University Timetabling: Bridging the Gap between Research and Practice

Barry McCollum

School of Computer Science, Queen’s University
University Road, Belfast BT7 1NN, N. Ireland
b.mccollum@qub.ac.uk

eventMAP Ltd
21 Stranmillis Road, Belfast BT9 5AF, N. Ireland

Abstract. As an academic in the School of Computer Science at Queen’s University, a visiting researcher to the Automated Scheduling, Optimisation and Planning (ASAP) group within the School of Computer Science and IT at the University of Nottingham and Managing Director of eventMAP Limited, a university technological spin out company, the author is in a unique position to provide comments on both the practice and theory of timetabling (automated and otherwise) within the university sector. The study of the relationship and interaction between the work carried out in the academic literature and the requirements of university administrators is essential if ideas generated by research are to benefit everyday users. Conversely, it is crucial the needs of the timetabling community influence the direction taken by research if high quality practical solutions are to be produced. A main objective of the work presented here is to provide up-to-date information which will enable researchers to further investigate the area of timetabling research in relation to the generation of robust and flexible techniques which can cope with complexities experienced during implementation in ‘real world’ scenarios. Furthermore, although not discussed here in detail, it is essential, from a commercial perspective, that these developed leading edge techniques are incorporated and used within general applicable timetabling tools. The aim of this paper is to motivate the discussion required to bridge this timetabling gap by bringing timetabling research and educational requirements closer together.

1 Introduction and Context

EventMAP Limited was formed in 2002 to exploit the commercial potential of the educational timetabling research carried out by the Automated Scheduling, Optimisation and Planning (ASAP) group at the University of Nottingham and the Knowledge and Data Engineering (KDE) Group within the School of Computer Science at the Queen’s University of Belfast. The Company is based in Belfast within the Institute of Electronics, Communications and Information Technology (ECIT) at Queen’s University. The Institute, which officially opened in May 2005, represents a new £40M world class centre with a unique focus on blue skies, strategic and
industrial research projects. The Centre brings together internationally renowned research groups specializing in key areas of advanced IT, digital and communications technology. A key feature of the Centre’s overall remit is the “spinning out” of industrial based companies exploiting advancements made in research.

The decision to form a company followed identification of the market need for a high quality research led software solution to the scheduling difficulties experienced within the educational sector. The focus of eventMAP Limited is to develop, market and sell examination, course scheduling and space management and planning software into the worldwide higher and further education sector. The preface to the Selected Papers Volume from the Gent PATAT conference [6] stated that “The goal of developing interactive and adaptive systems that build on human expertise and at the same time provide the computational power to reach high-quality solutions continues to be one of the key challenges that currently faces the timetabling research community” This goal is very much shared by eventMAP Limited whose approach is to incorporate knowledge of the extreme complexity of timetabling problems with commercial skills and practical experience with the overall aim of developing and building upon the most recent research in Artificial Intelligence and Operational Research technologies.

The Company aims to develop and implement new practical methodologies and associated algorithmic techniques to enhance the solution of educational timetabling problems across a wide range of real world scenarios. At this early stage of the company’s existence, consultancy has been provided and systems implemented in Europe, Australia, New Zealand and America. The fact that work has taken place on a global scale at such an early stage in the company’s history is both promising and challenging from a company growth point of view.

In the recent international review of Operational Research in the UK (commissioned by the Engineering and Physical Sciences Research Council), a major identified weakness in the current approach to Operational Research is described as follows, “a gap still remains between the output of a successful research project and what is needed for direct use by industry” [1]. In general, the area of educational timetabling is one such area. The Company has an important role to play with respect to this ‘gap’ as it is in a unique position to integrate leading edge research techniques with the requirements of the user base in the provision of timetabling solutions. One of the primary overall aims of current efforts within the Company is to implement software which acts as an enterprise recourse planning tool as well as a management information service, informing on strategic ways forward for the need for, use of and allocation of resources within an institution. A major aspect of the adopted strategy for achieving this is to highlight the important aspects of institutional requirements to researchers in the field while continually updating algorithmic techniques within the software, thus enabling solutions to be produced which are both workable and of a high quality. The intention of this paper is to focus on the initial part of the strategy by reporting on the needs of educational institutions from a practical point of view in terms of two of the main areas which the company is involved with i.e. examination and course timetabling. In each area, a number of challenges are detailed which are based on experience of working in the area from both an academic and practical view point. It is stressed that these challenges certainly do not represent all of the issues that require work from researchers, rather they represent a selection of key themes...
which will help bridge the gap and move the area of educational timetabling to a new level both in research and practical terms.

2 Examination Timetabling

The examination timetabling problem, studied in numerous papers in the PATAT conference series [2,4,5,6,7], is characterized by a set of students taking a set of exams over a specified time period within the context of various constraints. The quality of the timetable is normally measured as a function of best spread of examinations per student though some notable exceptions do occur [8,9]. Various algorithms have been used with their effectiveness being measured in relation to a standard set of benchmark data. An up-to-date review is provided in [10]. In addition to the PATAT Conference series, many papers have been published on specific techniques along with reporting of various surveys [11,12]. It is worth noting that research in this area has been instrumental in the continued development of the field of search methodologies and, in particular, metaheuristics. Although it is not intended to provide a general commentary on the approaches adopted to date it is possible to argue that the nature of the gap between research and practice has not been aided by the simplicity of the current datasets e.g. the lack of substantial benchmark data with sufficient room, constraint and solution modelling data. It is expected that the release of six new datasets [13] along with a dedicated web service to the research community via the web site at http://www.cs.nott.ac.uk/~rxq/data.htm will go a long way to remedying this situation. This service will also act as a repository of information relating to techniques and solutions generated and will enable researchers to easily and accurately test and compare approaches.

From a Company perspective, the latest version of its flagship examination product, OptimeXam, was released in January of this year. An earlier version of the software was presented at the PATAT conference in Konstanz, 2000 [2]. The additional functionality made available through this new version will be discussed at the conference during a software presentation [14]. In general, the aim of improving OptimeXam is to make the system as intelligent and intuitive as possible, providing maximum information to the institutional administrator, allowing informed strategic and managerial decisions to be made. This has been achieved through the inclusion of the user in all stages of the ‘examination modelling’ process. It is important to note that although not described in detail here, the ‘gap’ between the needs of the user and the provision of software is also being tackled within the company by the development of a close working relationship with users. Feedback from this process which is relevant to researchers includes modelling aspects of the information, algorithmic and solution development, all of which represent significant challenges for the research community. The following discussion is concentrated around this reported examination modelling process.

2.1 Building the Institutional Model

The development of examination timetables within institutions is a multi phase procedure that is dependent on varying criteria at each stage. Firstly, a structure has to
be decided on before exams and students are assigned e.g. the length and format of the time period together with the ‘diet’ of rooms which are to be made available. Secondly, data on exams and associated constraints have to be added before the student information is considered. The stage and degree of automation is highly dependent on the procedures adopted within the institution. This multi-stage process is referred to here as building the ‘institutional model’. This process encompasses two main aspects i.e. information and solution modelling.

2.1.1 Information Modelling

Information modelling can be divided into data and constraint modelling. The base examination data from which a workable solution is achieved is composed of student enrolment, exam and space data. In addition, the construction of an overall solution is phased due to the information environment within which the examination process takes place. In practice, a solution is often attained based on a percentage of the actual data due to incomplete and inaccurate data from the student administration systems. Ultimately the algorithms applied must therefore construct solutions working with a degree of uncertainty. The inadequacies of the data set up therefore represent the first challenge to the timetabling community. It is suggested that there are two possible approaches to solving this problem i.e. either solutions are sought with associated repair mechanisms or robust optimisation techniques are used which produce solutions that are ‘good’ for an agreed range of input values. Under this scenario, a solution would be sought that remains feasible for all potential input data values. Although some work is evident in the literature in relation to the first of these approach in relation to educational timetabling [15, 16], little attention has been paid to the second.

Constraint modelling involves setting up a range of criteria which effectively describes the boundaries within which a solution should be constructed. Constraints used in institutions have been reported in 1996 [17]. Since then, in the UK in particular, there has been a steady increase in complexity regarding this issue with the implementation of increasingly flexible modular course structures by many universities. The central production and coordination of the associated examination timetable has become increasingly difficult with more examination offerings having to be timetabled in such a manner so as to offer students maximum spread throughout the session while ensuring space usage is maximised. In addition, many new constraints have been added to the overall problem to accommodate all types of special needs of students. An example of this was reported in the Times Higher in March of 2006 where students from a Muslim background require Fridays free of examinations [18]. This and other additional soft constraints further complicate the modelling process and the scope of potential solutions. It is essential these are documented and incorporated into the modeling process as, for example, at our leading implementation site, 9% of students in the 2004/05 academic year had special needs with regards to their examination requirements. The second challenge is therefore to redefine the problem in terms of recent identified changes. This can be achieved by getting access and reporting on practical examples of constraints and the processes involved. The PATAT conference series and the close link with eventMAP limited is of particular relevance here as practical issues as well as datasets can be added to the research base on a continual basis. Another important aspect of
constraint modelling is the structure of the examination session i.e. session modelling. Two features of this are detailed below.

In establishing an institutional model for the examination process, one of the major issues for many institutions is the potential relaxing of a constraint which has hitherto been considered ‘hard’ i.e. the imposing of certain time periods within the day structure. For example, a day may be split into two periods of three hours in length, one beginning at 9am and the other beginning at 2pm. Analysis of various solutions produced by eventMAP has shown that this is the single biggest factor in relation to poor usage of time and space and hence a major contributory factor to poor overall solutions. This is chosen here as it is an excellent example of a hard constraint which needs to be changes to move the examination timetabling forward from a practical point of view. Before leaving the established ‘period based’ approach to one side, it is essential to understand the required needs and the extent of ‘non period’ based timetabling. The period based nature of the problem needs to be investigated to establish a model where examinations can be scheduled during any part of the defined day. This issue is related to recent work with respect to a redefinition of the nurse scheduling problem [49] where metaheuristic techniques which have been used to manage this time interval coverage have produced the best results so far on the presented data. Due to the similarity of the nurse rostering and examination timetabling problems it is considered appropriate that these techniques are investigated. The concept of ‘time interval’ was introduced, where instead of formulating the staff requirements as the number of personnel needed per shift type for each day of the planning period, time interval requirements allowed for the representation of the personnel requirements per day in terms of start and end times of personnel attendance. As with the nurse scheduling example, an updated formulation would enable the provision of a greater number of time slots and would reduce the amount of unproductive time currently in existence.

It is clear that institutions involved in the process of carrying out the initial stage of the institutional modelling process often do so blindly. That is to say, they base the timetable on new data but attempt to superimpose this on existing models of how the examination sessions should progress. For example, an existing model for a particular institution may be a certain number of periods over a designated time period with a certain number of rooms. This, in part at least, is related to inadequate methods which allow users to understand how solutions are being created. For example, space considerations are often an afterthought with the primary aim being the actual creation of a timetable. No help is afforded to the users in directing them towards a solution which is ‘right’ for the Institution. Before going on to the important issue of solution modelling in the next session it is important to note that the investigation of similarity of data to previous datasets from the same or indeed other institutions is important if efficient and effective models are to be found. Continuing on from recent work [21,22] on similarity measurements between datasets, novel techniques need to be investigated to establish how changes in individual data sets from year to year effect the nature of the examination set up and ultimately the algorithmic methods applied.
2.1.2 Solution Modelling

Solution modelling is concerned with the construction of a solution in terms of what is deemed important to the institution. Currently, the majority of the work in evaluating a solution is based on the production of a single solution from each execution of the algorithm whose value is measured by a single objective weighted sum of soft constraints. There are some exceptions though, for example, in paper [9], the quality of a constructed timetable is considered in terms of the average penalty per student and the highest penalty imposed on any one student. Although research has been carried out in modelling the problem as a multi-criteria/objective problem [54, 55] this work has not yet been implemented into a generalised tool. The responsibility is currently on the user to model the problem accurately at the constraint modelling phase and subsequently ‘leave’ it to the algorithm to produce the ‘best solution’. This has the effect of the user feeling ‘frozen’ out of the solution construction phase and gives the impression that this is the best solution based on the constraint set up process. Of course, this is not the case with many solutions being possible which ‘best’ fit the constraints set up. Paquete et al [19] carried out work in which individual constraints were given preference at various stages of the process. This is similar to how the process of solution construction is carried out in a number of institutions with, for example, the effectiveness of a solution being measured as the ‘number of students with two examinations in a day’. It is clear that the user requires a number of solutions to be presented with the differences explained intuitively, thus allowing the user to decide on what solution is the ‘best’ to meet the institutions needs. It is suggested here that this could be achieved by a combination of techniques incorporating pareto optimization and fuzzy techniques e.g. the user chooses the characteristics of the solutions they would like to see from a number of fuzzy sets. This could possibly be translated into a choice function for discriminating between the non dominated pareto solutions generated by a multi objective algorithmic technique. It is stressed that this is only one possible approach which could be used to address this important issue. More work is required on how the quality of solutions are measured. The challenge for researchers is the provision of a solution where the user understands the trade offs between the original objectives.

Once a solution is being generated, it is normal to have a construction phase followed by an improvement phase. In both cases there have been many heuristic techniques applied (see [11]). Recent work has shown promise in relation to using a combination of heuristics in relation to the initial construction [20]. Results on the benchmark datasets have got increasingly better over the years as more and more metaheuristic techniques have been applied and domain specific knowledge has been increasingly incorporated into the approaches [10, 11]. One criticism of this approach is that the developed techniques have become specialised in relation to the benchmark datasets at the possible cost of generality i.e. techniques which can produce ‘good’ results when applied across a wide range of other real world scenarios. Recently, in terms of metaheuristics, it has been shown that changing the neighborhood structure has been effective. It is felt that Hyperheuristics approach (heuristics to choose heuristics) [56] undoubtedly offers promise as this methodology is based on raising the level of generality by aiming to automatically apply the correct heuristic or metaheuristic at the correct stage of the problem be that in the construction or indeed the improvement phase. Currently, Optime enables the timetabling algorithm to be
varied depending on the user algorithmic modelling process. These observations are the result of a close working relationship with five principal users in the UK and they currently represent the basis of further research [13]. Currently the combinations of algorithmic structures available are Saturation degree (Heuristic Method) [25], Adaptive [26] and Great Deluge during an additional improvement cycle [27]. The algorithm set up thus enables the user to have control over the time spent on various aspects of its operation. This is a first step in involving the user at a higher level of the algorithmic modelling of the problem and is in response to the observation that various algorithmic set ups perform better on different datasets. It is important to understand why various metaheuristic and combination of metaheuristics work better in particular situations. One challenge to the research community is therefore to explore how new search methodologies can underpin the development of more widely applicable timetabling systems. Indeed this is one of the main motivating factors for the current level of interest in hyperheuristic research [74].

3 Course timetabling

The University course scheduling problem is concerned with groups or classes of students following a particular defined pathway or course which has associated events that require the allocation of time and resources. Recent definitions of the course timetabling problem can be found in [12,29]. As with the university examination problem, a solution requires a number of hard and soft constraints to be satisfied. Similarly, the central production and coordination of the course timetable is essential as more modules and associated events have to be timetabled in such a manner as to, firstly, offer students maximum flexibility of choice, secondly, to provide flexibility for staff and, thirdly, to ensure that teaching space is used effectively. Universities, struggling with rising student numbers, have increasingly relied upon the automation of this task to produce efficient timetables which satisfy these constraints [11]. Much of the software assistance that is currently available is either a commercial product or has been designed specifically for the institution in which it was developed [30,31,32]. In both cases the timetabling process often involves significant human interaction which, in practice, can turn the process into a room booking exercise [33,34]. Therefore, the construction of a solution is often categorised by finding any timetable that satisfies all of the constraints [12]. From a software point of view, any solution is often seen as a good solution and, indeed, the notion of an ‘optimised solution’ is usually not a main objective of incumbent university administrators. The reasons for this are diverse and complicated. One issue is that as too much assumed and incomplete knowledge surround the entire process and their exists many staff, with differing view points involved. The data required for the process is often difficult to obtain and, as with the examination process, it is often ‘sketchy’ [45,64]. From a staff point of view, fixed views exist on when and where teaching should take place within a predominantly ‘territorialism’ culture [34]. These issues will be further explored in the remainder of the paper with challenges presented as to how this area can be moved forward from a research point of view. It is important to note that, within the majority of universities which use automated systems, the process of the
production of a workable timetable remains firmly with a combination of lecturing and administrative staff rather than the sole use of the automated component. Recent years have seen significant research efforts to improve this situation. The following papers represent a small selection of these contributions [16,29,31,33,34,35,41,42,45]. Carter [42] stressed the importance of taking into consideration and dealing with the human factors associated with the process of constructing an institutional wide timetable. However, when dealing with the issue of course timetabling, it is often the case that many of the papers ignore the human factors all together, choosing to deal with ‘sculpted’ data sets in order to evaluate particular techniques and approaches. Some real world aspects have been discussed in the literature but these tend to be in conference abstracts (as a small selection, see [40,63,64,66,67]) rather than full papers. If one of the strategic goals of timetabling research over the next few years is to close the gap between theory and practice then these issues have to gain more prominence in the mainstream literature.

Although many advancements have been made with respect to the development of search techniques on benchmark data sets [29,36,41,57,58], there is not much evidence that the work has been translated into actual implementations within a significant number of institutions. Indeed Carter and Laporte [31] comment that they were “somewhat surprised to discover that there are very few course timetabling papers that actually report that the (research) methods have been implemented and used in an institution”. Although this was reported almost a decade ago, the situation largely remains unchanged. They go on to say that they expected to see a number of implementations in the near future. Once again unfortunately this has largely not been the case.

In relation to this area in general, it is suggested here that, there has been insufficient investigation of real world issues and therefore understanding of the methodologies used by expert timetablers. More work needs to be carried out on the formulation and modeling of the problem. This latter issue is particularly challenging because different institutions must satisfy a range of different constraints in generating an institution-wide timetable [35, 31] which means that a generally applicable solution to this complex problem is extremely difficult. Given the complexities of real world course scheduling, many researchers have developed approaches which rely on various simplifying assumptions in modelling the problem. While it can be argued that this is valid as an initial research test bed, which has resulted in useful and powerful search techniques, such an approach needs to be supplemented by methods which addresses the true complexities of the problem that must appear in real world applications. By way of illustrating this point, recent work carried out on practical course timetabling by the Metaheuristic network [36] used generated datasets. It was stated that

"The problem we are studying in the Metaheuristics project is one that is closely based on real world problems, but simplified. We are not entirely happy about using a simplified problem, but the reasons are two-fold: We want to be able to see more clearly what is going on in algorithms designed to solve the problem. Real data is too complicated, and real problems have too many soft and hard constraints to allow researchers to properly study the processes and; The large number of soft and hard constraints in real data (and the differences between them at different institutions)
make it a long process for researchers to write code to solve them, or to adapt existing programs to be suitable.”

Although this has been useful, from a practical point of view, the results obtained do not seem relevant in practice. In addition, the impression is often that benchmark course timetabling datasets \([36,57]\) are seen as data which can be used in addition to examination data sets to prove that certain search techniques are of benefit. Although successful in this regard the gap between research techniques and the software required for actual implementations is much wider than that seen with examination timetabling. Whereas this paper has spent the opening sections detailing challenges which will help narrow the gap in relation to examination timetabling, the rest of the paper will concentrate on describing course scheduling from a practical point of view with the hope of identifying what is required if a relevant and comprehensive formulation of the problem is to be reached. It is felt that this view of the course timetabling problem will better serve the purpose of making timetabling research more relevant to real world practice. It is stressed that the contribution of timetabling research must address more wide ranging issues than the tuning of algorithms to work well on particular datasets. Rather, the modelling issues related to the complexity of real world implementations must be recognized and dealt with. The most realistic formulation of the problem which currently exists can be found at \([24]\). Further work is required to build on this to allow the full complexities of the problem to be explored and to narrow the current gap. With this aim in mind, it is essential that more comprehensive representative benchmark datasets are made available along with information on the aims of the associated institution.

3.1 A Very Different Timetabling Problem

University course timetabling is often reported in the literature as a variance of the related examination timetabling problem \([12]\). Indeed it is the author’s impression that many pieces of research default to talking about examination timetabling when they are talking about university timetabling in general. Although some of these issues are further described in subsequent sections of the paper it was felt worthwhile to draw out the major differences between the two types of timetabling at this early stage in the discussion. The reported difference is often the addition or removal of particular constraints e.g. more than one event cannot take place in the same room and lectures should be avoided in the last period of the day \([41]\). In addition, the term ‘best spread’ of events has an entirely different meaning.

A major difference with the examination timetabling process is the environment in which the construction process is carried out. This is a dynamic, multi-user distributed environment with various cohorts of schools and departments who often operate quite autonomously. Although issues in relation to this have been studied, for example \([64,69,70,71]\), much more work is required on understanding the issues involved and the interplay between user interaction and managing the information with the goal of producing a workable solution and the extent to which techniques can be used in an automated process. These issues will be discussed further at various places under the heading of ‘building the institutional model’.
Another difference that is often overlooked is, as with the examination problem, course timetabling does not take place at the module or course level. The following presents a discussion on the effects of this. Consider the module ‘Introduction to Computer Science’ with associated module number 110CSC101. The associated examination for the module will normally take place at the end of the semester in which the module is given and will be timetabled by the rules employed by the institutional examination officer which are generally those governing the body of research which has taken place over the last decade or so. Therefore, in this case the ‘gap’ which exists between what is required by the institution and the techniques researched from an academic sense, is small. The course timetabling issues with the module 110CSC101 are more complicated. The module can be broken into a series of events which require timetabling e.g. lectures, seminars, tutorials, practical classes and laboratory classes. A subset or indeed all of these ‘event types’ require timetabling in a manner which provide the group of students associated with the module, firstly, a feasible solution and secondly, a ‘good’ timetable. A feasible solution is achieved by ensuring that individual students can attend all event types associated with each of the modules that constitute the overall pathway they are enrolled on e.g. year one of BSc in Computer Science. Secondly a ‘good’ solution is one which satisfies the soft constraints as defined by the institution e.g. Lectures should be in the morning in a particular time or room. It is clear that these soft constraints require a higher investigation as they can vary from one institution to another and indeed from one event type to another belonging to the same module. Furthermore, in setting up the problem, these events have different individual requirements, ordering and constraints. The following section outlines some of the associated issues.

The simplest example is that particular event types are usually associated with certain types of space e.g. a computer laboratory class must take place in a computer laboratory. Also, lecture events represent the entire group of students on the module whereas the other event types represent subgroups as students are divided into smaller groups for different types of study. This issue of event subdivision is further explored in the following section. From an ordering perspective, it is often the case that particular orders of events over a defined time period e.g. a week, are defined to achieve the desired combination of teaching and learning skills. It is also often the case that particular events are related to each other in relation to the time which separates them in this ordering e.g. seminar classes should be timetabled in the afternoon following the lecture activity. In addition there is an associated hierarchy with the event types e.g. lectures are timetabled as a priority in the first instance to ensure that the entire group can be brought together. It is often the case that this situation means that lectures will be timetabled first with all other events timetabled after week one of the semester. Of course, there are many variations of this related to when the timetable is produced in relation to student enrolment i.e. pre enrolment or indeed post enrolment. Event types may also have a particular life span associated with them throughout the semester. Whereas the lecture event may run in a particular format throughout the entire semester, other event types may begin and end in particular weeks. In addition they may have an associated pattern which is individual to the event type e.g. lectures may run twice a week for 12 weeks whereas lab classes may begin in week three and run for a three hour afternoon slot every two weeks for
six weeks. Currently, research does not take these considerations when either defining the problem or applying techniques to help solve the problem. This has been detrimental to the overall practical area and has meant researchers, in many cases, have been working on oversimplified problems.

Course scheduling, much more the examination timetabling, must be seen in the wider context of the use and availability of institutional space either existing or in the planning stage. This linkage allows measured and improved utilisation while identifying the needs for particular types of space across the Institution. The Company aims to model how increases in course delivery, through effective timetabling, can affect the overall nature and structure of the campus. Ultimately, this would allow for strategic decisions to be taken in relation to room types, sizes and quantities across all space types within the Institution. The course timetabling system is therefore a fundamental part of the strategic computing systems within the institution.

Another major difference with the examination timetabling problem is not only related to differences in the nature of the information and constraints but in the style in which the solution is constructed. Overwhelmingly in all consultancy and implementation undertaken to date within the Company, the timetable is constructed prior to student enrolment and therefore optimised on projected student numbers taking particular combinations of modules. In many cases the goal of optimisation is sacrificed for the sake of getting a solution which is workable. Student clashing is related to defined course structures as opposed to the examination counterpart which is based purely on student enrolment to assessment events. Regarding soft constraints, the emphasis is on the ability to offer as many options as possible as opposed to best spread across a particular examination session. Administrators employ heuristics that suggest what modules should be made available to particular courses and which ones should not. Indeed, this information can often be inferred from previous year’s data or obtained directly from members of particular schools. Because the timetable is constructed pre enrolment, inefficiencies occur which are allowed to ripple throughout the rest of the year. Based on the initial construction and space utilisation, potentially the problem could be is reshuffled or indeed amended based on a different measure of optimisation. This optioned is not presently favoured by institutions due to the disruption that would be caused. There are a number of reasons timetabling pre enrolment; if it were left entirely to student choice there is no guarantee that a feasible timetable could be constructed and secondly, more and more emphasis on opening access to universities dictates that students with busy lives need to know timetables before choosing optional parts of the course. Many universities used a phased approach which is a combination between pre and post enrolment. More work is required to understand the issues involved and where, what and how search techniques and indeed what measures of optimisation can be used.

It is clear that the improvement of solutions will come about through the combination of high level heuristics and optimisation techniques. The research challenge is therefore identified as the requirement for detailed studies of how the aims, objectives and practicalities of timetabling within institutions interlink.
3.2 Building the Institutional Model

As with examination timetabling, the timetable construction process can be broken down into a series of information and solution modelling. Even more so than with the examination problem, this process is complicated. As stated, this is related to the number of interested parties and diversity of the data requirements. Attempts have been made to provide a general framework to aid this situation. For example, work has been carried out proposing a generic architecture for the production of a timetable by examining the full range of procedures and the associated characteristics [64]. Also in [65], a framework was presented allowing the researcher to combine many different solution methods in arbitrary ways in the solution of a single problem. Such contributions have provided an important platform upon which we can build. A more complete description to enable understanding of the specific needs of the modelling process is required. The following impacts on a number of key issues.

In the case of course timetabling, information modelling can be broken into data, constraint and course structure modelling with solution modelling being dominated by factors related to optimisation and evaluation. Although it is an important issue, algorithmic modelling is not discussed here because the focus of this discussion is concerned with highlighting the high level challenges that need to be addressed if the gap between theory and practice is to be closed. In many respects, the key to narrowing this gap in relation to course scheduling is related to the modelling of the entire problem, thus identifying where and when in the process search techniques may be of use.

3.2.1 Information Modelling

In terms of information modeling, the main differences with examination timetabling is the much more incomplete nature of the data requirements [45,64] which are much more substantial. Data is required on events, course structures, the estate and the lectures / instructors availability and expertise. From the author’s experience, it is evident that a combination of poorly implemented information strategies and reluctance of staff within the sector has led to a position where this information is difficult to obtain. This situation inevitably leads to significant changes in the timetable formulation at the beginning of the period in which it is required. Work has been carried out on ensuring a changed solution is close as possible to the initially modeled solution after changes in the original definition. For example see [45].

In many instances, expert timetablers have dealt with the initial construction by adopting a series of high level heuristics. For example some institutions use a centralised approach initially, timetable a percentage of the required events in a percentage of the available centrally ‘owned’ rooms thus allowing individual schools / departments to ‘fill in the blanks’ in the remaining rooms or indeed in departmentally ‘owned’ rooms [34]. Many such high level heuristics are used within institutions during the construction process, little of which (to the author’s knowledge) have been reported in the literature. In general, these relate to space usage and decomposition within both the information and solution modelling process. This emphasises the fact that an important challenge for the research community is therefore to review real applications of course scheduling techniques and software with the aim of identifying the major themes which will facilitate the construction of robust initial solutions. High
level heuristics need to be identified, analysed and modeled in terms of constraints and evaluation. In general these usually relate to student and staff preference and space usage.

### 3.2.2 Course Structure Modelling

Modelling the course structure is a difficult and important aspect of the information modelling process. This aspect is completely unnecessary in the examination counterpart. Course timetabling raises a variety of issues relating to when staff / rooms are available and what events should be timetabled with which others. The later of these issues becomes more difficult when, as discussed earlier, it is dictated that a timetable must be ready before student enrolment. The research challenge is therefore in identifying easy intuitive ways of representing constraints. Attempts have been made to specify a standard timetabling data format that is complete and universally applicable [51,52,53,68]. This work needs to be extended and made more readily available to enable users to identify and model constraints thus allowing the interface between users and researchers to become better defined.

Another important issue is the division of students attending a lecture into sub events such as tutorial classes. In examining this in detail a number of key issues are explored. Consider the case involving the separation of students enrolled on a particular course into tutorial classes. Consider, also, a lecture event which has $x$ students. If the preferred size of tutorials is $y$, then it is trivial to calculate that $x/y$ tutorial slots are required. The interesting research issue considered here, however, is in what way to split the $x$ students into groups while ensuring that maximum flexibility is introduced into the timetable i.e. what are the best combinations of students to be timetabled in which slots. In addition this must be done in a manner to allow room usage to be maximized while ensuring that students are allocated throughout the week with cognisance taken of their existing commitments on events related to other courses. This is often done manually by allowing students to self-select particular slots from a set of pre-established time slots. In the course timetabling literature, the majority of influential work on course sectioning (sometimes termed 'splitting') has concentrated on timetabling courses, where lectures, tutorials and laboratories etc. are not distinguished between each other [42, 37,31,39,62]. Apart from a few notable exceptions [40], courses or groups of students are subdivided into groupings for the purpose of offering student choice as opposed to reflecting the structure of events which constitute the structure of the course. The objective is normally related to balancing the size of the groups while offering students maximum choice, this enabling them to enroll on their choice of modules.

Within the UK in particular, universities subdivide students in line with course structures. The main problem with this current definition of course splitting is that sub events do not inherit parental clashing constraints [59], apart from where a lecture event is subdivided. There are also some work dealing with students sectioning problems dated back to 80s [39, 43]. Once again, this work is different from what we are considering here, where students are divided into sub-groups as opposite to multi-groups. More recently, Fuzzy algorithms have been used [44] to cluster students in large classes into groups which may later lead to the fewest possible conflicts in timetables. Beyrouthy et al [59] considered the problem of splitting in relation to space objectives by investigating splitting of courses of same type event into sub
events of that type for the purpose of fitting into particular room profiles. During the years little has been done on partitioning the students into actual sub events as dictated by the course structure. In [40], meta-heuristics are proposed to address the Availability-based Laboratory/Tutorial Timetabling Problem (ALTP). This offers a very promising platform for further exploration into the automatic constructing of timetables while providing a solution which assigns students to the ‘best’ timeslot based on a defined week range. In should be noted that in doing so, it is important that the needs of all parties need to be addressed. This raises the interesting concept of how an attained solution should be measured. When producing a course timetable within an institution, it is important that the timetable produced is seen to be fair and equitable to all interested parties. The challenge to research is investigation of these and other information modelling issues. This will be further discussed in the next section.

Another aspect of course structure modelling is related to the timetabling of associated events together. It is important to provide the ability to link particular events under the notion of course structure and schedule them as a ‘package’. This concept is similar to Kemp chains in examination timetabling [46]. This macro event scheduling process will allow the basic building blocks of the course timetabling problem to be sustained throughout the process. This approach has the advantage of reflecting organisational and course make up. In addition it may be possible to decide which events / courses have similarities and can be linked together when timetabling based on individual of indeed groups of characteristics. For example, pathways within a particular school could be timetabled together at the same time using the same departmental space. This mimics the construction process already in existence within an institution where the overall timetable is broken into a number of sub units which are timetabled at a particular time by a particular person. This subdivision or decomposition of the timetabling is a challenging research aspect which needs further investigation. Macro events may be based on a combination of course structure and clusters. Academic timetabling problems tend to show signs of clustering related to the organisational structure. For instance modules from a Math’s school will clash other modules from that school. Further to that those modules will tend to clash with other science subjects such as physics and chemistry. What is required is a way of splitting such problems into smaller sub-problems in such a way that any crossover between events in different sub-problems is kept to a minimum.

### 3.3 Solution Modelling

Within the context of developing and delivering an institutional wide timetable, it must be clear what the optimisation issues are and how they are to be measured. The measurement of optimisation itself is quite different from the measure needed for the examination problem. There is sometimes a view in the research community that it is possible to define the course timetabling problem by simply altering the optimisation function used within the examination timetabling problem. However, this formulation does not define how institutions view the quality measure of a particular course timetabling solution. Institutions are interested in a combination of room usage, staff and student satisfaction. The first of these is measurable by multiplying occupancy
by frequency e.g. how many students use a room how often. The measurement of utilisation is an average of multiplication of occupancy and frequency over a set 40 hour week. Staff satisfaction is measured by the extent to which teaching duties can be ‘bunched’ together leaving time for research and other activities. In many cases, academic staff members insist on the concept of a ‘research day’. As a further advantage, it is often considered advantageous if undesirable hours can be identified and minimised per member of staff. This is termed here as the ‘share bad hours’ heuristic and is an example of a new soft constraint to be considered when optimising the construction and improvement of an institutional course timetable. Student satisfaction can be measured by the spread of events and the availability of choice within a particular course structure. As already mentioned, ‘best spread’ has quite a different meaning in this context. A number of other issues are relevant to the overall construction problem but not the optimisation problem e.g. staff satisfaction can further be measured by the ease at which information is gathered from them.

As previously stated, in many cases optimisation is sacrificed for the sake of getting a solution which is workable e.g. the definition of a ‘good’ solution is driven by the need to have any solution based on a subset of the actual event types which are required [47]. This has the effect of meaning that a feasible solution is judged at an early stage in the construction process as opposed to answering the question as to whether or not the solution is actually workable e.g. can all additional events not timetabled be accommodated after student enrolment. When students arrive and populate the skeleton structure of the timetable, solutions to individual problems of over subscription are obtained through negotiation and compromise. The overriding factor which makes the entire process workable is the fact that currently universities utilise on average about 30 percent of their space effectively [61,63]. One explanation for this is that space utilisation is low because of the inherent flexibility within the timetable i.e. staff and students have a lot of choice. Unfortunately, this is not always the case as timetabling concerns rate highly in both student and staff surveys [38]. Further evidence of the inflexible nature of the course timetable is the fact that universities are not able to accommodate more students easily or indeed plan new or change existing course delivery. The author’s view is very much like that of Carter [42] e.g. More work needs to be completed to understand the relationship between space usage, staff flexibility and student choice. It is therefore essential that metrics are produced to measure the effectiveness of timetables from all perspectives.

It is suggested that the optimisation function used to measure the quality of the problem solution must be constructed in such a manner as to take in the multi criteria associated with each area. Whereas, optimisation is relatively easily defined for examination scheduling, it is difficult to define for course scheduling. From the author’s experience, it can be defined as a balance between keeping all the stakeholders happy e.g. student choice, staff flexibility and room usage. Therefore, to aid with the automation of the task, the construction and optimization of the solution must take into consideration three distinct areas as an absolute minimum. In addition, in evaluating a given solution to the course timetabling problem within an institution, the users need to understand the situation in terms of the outcomes of individual constraints associated with all identified areas. The multi-objective approach has received significant recent [48,49,50] interest with respect to timetabling and, with
respect to course timetabling, will be able to better express and illustrate the features of a solution to a problem.

4 Conclusion

This paper outlines the major challenges which face those researchers working in the area of university exam and course timetabling. While not trying to exhaustively referencing the literature, detail is provided of the relevant research in both areas. The challenges are presents from the perspective of the author’s experience and experience of working closely with the educational sector. The intention is to stimulate debate in the literature by providing opinion based on practical implementations. The aim is the improvement of techniques and hence software tools available to the sector to help with this most difficult and time consuming aspect of university administration.

In relation to examination scheduling the identified challenges to researchers in the area include the following;

(i) New datasets becoming available on a regular basis encompassing more real world requirements.
(ii) The development of robust techniques which are able to deal with the information poor environments within which examination timetables are often developed.
(iii) Investigation of a reformulation of the problem, including new hard and soft constraints which better reflect the real world environment.
(iv) Identification and comparison of key dataset characteristics and potential linkages with the likely best search approach to be taken.
(v) The investigation of all aspects of solution quality in the provision of the ‘best’ solution for the institution.
(vi) The exploration of new search technologies in establishing how developed systems can be made more general.
(vii) Investigation of how to incorporate user interface design with the inherent complexity of the problem.
(viii) Wide ranging Investigation of different neighbourhood structures and fitness landscape within the context of real world problem solving environments.

In relation to course timetabling, the following research themes are highlighted;

(i) Investigation of techniques to deal with the distributed, information poor environment in which course timetables are produced.
(ii) Standardisation of datasets, constraints and modeling languages influenced by real world scenarios.
(iii) Investigation of the role in user interaction in the design of decision support system for course timetabling.
(vi) Investigation of the need for the reformulation and modeling of the problem. It should be noted that this represents a far greater challenge within the context of course timetabling than it does for examination timetabling.

(v) Identification and adaptation of high level policies and practices that are employed by administrators within institutions to construct initial solutions.

(vi) Experimentation related to heuristic approaches to subdivision of events.

(vii) Investigation of the effect of pre- and post-enrolment production of the timetable on the approaches taken to optimisation e.g. penalty used.

(viii) Undertake an investigation into the delivery of more sophisticated models which capture the complexity and multi-objective nature of timetable evaluation in the real world.

(ix) Investigation of the important linkage between space usage and flexibility within the academic timetable.

(x) Investigation of approaches involving decomposition and ‘macro event’ timetabling.

In summary, this paper has outlined a number of significant research challenges which provide a rich area for research into automated search methodologies for educational timetabling. Moreover, by addressing these demanding research issues, the scientific community will be taking a step towards closing the gap between theory and practice which has existed for so long.

Acknowledgements

I would like to thank Professor Edmund Burke for his contribution throughout the preparation of this paper.

References


8. S. Petrovic, V. Patel, Y. Yang, Examination timetabling with fuzzy Constraints, In [7], pp. 313-333.

9. H. Asmuni, E.K. Burke, J.M. Garibaldi, B. McCollum, A Novel Fuzzy Approach to Evaluate the Quality of Examination Timetabling, Accepted for full paper at PATAT06.


13. E.K. Burke, B. McCollum, and P. McMullan, Examination Timetabling: A New Formulation, Accepted for abstract at PATAT06.


35. B. McCollum, Bridging the gap between research and practice: University timetabling in the real world – KEYNOTE, Proceedings of the 47th Annual Operational Society Conference (OR47), September 2005, Chester, UK.
38. B. McCollum, 2003-2004 Academic Timetabling: Analysis of Staff and Student perception. Internal report eMAP04/02/01.
42. M.W. Carter, A Comprehensive Course Timetabling and Student Scheduling System at the University of Waterloo, in [5], pp. 64 - 84.
44. M. Amintooosi, J. Haddadina, Feature Selection in a Fuzzy Student Sectioning Algorithm, in [7], pp. 147-160.
45. T. Muller, H. Rudova, R. Bartak, Minimal Perturbation Problem in Course timetabling, In [7], pp. 126-146.
47. B. McCollum, M. McKillop, P. McMullan, Course Scheduling: The Division of Lecture Events into Tutorials. Internal Report CSS/04/12abs.


56. EK.Burke, B.McCollum, A.Meisels, S.Petrovic, and R.Qu, A Graph-Based Hyper Heuristic for Educational Timetabling Problems. Accepted for publication in the European Journal of Operational Research, 2006.


60. C.Beyrouthy, E.K.Burke, J.Landa-Silva, B.McCollum, P.McMullan, A.J.Parkes Understanding the Role of UFOs Within Space Exploitation. Accepted as an abstract for PATAT06.


68. L.P. Reis, E. Oliveira, A Language for Specifying Complete Timetable Problems, in [6], pp. 322-341.


