Data Formats for Exchange of Real-World Timetabling Problem Instances and Solutions

Discussion Session (Abstract)

Jeffrey H. Kingston

Keywords Timetabling · Data Formats · Benchmarks

1 Abstract

The problem of exchanging timetabling data has been a perennial topic at PATAT conferences. Indeed it was discussed at the very first (Cumming and Paechter 1995).

The principal difficulty is the large variety of kinds of requirements. To quote from a recent paper (Post et al. 2008): ‘This complexity with the specification of the problem has been addressed in several ways. Some papers have tried to generalize and unify the constraints (Chand 2004; Kitagawa and Ikeda 1988). Others have adapted existing technologies in which constraints may be expressed, such as XML and the semantic web (Custers et al. 2005; de Causmaecker et al. 2000, 2002; Özcan 2003), or object-oriented modeling and frameworks (Grobner et al. 2003; Ranson and Ahmadi 2006). Others have expressed constraints as logic expressions within specifically designed specification languages (Burke et al. 1998; Kingston 2001; Mata et al. 1997). There has been at least one attempt to simply enumerate every possible constraint (Reis and Oliviera 2001).

‘Another approach is to restrict the problem domain to one particular kind of timetabling, then use judicious simplification to further reduce the specification burden while maintaining the essence of the problem. The Carter data sets for examination timetabling (Carter et al. 1996) omit many details, notably all data related to rooms, and similar simplifications appear in the Traveling Tournament Problem (Easton et al. 2001) and the International Timetabling Competition (Paechter 2003). These are some of the most successful examples of timetabling data exchange. However, judicious simplification has been criticized for contributing to the gap between research and practice (Burke et al. 2006), at least in examination timetabling; and the data transfer has almost always been in one direction only.’

The paper from which this quotation is taken goes on to define an XML format for exchanging real high school timetabling problem instances and solutions. Since that paper

Jeffrey H. Kingston
School of Information Technologies
The University of Sydney, NSW 2006, Australia
`http://www.it.usyd.edu.au/~jeff`
E-mail: jeff@it.usyd.edu.au

was written, the format has been refined, and a set of instances and solutions from widely varying institutions around the world has been collected and stored in the format (Post 2009). An evaluator, which compares solutions against instances and produces badness values, has been made available as a web service (Kingston 2009).

The purpose of this discussion session is to explore the idea that the general approach taken by the high school timetabling project should be adopted for real-world data exchange in other sub-disciplines of timetabling.

Two key points define this approach. First, the format was developed in consultation with several groups of researchers, and is avowedly inclusive: if the accurate expression of real-world instances requires that certain features be present, there is a commitment that they will be added. Second, the format is not overly general: it is limited to high school timetabling, and it does not attempt to define constraints by logic expressions that researchers are unlikely to want to read and interpret. Instead, there is a fixed list of constraint types of predetermined meaning. At the time of writing there were 15 constraint types; more will be added as required.

It might seem that this proposal offers nothing new compared with the formats used by the Carter data set or the International Timetabling Competition cited above. The new point is the demonstration that it is feasible to specify *real-world instances* from *disparate sources* in *complete detail*, by limiting attention to a single sub-discipline and a fixed list of constraint types.

The high school project was influenced by an earlier project in nurse rostering. After the need for exchanging nurse rostering data had become clear (Burke et al. 2004; Cheang et al. 2003), an XML data format was defined and a web site (Curtois 2009) set up in the year 2005 (personal communucation from T. Curtois). This is the earliest example known to this author of real-world instances and solutions from disparate sources being brought together.

Another point for discussion is whether specific features of the high school format could be re-used in formats for other sub-disciplines. To support this discussion, the remainder of this abstract introduces the high school format.

The format allows any number of instances of the high school timetabling problem to be stored in one file, along with any number of sets of solutions to those instances, each set contributed by one researcher. The file may thus be a comprehensive archive, and the evaluator may produce tables comparing the solutions, of the kind frequently seen in research papers. The existing evaluator (Kingston 2009) does this.

Each instance consists of a `<Times>` section (the concrete format is XML) listing the times of the cycle in chronological order, a `<Resources>` section listing the resources (typically but not necessarily student groups, teachers, and rooms), an `<Events>` section listing the events, and a `<Constraints>` section listing the constraints. Named sets of times, resources, and events may be defined.

An event represents a meeting of arbitrary but fixed duration between any number of resources, beginning at a particular time. It contains one time variable and any number of resource variables, each of which may either be preassigned or left for solutions to assign, subject to constraints.

Each constraint contains its type (one of the 15 types mentioned above), whether it is a hard or soft constraint, its weight, the resources, events, or sets of events that it applies to, and possibly other parameters specific to the constraint type. For example, there is an `<AvoidResourceClashes>` constraint which would typically state that it applies to all resources, and whose meaning, defined implicitly, is the usual one that those resources should not have timetable clashes. If they do, the constraint indicates the penalty for each occurrence. Other constraints require sets of events to have the same starting time, impose domain constraints on time and resource variables (e.g. requiring staff members to be suitably qualified), and so on. Everything affecting the evaluation of solutions appears explicitly as a constraint, allowing solutions to be evaluated unambiguously; this is the function performed by the web service (Kingston 2009).

A solution is a simple collection of assignments to the time and resource variables of the events of the corresponding instance.

The basic concepts of times, resources, events, and constraints seem applicable to other sub-disciplines, but the constraint types naturally vary. A recent comprehensive formulation of the curriculum-based university course timetabling problem (Cesco et al. 2008) is very similar to high school timetabling, the main addition being a limit on travel time between consecutive classes. Examination timetabling has times, resources (students and rooms), and events (examinations); its constraint types include constraints that limit the number of examinations that a student is required to attend over a short period of time. In the personnel rostering format (Curtois 2009) mentioned earlier, 'shifts' are times, 'shift types' are sets of times, 'employees' are resources, and their 'skills' are sets of resources; there are complex constraints on the timetable of each resource (at most one night shift per week, for example). Sports scheduling has times, resources (teams and venues), and events (matches); its constraints include the cost of travel between venues.

Despite some common structure, unifying all these problems into one is not advocated. It would be difficult in practice and would yield no practical benefit.

References

Burke EK, Kingston JH, Pepper PA (1998) A standard data format for timetabling instances. In: Practice and Theory of Automated Timetabling II, Burke EK and Carter M (Eds.), Springer Verlag Lecture Notes in Computer Science 1408:213–222

Burke EK, De Causmaecker P, Vanden Berghe G, Van Landeghem H (2004) The state of the art of nurse rostering. Journal of Scheduling 7:441–499

Burke EK, McCollum B, McMullan P, Qu R (2006) Examination timetabling: a new formulation. In: Proceedings of the Sixth International Conference on the Practice and Theory of Automated Timetabling (PATAT 2006), Brno, 373–375

Carter M, Laporte G, Lee ST (1996) Examination timetabling: algorithmic strategies and applications. Journal of the Operational Research Society 47:373–383

Cesco F, Di Gaspero L, Schaerf A (2008) Benchmarking curriculum-based course timetabling: formulations, data formats, instances, validation, and results. In: The 7th International Conference on the Practice and Theory of Automated Timetabling (PATAT2008), Montreal

Chand A (2004) A constraint based generic model for representing complete university timetabling data. In: Proceedings of the Fifth International Conference on the Practice and Theory of Automated Timetabling (PATAT 2004), Pittsburgh, 125–148

Cheang B, Li H, Lim A, Rodrigues B (2003) Nurse rostering problems: a bibliographic survey. European Journal of Operational Research 151:447–460

Cumming A, Paechter B (1995) Standard formats for timetabling data, Unpublished discussion session at the First International Conference on the Practice and Theory of Automated Timetabling, Edinburgh

Curtois T (2009), Personnel scheduling data sets and benchmarks, `http://www.cs.nott.ac.uk/~tec/NRP/`. Accessed 2009

Custers N, De Causmaecker P, Demeester P, Vanden Berghe G (2005) Semantic components for timetabling. In: Practice and Theory of Automated Timetabling V, Burke EK and Trick M (Eds.), Springer Verlag Lecture Notes in Computer Science 3616:17–33

De Causmaecker P, Demeester P, De Pauw-Waterschoot P, Vanden Berghe G (2000) Ontology for timetabling. In: Proceedings of the Third International Conference on the Practice and Theory of Automated Timetabling (PATAT 2000), Konstanz, 481–482

De Causmaecker P, Demeester P, Lu Y, Vanden Berghe G (2002) Using web standards for timetabling. In: Proceedings of the Fourth International Conference on the Practice and Theory of Automated Timetabling (PATAT 2002), Gent, 238–257

Easton K, Nemhauser GL, Trick MA (2001) The travelling tournament problem: description and benchmarks. In: Principles and Practice of Constraint Programming (CP 2001), Springer Verlag Lecture Notes in Computer Science 2239:580–585

Gröbner M, Wilke P, Büttcher S (2003) A standard framework for timetabling problems. In: Practice and Theory of Automated Timetabling IV, Burke EK and De Causmaecker P (Eds.), Springer Verlag Lecture Notes in Computer Science 2740:24–38

Kingston JH (2001) Modelling timetabling problems with STTL. In: Practice and Theory of Automated Timetabling III, Burke EK and Erben W (Eds), Springer Verlag Lecture Notes in Computer Science 2079:309–321

Kingston JH (2009) The HSEval high school timetable evaluator, `http://www.it.usyd.edu/au/~jeff`.

Kitagawa F, Ikeda H (1988) An existential problem of a weight-controlled subset and its application to school timetable construction. Discrete Mathematics 72:195–211

Da Mata JM, de Senna AL, de Andrade MA (1997) Towards a language for the specification of timetabling problems. In: Proceedings of the Second International Conference on the Practice and Theory of Automated Timetabling (PATAT'97), Toronto, 330–333

Özcan E (2003) Towards an XML-based standard for timetabling problems: TTML. In: Multidisciplinary Scheduling: Theory and Applications (First International Conference, MISTA '03, Nottingham, Selected Papers), 163–185

Paechter B (2003) International timetabling competition. `http://www.idsia.ch/Files/ttcomp2002/`

Post G, Ahmadi S, Daskalaki S, Kingston J et al. (2008) An XML Format for benchmarks in high school timetabling. In: The 7th International Conference on the Practice and Theory of Automated Timetabling (PATAT2008), Montreal

Post G (2009) Home Page, `http://wwwhome.math.utwente.nl/postgf/`

Ranson D, Ahmadi S (2006) An extensible modelling framework for the examination timetabling problem. In Practice and Theory of Automated Timetabling VI, Burke EK and Rudová H (Eds.), Springer Verlag Lecture Notes in Computer Science 3867:383–393

Reis LP, Oliviera E (2001) A language for specifying complete timetabling problems. In: Practice and Theory of Automated Timetabling III, Burke EK and Erben W (Eds), Springer Verlag Lecture Notes in Computer Science 2079:322–341