Towards a reference model for timetabling and rostering

Patrick De Causmaecker

Katholieke Universiteit Leuven
Department Of Computerscience
DDID group Campus Kortrijk
E. Sabbelaan 53, 8500 Kortrijk, Belgium

Abstract. In this talk we discuss the need for a reference model in the field of timetabling and rostering. The reference model should allow to describe the interaction between different levels of planning and between the subjects involved to provide the information for the definition of applications at the most detailed level. Furthermore it should allow to systematically study complexity and hardness of the problems in the domain. We present a notation for rostering that allows classification and argue that it can be used in such a systematic study.

1 Introduction

PATAT conferences have studied a large number of instances of timetabling and rostering problems. Among the many application domains figuring as suppliers of case studies were school and university timetabling, sports timetabling as well as employee timetabling, agenda building and more. Authors often presented a detailed description of a specific case. They developed data models, identified the relevant hard and soft constraints and defined goals for the algorithms which they subsequently developed and tested. Not surprisingly, given the model complexity and the required amount of flexibility, the algorithms were almost always heuristics designed according to some metaheuristic paradigm.

In doing so, timetabling researchers have achieved progress along at least three important lines:

- They discovered and studied a large collection of situations for which they created sophisticated models
- They built an interesting playground with large and difficult problems for metaheuristic algorithms, studied algorithmic behaviour and proposed new techniques and devices
- They exhibited the need for - and partly established - systematic testing and evaluation procedures based on well balanced benchmarks, often drawn from real world situations

The presently available corpus on timetabling is sufficiently mature to find its way to industrial applications. Examples of such applications have emerged. Two important questions arise:

– Which specific input is needed and in which form should it be presented to allow industry to make optimal use of the present and future scientific state of the art.
– How can we facilitate future developments and position them in the corpus of knowledge.

The last PATAT conference enjoyed two plenary talks addressing these questions. Barry McCollum, from an industrial perspective, diagnosed and described treatment for the gap between theory and practice in university course timetabling [18]. Andrea Schaerf reported on work in collaboration with Luca Di Gaspero in an inspiring contribution on measurability and reproducibility in university timetabling research [21]. It is the aim of the present talk to extend on these two lines of thought.

We will discuss the development of a consistent reference model and a structured categorisation of systems. Application developers will be able to position their product with respect to this model. They will use it to discover opportunities for new developments. It will guarantee communication, interchange and collaboration between components aiming at very different users. From a research point of view, the model will structure the ideas that have developed in the community and put forward new research opportunities. The categorisation system will allow researchers to evaluate and compare approaches. We want to improve the input of real world information, distinguish the different levels of abstraction and study the complexity of the resulting decision support problems.

Input The domain models and constraint sets have been collected by individual researchers in individual cases. The necessary information was often obtained through interviews with people in the field. This is true for the hard constraints and regulations, although these could often be inferred from the problem at hand, from available legislation or from contract agreements. It is even more true for the soft constraints expressing preferences and working comfort expectations of the personnel (teachers, nurses, football players) or of the subjects (students, patients, public). It is hard to bring structure to this material gathered in an ad hoc fashion without a scientific guideline. This guideline can come from other scientific disciplines within the OR community as well as in humanities. We investigate what law, social, economical and management sciences may have to say [9].

Levels Although timetabling is typically concerned with the most detailed level of decision making, it is not independent of decisions taken at more abstract levels. The goal for the timetabling activity - what determines a good working schedule - cannot be defined independently at the different levels of decision making. After all, these goals are mostly set at the more strategic levels. We will thus not be able to study and define general reference models without taking all these levels into some account. As a consequence, the reference model will ultimately not only describe how timetabling components for the most detailed, short term decision making should be conceived, but also how decision support tools for the mid and long term should behave and interact. We want to pinpoint the best level for each decision act.

Complexity Only a few authors have tried to define relevant complexity measures (e.g. [22]). If the reference model has to provide reliable information about the applicability of a specific technique for a specific problem, such measures are indispensable. Building a system of measures is a non trivial but rewarding exercise. We can build on results in other domains and problems.

2 Modelling, solution techniques and real world issues

Some proposals have come up in previous PATAT conferences. These contributions highlight the need for standardisation and classification and for interchangeable formats. Adoption of these approaches by the community could lead to a platform for comparative study of problems and algorithms. Until today however, such adoption has been very limited or non existing. Researchers in the field seem to feel that the domain has not been sufficiently foraged to settle down on a final definition.

2.1 Languages

A number of authors have presented language constructs to describe timetabling problems. We mention Kingston's STTL [5, 15] and Özcan's TTML scheme [19]. We proposed to devise an ontology for timetabling and used it in a framework [8]. The competitions have defined their own formats. These efforts are still far from being accepted and used in general. They were built from the experience of the authors and they do offer a systematic way to describe specific classes of timetabling problems.

2.2 Modelling devices

The analogue of the linear or integer program in commonly used metaheuristics is a model of the solution space and a neighbourhood structure. Timetabling has a number of general characteristics and much effort has been spent on catching these characteristics in metaheuristic approaches. Looking at the literature, authors often refer to each other's neighbourhood structures. Some authors have tried to develop generic neighbourhoods. An example is the tiles neighbourhood of Kingston [16]. It is intuitive for the users and thus gives them some understanding of the solution method. By allowing users to define the tiles, the algorithm can be given hints as to where to look for good solutions. Although Kingston introduced tiles in the context of high school course timetabling, they are generic as we demonstrated in [10] where we modified them for use in a university course timetabling problem.

The need to use experience and knowledge from the user is one of the reasons for the introduction of hyperheuristics. Acquiring data about the efficiency of individual steps in an algorithm may be used in the systematic study that we envision. (e.g. see [2, 6])

3 Levels

This community has been mainly concerned with the short term planning and scheduling of people's efforts. We distinguish the following parties and resources.

- Receivers of the services - e.g. students, patients, elderly - requiring a certain amount and quality of attention.
- Deliverers of services - e.g. teachers, nurses, care workers, doctors - with a certain qualification and able to perform a number of actions.
- Resources - e.g. rooms, operation theatres - with often specific equipment needed for certain actions.

It is then standard procedure to collect data concerning these available means and their associated constraints, as well as about requirements and expectations of the receivers. We have thus nearly always been active at the lowest operational level and suppose that decisions about organisational capacity and activity have been taken. The output of this higher level decision making process is considered input for our systems.

This situation is comparable to what is currently happening in production planning and scheduling where one traditionally distinguishes at least two levels.

- At the business level, decisions are taken based on orders and product production trees.
- At the lower level these production plans are decomposed in jobs for specific machines in specific timeslots.

The situation is further complicated when multiple companies are involved in such a production system. The *supply chain* defines a planning problem that crosses company borders. The detailed scheduling problem has been studied in great depth (see e.g. [14]). Business systems and planning models at this level have emerged leading to a number of standards. Supply chain management is a discipline of its own. It is clear that this decomposition leads to suboptimal solutions. Only recently, people have started addressing the global optimisation problem. It turns out to be hard to translate the detailed terminology of the short term scheduling to the mid term planning level. Stated differently, it has for a long time been unknown which quantities can be exchanged between the two levels. This situation has led to a specific niche, the domain of *manufacturing execution systems*, which plays the role of an intelligent interface between the business systems and the production line. Systems at this interface typically provide sophisticated logging, tracking and reporting together with some level of detailed scheduling.

In the field of timetabling the situation is comparable but further complicated by the fact that humans are involved. [1] One typically distinguishes decisions to be taken for the long term (years), mid term (months) and short term (weeks).

On the long term decisions are closely linked to selected strategic options. At this level the existence of a unit or the discontinuation of an activity are decided. Expected costs and revenues are weighted. Information that is needed is at what cost a specific unit can operate, what its relative impact may be on the other activities and whether or not profitable synergies can be expected. The market evolution, expected demand and prospects for the future are taken into account.

Mid term decisions are tactical and influence the operational capacities of the units as they exist. Tuning allocation of staff to expected performance and productivity is the main subject at this level of decision making.

Short term decision making starts in general from a given man power and demand to produce feasible rosters and working schemes that meet both the demand and the constraints on the individual working schedules.

These three levels clearly do not operate independently. Among the subjects to be studied, we mention:

- We need instruments to feed the higher level decision making from below. Experience in building rosters at the short time level should be aggregated to serve as input for the mid and long.
- Decisions at the staffing level should take specific constraints into account. Feasibility of a certain construction should be checked. We need standard measures and simulation tools. (e.g. [12])
- Rules for 'good' rosters should include, along with legal constraints and local habits or conventions, results from the theory of labour ergonomics (see e.g. [1, 3]).

A possible scheme for health care is presented in Fig. 1. Each rectangle in this scheme is a target for a line of research and/or for software developers. The interactions between these rectangles could become subject of an interdisciplinary study [9]. Some of the disciplines are shown on the third axis in Fig. 1.

4 Study of the hardness of rostering problems

One tool that could help at the level of staffing is a predictor for timetabling problems. Such a predictor could e.g. tell to what extent the constraints can be expected to be satisfied. Given a possible staffing decision, the decision makers could use the tool to judge its feasibility in practice. Whether or not such a predictor is feasible remains an open question. Predictors have been established in other domains where they could often be used in a preprocessing phase to decide which algorithm is the most appropriate for a certain instance [17, 23].

If we want to obtain generic results, we need a categorisation system of problems allowing to compare results from different application environments.

[1] The fields of production scheduling and timetabling are presently merging as can be seen from some recent publications

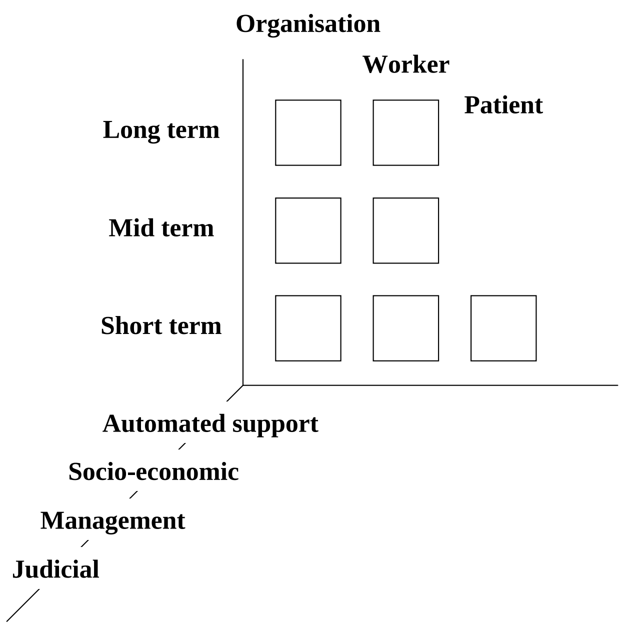

Fig. 1. Parties, planning horizons and disciplines

Research results could then be classified using this categorisation, so that they could be reused in similar circumstances. A classification in nurse rostering was developed in [7]. We extend on it to propose a notation for the timetabling sub field of rostering.

4.1 A categorisation of rostering problems

Rostering problems are defined by three elements:

- The available personnel, the people that can be assigned a certain task at certain points in time. We will call this the *personnel environment*.
- The work that has to performed, the set of tasks and the associated time structure. We will call this the *work characteristics*.
- The conditions to be optimised, the objective that is to be met by the schedule. We call this the *optimisation objective*

These three elements bear close resemblance to the $\alpha|\beta|\gamma$ categorisation in production scheduling (see e.g. [4, 13, 20]). We adopt this structure and notation for rostering problems, see Table 1.

	Personnel constraints		Skill	
	A	Availability	1	Single
α **: Personnel environment**	S	Sequences	N	Multiple
	B	Balance	I	Individual
	C	Chaperoning		
	Coverage constraints		Shift type	
	D	Determined	1,2,3	Limited
β **: Work characteristics**	R	Range	N	Multiple
	T	Time Intervals	O	Overlapping
	V	Fluctuating		
	Objective		Mode	
	P	Personnel regulations constraints	E	Exact
γ **: Optimisation objective**	L	Load and coverage constraints	O	Optimisation
	X	Number of personnel	M	Multi objective
	G	General		

Table 1. Classification of nurse rostering problems

- **Personnel environment** (α) The personnel environment distinguishes the categories A (availabilities), S (sequences), B (balance) and C (chaperoning). A describes the availability of each member of personnel and allows expression of preferences. S is used for constraints on the consecutiveness of assignments, e.g. the allowed number of consecutive night shifts. B allows the expression of balance constraints on the schedule, e.g. to enforce some level of fairness. C can be used to express the need for combined tasks where one member of personnel must necessary work together with another one. Independently from these categories, personnel may have individual skills.
- **Work characteristics** (β) The work characterised by the demand to be covered (coverage) and the shift structure. We distinguish D (definite), R (range, V (fluctuating) and T (time interval) types of coverage and 1 (single) N (multiple) and O (overlapping) types of shift structures.
- **Optimisation objective** (γ) Possible objectives are P (personnel constraint violation), L (demand) X (number of personnel) and G (general). These may be subject of O (optimisation) or M (multi objective optimisation) which may be E (exact) or not.

4.2 Example

A detailed description of the elements in Table 1 is published elsewhere [11]. In this paragraph we present a simple example. Suppose a ward is working with nurses in three skill categories, say x,y and z, where a nurse from skill type y can work as a nurse in type z (N). It has regulations constraining the number of consecutive assignments, assignments in weekends, minimal length of the breaks and so on (S). Nurses have to work a minimal number of hours and can ask for some days off (A). The workload for the nurses should be balanced (B). This activates labels A,S,B and skill type N in category α.

Furthermore, the organisation uses five shift types which partly overlap in time (O). The coverage is given as a range per shift, no time intervals are used in this specification, but the coverage varies per shift (R, V).

As an objective, the strategy is to always meet the coverage constraints, eventually violating the other constraints at some given cost which must be approximately minimised (P, O).

The resulting problem can thus be classified as $(A, S, B, N|R, V, O|P, O)$. This example may make it clear that our notational system is not closed. It should at any time be open for extension if a new problem comes up, and the level of detail should also be set according to the needs of the research at hand. We did keep the size of the problem out of the notation as we think that this should be a parameter in any systematic investigation of a set of rostering instances. What we do propose is to investigate the resulting categories systematically and exhaustively in order to determine the hardness of the instances included. This would allow to draw charts for problems and algorithms to be used by researchers and developers. We went through part of the nurse rostering literature since 2004 and were able to classify most of the problems discussed. We furthermore performed a systematic analysis of a specific category. These results will be published later [11].

5 Conclusion

In this talk we discussed the need for a reference model for timetabling and rostering.

An essential step in the development of such a model is a multi disciplinary effort. We think such an effort is necessary to build a model in which all levels of personnel planning systems can be integrated.

Another ingredient is a classification of problems that can be used for study and comparison. We propose a notation for rostering. It is sufficient to classify most problems in recent literature. Once a category has been defined, a systematic study on the influence of the parameters in the category on the hardness of the problem can be undertaken. This allows to draw charts that can be used in the development of automated rostering systems.

Future work includes an extensive study of the different categories, the extension of the notation to other fields of timetabling and the interdisciplinary effort to set up an integrated reference model.

In this contribution we wanted to join in and extend on ideas presented before at PATAT [18, 21]. We look forward to seeing more contributions along these lines in current and coming PATAT conferences.

6 Acknowledgements

The work reported on and referred to in this talk is a close collaboration with colleague Greet Vanden Berghe of KaHo. Kristof Salomez of VUB, Paul Gemmel of UGent, Christine Baaijens and Luc Berlanger of Déhora were excellent partners from respectively law, management and socio-economical sciences in the interdisciplinary discussions. PhD students Stefaan Haspeslagh, Burak Bilgin and Tommy Messelis did a great job in the analysis for the experiments and to collect the data.

References

1. Baaijens C., van Doorne-Huiskes J., Schippers J. Do Dutch employees want to work more or less hours than they actually do? In: B. Peper, J. van Doorne-Huiskes, L. den Dulk (Eds.), Flexible working and organisational change: the integration of work and personal life. Cheltenham: Edgar Elgar, 2005
2. Bilgin B., Özcan E., Korkmaz E.E. An Experimental Study on Hyper-heuristics and Exam Timetabling. In Burke E. K., Rudová H. (eds.): Practice and Theory of Automated Timetabling VI, Selected Papers. Lecture Notes in Computer Science, Vol. 3867, Springer, Berlin (2007) 394-412.
3. Bourdais S., Galinier Ph., Pesant G.: HIBISCUS: A Constraint Programming Application to Staff Scheduling in Health Care, F. Rossi (Ed.): CP 2003, LNCS 2833, 153-167, 2003
4. P. Brucker (1998): Scheduling algorithms, 2nd edition, Springer, Heidelberg.
5. Burke E. K., Pepper P. A. and Kingston J. H., A Standard Data Format for Timetabling Instances, In Burke E.K., Carter M. (eds.) Lecture Notes in Computer Science, Vol. 1408, Springer, Berlin (1997) 213-222.
6. Burke E.K., Kendall G. and Soubeiga E. A Tabu-Search Hyper-Heuristic for Timetabling and Rostering. Journal of Heuristics, 9(6), 451-470, 2003
7. Burke E.K., De Causmaecker P., Vanden Berghe G., Van Landeghem H.: The State of the Art of Nurse Rostering, Journal of Scheduling, 2004, Vol. 7, No. 6, Nov/Dec 2004, 441-499
8. De Causmaecker P., Custers N., Demeester P., Vanden Berghe G.: Semantic Components for Timetabling, E.K. Burke, M. Trick (Eds), Revised Selected Papers of 5th International Conference on Practice and Theory of Automated Timetabling V, LNCS 3616, 2005, 17-33
9. De Causmaecker P., Vanden Berghe G., Baaijens C., Gemmel P., Salomez K.: Workforce Planning and Scheduling: an Integrated Approach for Health Care, K.U. Leuven working paper (2007).
10. De Causmaecker P., Demeester P., Vanden Berghe G.: A decomposed metaheuristic approach for a real-world university timetabling problem, European Journal of Operational Research, accepted for publication, 2008
11. De Causmaecker P., Vanden Berghe G., Haspeslagh S., Bilgin B., Messelis T., Categorisation of rostering problems for hardness studies, K.U. Leuven working paper (2008).
12. Gemmel P., Vandaele D., Tambeur W.: Hospital Process Orientation (HPO): the development of a measurement tool, accepted for publication in Total Quality Management and Business Excellence, vol.19, 2008.

13. R.L. Graham, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan (1979): Optimization and approximation in deterministic sequencing and scheduling: a survey, Ann. Discrete Math. 4, 287-326
14. Halldorsson, A., Kotzab, H., Mikkola, J. H., Skjoett-Larsen, T. (2007). Complementary theories to supply chain management . Supply Chain Management: An International Journal, Volume 12 Issue 4 , 284-296
15. Kingston J.H., Modeling timetabling problems with STTL, In Burke E. K., Erben W. (eds.): Practice and Theory of Automated Timetabling II, Selected Papers. Lecture Notes in Computer Science, Vol. 2079, Springer, Berlin (2001) 309, 2001.
16. Kingston J. H.: A Tiling Algorithm for High School Timetabling. In Burke E. K., Trick M. (eds.): Practice and Theory of Automated Timetabling V, Selected Papers. Lecture Notes in Computer Science, Vol. 3616, Springer, Berlin (2005) 208-225.
17. Leyton-Brown K., Nudelman E., Shoham Y. Learning the empirical hardness of optimization problems: The case of combinatorial auctions. In Constraint Programming (CP), pages 556-572, 2002.
18. A Perspecive on bridging the Gap Between Theory and Practice in University Timetabling. In Burke E. K., Rudová H. (eds.): Practice and Theory of Automated Timetabling VI, Selected Papers. Lecture Notes in Computer Science, Vol. 3867, Springer, Berlin (2007) 3-23
19. Özcan E., Towards an XML based standard for Timetabling Problems: TTML, Multidisciplinary Scheduling: Theory and Applications, Springer Verlag, 163 (24), May 2005.
20. Pinedo M. L., Planning and Scheduling in Manufacturing and Services, Springer Series in Operations Research, Springer Verlag 2006
21. Schaerf A., Di Gaspero L., Measurability and Reproducibility in University Timetabling Research. In Burke E. K., Rudová H. (eds.): Practice and Theory of Automated Timetabling VI, Selected Papers. Lecture Notes in Computer Science, Vol. 3867, Springer, Berlin (2007) 40-49
22. Vanhoucke M., Maenhout B.. Characterisation and generation of nurse scheduling problem instances. working paper, Gent, 2005.
23. Xu L., Hoos H.H., Leyton-Brown K.. Hierarchical hardness models for sat. In Principles and Practice of Constraint Programming (CP-07), Lecture Notes in Computer Science 4741, pages 696-711. Springer Berlin, 2007.