        remove(s, remainingNightShifts)
        continue
```

```
    else
        doctor ← getRandomDoctor(feasibleDoctors)
        allocateAssignment(doctor, , a)
        if extraNightShiftsNeeded(doctor) = 0 then
            remove(doctor, doctorsNeedingNightShifts)
        end if
        remove(s, remainingNightShifts)
    end if
end while

while isNotEmpty(remainingDayShifts) AND isNotEmpty(doctorsNeedingDayShifts) do
    s ← getRandomShift(remainingDayShifts)
    a ← getRandomAssignment(s)
    feasibleDoctors ← getFeasibleDoctors(a) INTERSECT doctorsNeedingDayShifts
    if isEmpty(feasibleDoctors) then
        remove(s, remainingDayShifts)
        continue
    else
        doctor ← getRandomDoctor(feasibleDoctors)
        allocateAssignment(doctor, a)
        if extraDayShiftsNeeded(doctor) = 0 then
            remove(doctor, doctorsNeedingDayShifts)
        end if
        remove(s, remainingDayShifts)
    end if
end while
```

# Appendix F

# Comparative Study User Guide

## F.1 Introduction

Firstly, thank you very much for agreeing to take part in this study, it is greatly appreciated. Before you get started, please read and sign the consent form included in the onedrive folder. If you do not do so, I will not be able to use your results in the study.

As outlined in the consent form, my dissertation project aims to investigate the applicability of hyper-heuristic methods to the problem of scheduling junior doctors; the aim being to reduce the time spent by staff, improve the utilisation of permanent staff, and to incorporate the preferences of the doctors (something that is not currently done). As part of an evaluation of the performance of the hyper-heuristic based system, I am conducting a study comparing the quality of timetables produced by the system to those produced by humans.

In order to mitigate the impact of your potential lack of experience in timetable creation, a graphical interface has been provided that limits assignments of doctors to feasible ones, and gives a view relating to data associated with each doctor (average hours worked per week, number of types of shifts worked, preferences, targets, etc.). The hope is that this will make it easier for you to create timetables; at the very least it will ensure that the timetables you produce are legal. I apologise for the threadbare nature of the interface, and its potentially unwieldly nature, this stems from the fact that it is just a repurposed version of the code that the AI algorithms interact with: my project doesn't involve UI design, and I have had to prioritise where I'm putting my efforts.

## F.2 Using the Interface

Shown in Figure F.1 is the overall view of the interface, yours may look slightly different depending on the department that you are making timetables for (the number of shifts per day varies). The larger box on the left is your view of the days within the timetable, the shifts, and the assignments that can be made. On the right is the view for information regarding the doctors. Both views are scrollable – feel free to give it a go.

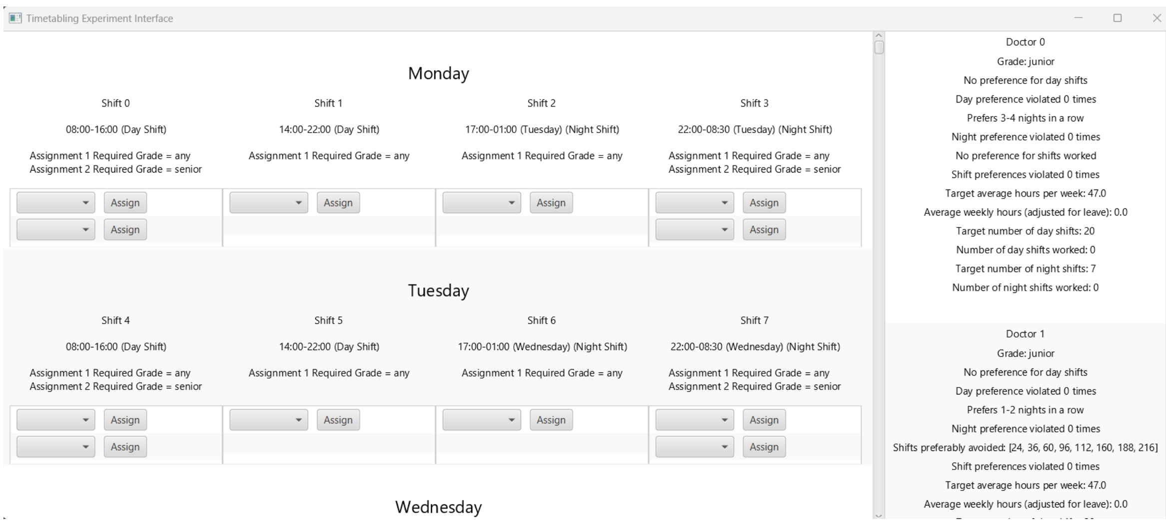



Figure F.1: The overall view of the interface

Figure F.2 shows the view for a single day. As you can see there is a row of shifts. In each column of this row, you get information about the shift (id, time, and the required grade of doctor for each assignment), and dropdown boxes for the assignments of that shift.

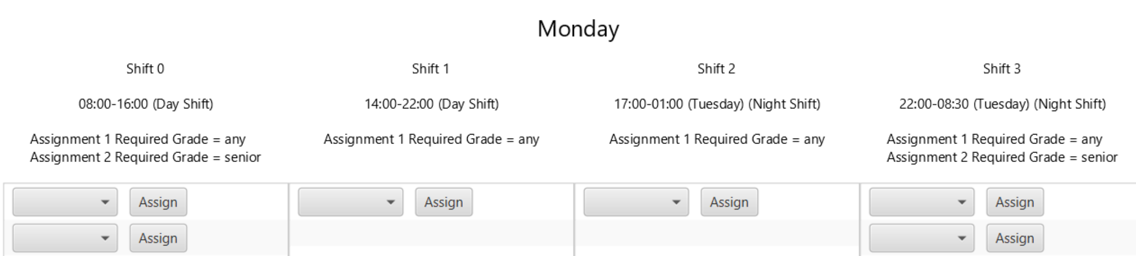
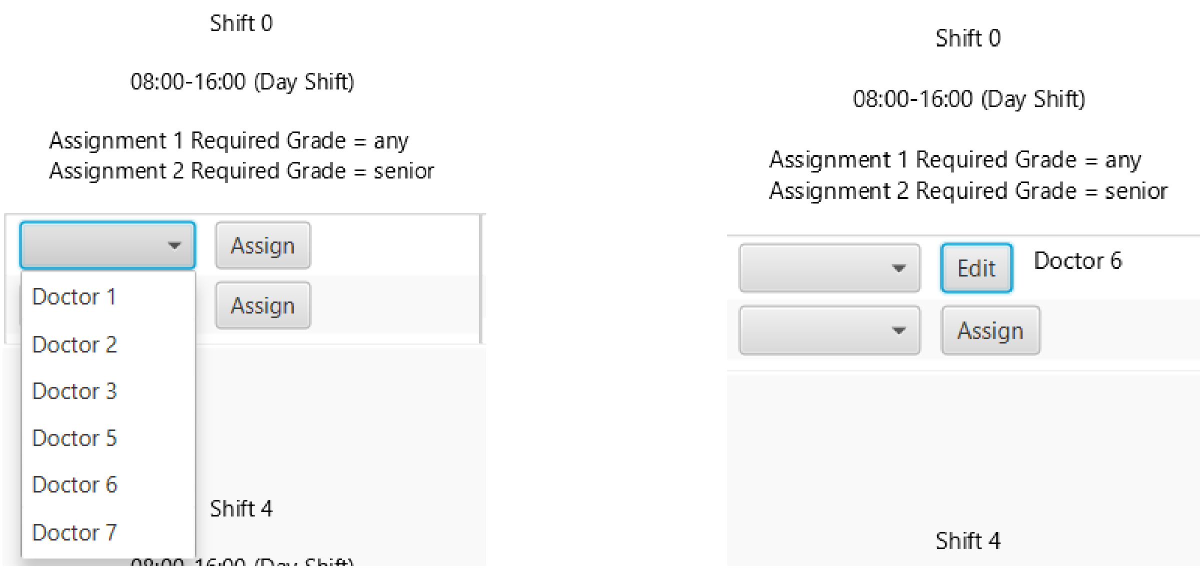



Figure F.2: The view for a single day

The "Grade" of a doctor refers to their level of experience, the number and classification of grades varies between hospital department. In the instances for this study we only consider two levels of experience: junior and senior. A shift will either have no requirements for experience, or will require one more experienced doctor. This requirement is implemented in the form of Assignments: Assignment 1 (top dropdown box) has no requirement for experience whereas Assignment 2 (second dropdown box) needs junior doctors of the "senior" grade.

There are two things to note here. The first is that the second assignment will be harder to allocate a doctor to, as you will have fewer doctors to select from. The second is that this does not mean that you can only assign one senior doctor to the shift: the requirement of "any" for the first assignment means that you could conceivably have two senior doctors assigned to it (although this might make it harder to cover shifts where a senior doctor is required).




(a) The dropdown menu of an assignment

(b) The view after allocating a doctor

Figure F.3: Sates of Allocation

Clicking on the dropdown box of an Assignment (Figure F.3a) will present you with the doctors available to be allocated. Select the doctor you wish to assign and click on the "Assign" button next to the drop-down menu. Notice that the information for the assigned doctor has been updated in the view on the right. The options for the other Assignments have also been updated, in line with the rules for creating legal rotas (Section F.4). In order to deallocate the Assignment, click the "Edit" Button (Figure F.3b). You can now select a different doctor, or even reassign the one you just removed from it.

If you scroll to the bottom of the Assignment section of the interface, you will come to the "Submit" button (Figure F.4. Pressing this button will save all of your assignments and either take you to the next problem instance, or quit the program. In this study you will create timetables for two problem instances. Do not worry about having to do both of the timetables in one sitting, if you wish to do them one at a time, simply submit the first one and close the program. When you are ready to create the second timetable, rerun the program and click submit without making any assignments. This will create another file, but it will be easy for me to tell which file contains your submission for the study. On this note, please ensure that you send me all of the files produced by the program, and please do not delete any of them.

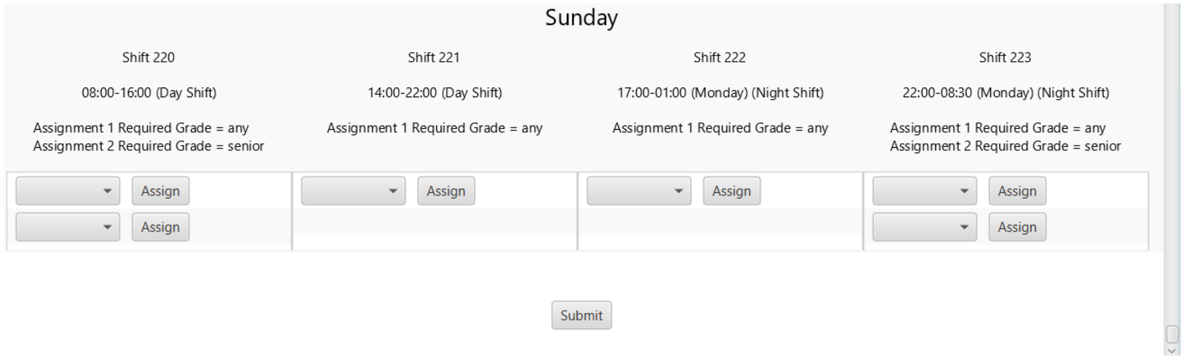



Figure F.4: The "Submit" button

## F.3 Objectives When Creating a Timetable

Now that you understand how to use the program, it is time to explain what you should be aiming for when creating a timetable. Your primary objective is to cover as many shifts as possible, things like preferences are secondary to this goal. It is also more important to ensure that each day has and each shift has at least one doctor working on it than it is to ensure full coverage. i.e. the selection in Figure F.5 is better than that in Figure F.6.

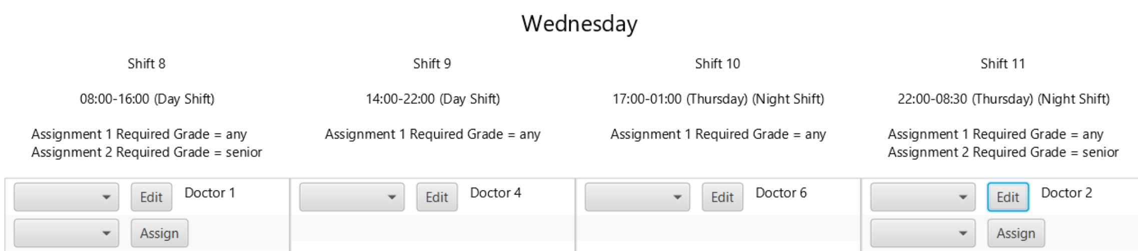



Figure F.5: A better spread of allocation

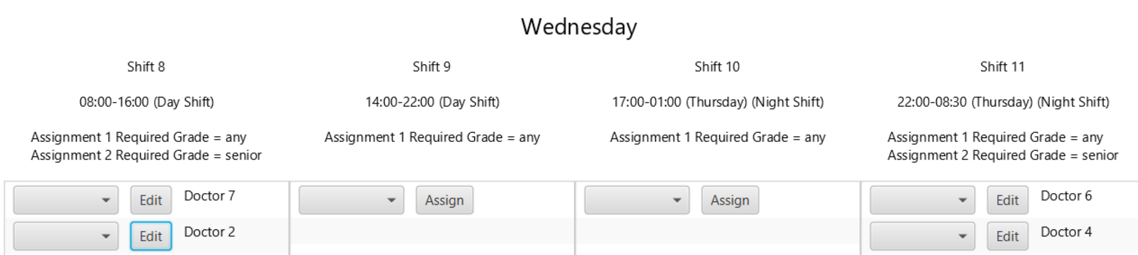



Figure F.6: A worse spread of allocation

The next most important factor is meeting the targets for hours worked and number of shift types worked by each doctor each week. These values vary according to whether a doctor is part time or full time. It is better to be under these values than over them (I was told that shifts are often deallocated to avoid going over the expected amount of work) but you should aim to be close; a doctor's pay is partially contingent on the number of night shifts worked, for instance, so these numbers do matter.

Finally, you might want to try and consider the doctor's preferences. This is considered a soft constraint (shift coverage is the priority) and it is quite hard to do, so don't worry if you aren't able to accommodate them. Preferences are split into three categories: those relating to days in a row, those relating to nights in a row, and those relating to shifts that a doctor would prefer not be assigned to. Not all doctors have preferences in these areas, and the doctor view of the interface will let you know whether they do.

What is very important, however, is to avoid ignoring one doctor's preferences while granting those of another. You can keep track of this by looking at the number of violations for a doctor's preferences and trying to maintain a similar number for all doctors with that category of preference.

To summarise, your aims are to (in order of priority):

1. Minimise the number of days without any coverage
2. Minimise the number of shifts without any coverage
3. Minimise the difference between target hours and shifts worked for each doctor and the numbers actually worked, without going over the target
4. Minimise the number of unassigned assignments
5. Minimise disparity between doctor preferences being granted
6. Minimise violations of doctor preferences

Please do not be put off if you find it difficult to meet these objectives, this is a problem that is very difficult for humans to find solutions for: the primary aim of this study is to show that a hyper-heuristic approach is capable of finding better solutions than could otherwise be obtained.

## F.4 Rules Regarding Rest

The interface provided enforces adherence to the rules necessary for creating a legal timetable, however, it would still be beneficial to understand what makes a shift feasible or infeasible.

There are three basic causes for a shift to be infeasible: rest, training, and leave. All three causes are taken into account by the interface, and rest is the only cause of infeasibility that can be impacted by your actions; a shift that coincides with a doctor's leave or training cannot be assigned to them, regardless of what is timetabled before and after it.

The rules regarding rest are as follows:

- Doctors must be given at least 11 hours rest between shifts (starting at the end of the worked shift)
- Doctors must be given either 24 hours of rest once a week, or a 48 hour stretch within a period of two weeks
- Doctors can work no more than 7 days in a row, and are given a minimum of 48 hours of rest after such a stretch
- Doctors can work no more 4 night shifts in a row
- Doctors must be given 48 hours of rest after the end of a night shift or stretch of night shifts
- If a doctor works on a weekend (Midnight on Friday – Midnight on Sunday), they must be given two weekends off (1:2 ratio)
- A doctor can work no more than 4 long shifts (longer than 10 hours) in a row, 48 hours of rest must be given after the 4th long shift

**Notes**:

**1.** It is often not possible to cover all the weekend shifts

According to NHS workers that I have had contact with, ever since the weekend stipulation has been added to the junior doctor contract, it has been exceedingly difficult to cover weekends and external workers often have to be hired. You therefore should not troubled if you find that this is the case for your timetables.

**2.** On the allocation of night shifts

You can have rows of night shifts that are 1-4 nights long; note that the requirement of 48 hours of rest only applies to the last night of the row in question. It also means that if your row is less than 4 nights long, the next night after the row can be assigned, but the night after it cannot. Figure F.7 tries to illustrate this concept.

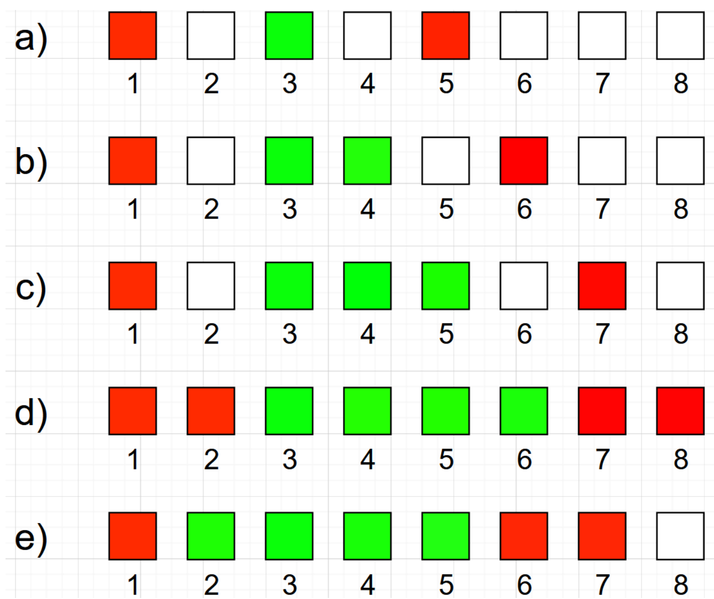



Figure F.7: A few scenarios that might arise when allocating nights

Each square illustrates a night; a red square indicates an infeasible night, a green square an allocated one, and a white square indicates an unassigned night that could feasibly be worked. a) shows the scenario after after night 3 has been allocated. At this stage the subsequent night shifts are not worked, so it has 48 hours of rest. If you were to allocate night 4, you would create a row of two nights with 4 at the end: the rule for 48 hours rest would now only apply to shifts after night 4. As the night shifts 48 hours after it are not worked it can feasibly be assigned. Night 5, however, would not count as part of the same row as night 3 (as night 4 has not been assigned) and would thereby violate the requirement of 48 hours of rest after night 3; hence, it is made infeasible. Night 1 is infeasible as night 3 prevents it having the requisite 48 hours of rest. Note that night 2 is still feasible: were it to be assigned you would have a row of two nights, 2 and 3, and the 48 hours of rest would still start at the end of night 3.

b) shows what happens if night 4 is assigned after night 3. Notice that night 5 has now been made feasible. This is because we are now only considering the 48 hours after night 4: the row is less than 4 nights long, and there would be 48 hours of rest after night 5; it can therefore now be assigned as it would join the same row as nights 3 and 4. The allocation of night 5 is shown in c).

d) and e) show the two possible allocations that can be made to add to the row of nights after step c). As you cannot legally have a doctor work for more than four nights in a row, the nights before and after the resultant row also become infeasible.

It should also be noted that night shifts can overlap into the next day, for example a shift might run from 8pm on Friday to 8am of Saturday. The doctor would then have to be counted as working on both Friday and Saturday - this is just something to keep in mind as you cannot legally allocate more than 7 days in a row.

## F.5 Taking Part in the Study

In order to take part in the study, please ensure that you have downloaded all of the files from the OneDrive folder. If you have done so, your folder should look like that in Figure F.8. In order to run the interface, simply double click on ”experimentInterface.exe”.

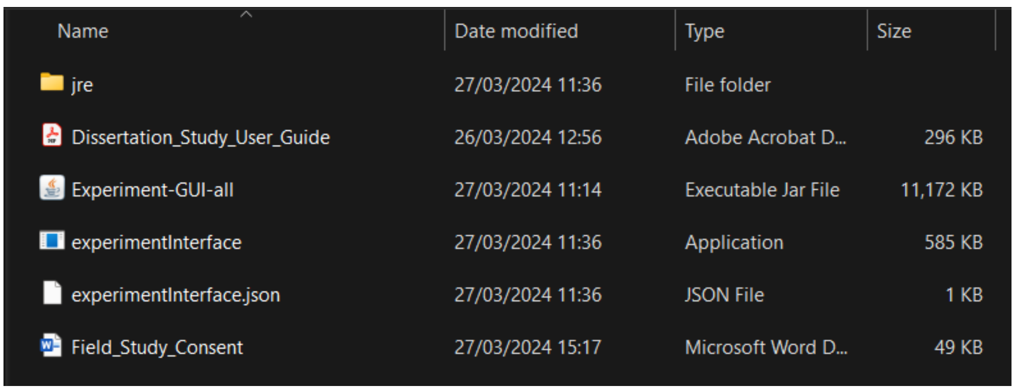



Figure F.8: How your downloaded folder should look

For my dissertation I have created models of two departments in the hospital, and have written functions capable of generating instances for which timetables can be created. An instance contains information like the number of shifts to be covered, the number of doctors, their preferences, training schedules, leave, etc. For the experiment, you will be given a two instances, one from each department. One will be easier: it will have the standard amount of doctors with a minimal number of them working part time or on leave during the planning period. Whereas the other will be harder, providing you with fewer doctors, more of which will be on leave or will work part time.

These instances are hard coded into the interface, when you run it, the easier interface will be loaded first. Upon submitting your solution for the first instance, you will then be shown the second instance. As mentioned before, you do not have to do both instances in one sitting. If you want to do them one at a time, simply rerun the program and click “Submit” on the first instance without assigning doctors. Upon submitting your solution for the second instance, the program will close itself. At this point, please email all of the produced files and your signed consent form to psydf3@nottingham.ac.uk. The files will be found in the same folder as the interface (Figure F.8) and will be text files (.txt).

Thank you again for agreeing to take part in this study. If you have any questions about the study or problems with the software, please email me and I’ll try to iron out any issues as quickly as possible.

# Appendix G

# Workplans

## G.1 Initial Plan

**The breakdown of tasks is as follows follows**:

(A) Enquire about the feasibility of obtaining data and send out surveys

(B) Outline initial definition of the problem using available information

(C) Encode representation of the problem and a basic objective function

(D) Define a format for problem instances and encode an instance reader

(E) Encode low level heuristics from literature

(F) Create process for initial solution generation

(G) Outline final formal definition of the problem

(H) Produce mechanism for creating problem instances

(I) Update objective function and low level heuristics

(J) Analyse the performance of existing hyper-heuristics on the problem

(K) Write the interim report

(L) Write the final dissertation

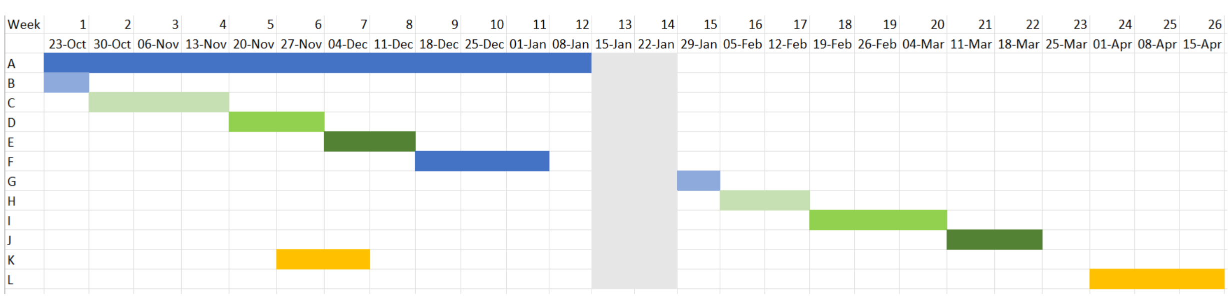



Figure G.1: A Gantt chart showing the project timeline

## G.2 Interim Report Plan

**The breakdown of tasks is as follows follows**:

(A) Enquire about the feasibility of obtaining data and send out surveys

(B) Outline initial definition of the problem using available information

(C) Encode representation of the problem and a basic objective function

(D) Define a format for problem instances and encode an instance reader

(E) Encode low level heuristics from literature

(F) Create process for initial solution generation

(G) Outline final formal definition of the problem

(H) Produce mechanism for creating problem instances

(I) Update objective function and low level heuristics

(J) Analyse the performance of existing hyper-heuristics on the problem

(K) Write the interim report

(L) Write the final dissertation

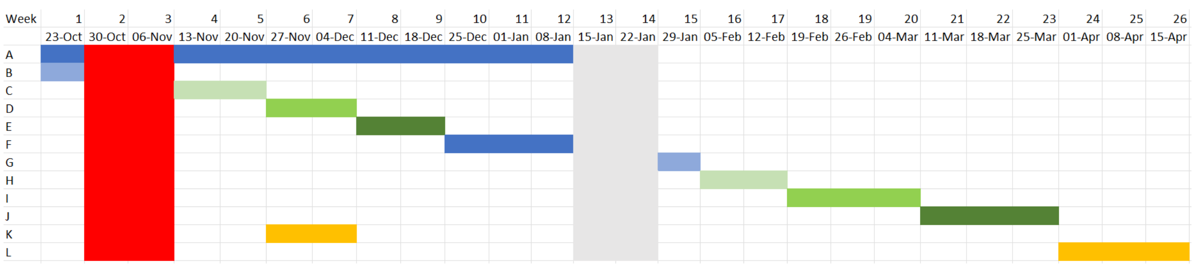



Figure G.2: The updated gantt chart for the project