A Tabu Search Algorithm for Solving the Timetabling-Problem for German Primary Schools (Abstract)

Thorsten Desef[1], Andreas Bortfeldt[1] and Hermann Gehring[1]

[1] University of Hagen, Germany
andreas.bortfeldt@fernuni-hagen.de

1 Introduction

This contribution deals with the generation of timetables for German primary schools. Clearly, the underlying combinatorial task belongs to the group of educational timetabling problems. Moreover, it falls into the category of school timetabling (also known as the class/teacher model), since it shows the characteristic features of this class. There exist disjoint and stable classes and all classes at the same grade level have to complete essentially the same teaching program. A weekly schedule is required for all the classes of a school that avoids a teacher meeting two classes at the same time, and vice versa (cf. SCHAERF 1999).

School timetabling problems are generally NP-hard if formulated as optimization problems (cf. EVAN et al. 1976, SCHAERF 1999). Therefore, most of the solution methods proposed in recent years are based on metaheuristic approaches such as genetic algorithms (e.g., DREXL and SALEWSKI 1997), simulated annealing (e.g., ABRAMSON 1991) or tabu search (e.g., SCHAERF 1996). Furthermore, constraint based approaches (e.g., MARTE 2002) as well as other heuristic procedures have been presented. Since the regulations of the educational systems vary from country to country and since there are different school types, e.g., primary and secondary schools, the school timetabling problems tackled in the literature differ considerably. In this work it is intended to propose a method for solving the timetabling problem at German primary schools (GPS-TTP) that covers most of the practical requirements of this problem type.

2 Description of the GPS-TTP

At first the GPS-TTP is described informally. The description is based on interviews with headmasters/headmistresses of various primary schools. Prior to the definition of the relevant constraints some basic facts and concepts are presented:

- There are four grade levels (1 to 4) with nearly uniform teaching programs per level (set of subjects, number of weekly lessons per subject).
- For a given class, a teacher may be responsible for more than one subject, while there is only one teacher per subject and class. The assignment of teachers to sub-

jects and classes can be viewed as a separate task rather than an integrated part of the timetabling.
- Each class has a dedicated class room. Room conflicts may occur only for special rooms such as sports halls. Moreover, there is generally no room selection problem.
- Most of the lessons can be represented as combinations of one subject, one class, one teacher and one room (hereafter called lesson units).
- A school week consists of five or six days and the number of available hours per day is a constant. A day and one of its hours form a period (or time slot).
- The scheduling requires that one and only one period is assigned to each of the predefined lesson units; furthermore specific constraints are to be met.
- The following constraints are arranged according to the classification from MARTE (2002).

C1 Clash constraints

1.1-3 No clashes with regard to classes, teachers and rooms, respectively, are allowed.

C2 Availability constraints

2.1-2 The time windows of teachers and (special) rooms, respectively, are to be met.

C3 Coupling constraints

3.1 Some lesson units, together involving several classes, teachers and rooms, are to be held at the same time. This requirement occurs, e.g., in religious education. Pupils of different classes are grouped according to their religious denomination and lessons for the different pupil groups take place simultaneously.

3.2 Some lesson units concern only one class but require two teachers and possibly two rooms. This situation happens, e.g., if the weaker pupils of a class receive special teaching while a "normal" lesson takes place for the remaining part of the class.

3.3 There are some subjects, e.g. craft lessons, for which 2-hour lessons are preferred, i.e. two lesson units of the same type are to be scheduled one after another.

C4 Distribution constraints

4.1 Daily teaching for all of the classes is required.

4.2 The daily lessons of each of the classes have to begin in the first period and the follo-wing three hours should also be occupied by lessons.

4.3 No idle periods between lessons are allowed for any class.

4.4 The daily lessons of each class should be finished as early as possible, i.e. all the lessons of a class should be evenly distributed over the days of the week.

Further distribution constraints concern, e.g., the avoidance of idle periods from a teacher's perspective and favoured daily sequences of lessons according to the sub-

jects. Apart from (4.4) all of the numbered constraints are regarded as hard in practice.

3 A model for the GPS-TTP

The modelling of the GPS-TTP leads to a binary optimization problem that then serves as a base for the development of the solution procedure. The model is developed in three steps.

At first the constraints to be considered are selected. Each of the numbered constraints listed above is included, while the further distribution constraints of lower importance are neglected. The distribution constraints (4.1) to (4.4) are handled as soft constraints, while the other constraints remain hard constraints in the model.

In the second step, the key concept of a complex lesson unit (CLU) is introduced. A CLU aggregates all lessons that are forced by the coupling constraints (3.1) and (3.2) to happen at the same time. In the most simple case a CLU corresponds to a normal lesson unit. A CLU can occupy either one time slot or two consecutive time slots if this is required by coupling constraint (3.3). Moreover, each CLU is assigned a time window that is composed of the time windows of all teachers and rooms which are involved in the CLU. It is assumed from now on that all weekly lessons are organized as a set of CLUs; the number of CLUs is called *nl*. There are two main effects of this organization. On the one hand, the coupling constraints (3.1) to (3.3) are satisfied automatically. On the other hand, only one availability constraint for CLUs has to be met instead of two separate availability constraints for teachers and rooms.

In the third step the model for the GPS-TTP is formulated in terms of the CLU set. A timetable is expressed by means of binary decision variables x_{ldh} ($x_{ldh} \in \{0,1\}$); x_{ldh} is equal to 1 if and only if CLU l is scheduled at day d and hour h. The model includes five (types of) conditions. Three of them relate to the clash constraints (1.1) to (1.3). The remaining conditions guarantee that the availability constraint for CLUs is met and that a single period is assigned to each CLU. Of course, a solution is called feasible only if it satisfies all conditions of the model.

Two objective criteria, called f_{IP} and f_{EF}, are defined. The criterion f_{IP} counts the number of idle periods for all the classes while the criterion f_{EF} measures the extent to which the daily lessons are finished as early as possible. To this end, f_{EF} is formulated, roughly spoken, as the sum of the variables x_{ldh} which are weighted by the hour indices h. Obviously, the criteria f_{IP} and f_{EF} correspond to the distribution constraints (4.3) and (4.4), respectively. But they also cover the constraints (4.1) and (4.2), since, e.g., idle days lead to higher values for criterion f_{EF} for a given number of lessons. Note that the criteria f_{IP} and f_{EF} are not fully correlated. If, e.g., a solution does not include idle periods and, on the other hand, the lessons of the classes are very unevenly distributed over the weekly days, the solution quality is rather differently judged by each of the criteria. Finally, two objective functions, f_1 and f_2, are introduced both combining the criteria f_{IP} and f_{EF} in a complementary manner. Within function f_1 the criterion f_{IP} is of higher priority while f_{EF} serves as the tiebreaker. The priorities are swapped around within the objective function f_2.

4 Design of a tabu search algorithm for the GPS-TTP

The tabu search algorithm (TSA) can be characterized by the following main properties:

- Both objective functions, f_1 and f_2, are minimized simultaneously, i.e. a separate best solution is determined for each of them. The best solutions form a so-called Pareto-front of non-dominated solutions with regard to the criteria f_{IP} and f_{EF} (cf. CARRASCO and PATO 2001).
- Generated solutions are represented by period vectors. The nl CLUs are arranged in a fixed order. A period vector consists of nl periods and the l-th period within the vector is assigned to the l-th CLU. The periods of two CLUs may be equal. The search is carried out in the space of all period vectors. Thus, the search space contains feasible solutions as well as infeasible solutions.
- A starting solution of reasonable quality is built by means of a special heuristic. At first the CLUs are sorted by the level of their difficulty. The more a CLU is constrained the higher its level of difficulty. The CLU list is then looked through one time per class. For each CLU including a given class and not scheduled so far, the earliest hour followed by the earliest day of the week are specified for which the CLU can be scheduled in a feasible way.
- Two types of neighborhoods (N1 and N2) are applied. The N1-neighborhood is well known from the literature (cf., e.g., ABRAMSON 1991). All neighbors of a current solution are gene-rated by shifting one CLU from its old period to another one. The neighbors of a N2-neighborhood are calculated by the same move type. However, only CLUs that violate the conditions of the model are allowed to be moved. The neighborhood type used within an iteration depends on the level of violations of the conditions caused by the current solution.
- The best neighbor per iteration is determined with the help of a special cost function that is stipulated as a weighted sum. Each of its parts relates to a different type of possible conflicts in a solution or to an objective criterion, respectively. In order to steer the search, the weights of the sum are dynamically changed. The neighbor with the minimal cost is regarded as the best one. A random selection acts as the tiebreaker.
- Two different tabu lists are employed. The move list includes the (inverse) moves having been performed recently. The frequency list contains the shift-frequencies of the CLUs. A move (and its corresponding neighbor) is excluded from the neighbor contest, if it proves to be an element of the move list or if the frequency of the related CLU exceeds a frequency threshold. Exceptions are made if the current best values of the objective functions are improved (aspiration by objective). Further features concern the periodic reinitialization of each of the tabu lists and the periodic random change of the move list's length.
- Some further details should be mentioned regarding the multi-objective search. At the beginning, each of the both best solutions (see above) is initialized by the starting solution. After a neighborhood has been examined both best solutions are updated by the best neighbor if necessary. The case may occur that, e.g., the best solution w. r. t. f_1 has to be changed and the best solution w.r.t. f_2 has not.

5 Numerical results

45 instances of the GPS-TTP were tested on a 2.5 GHz PC. Some instances stem from practice and the other instances were derived from them by introducing tighter constraints and/or changing some minor details. The number of lessons to be scheduled ranges from about 150 up to about 500. The results of the calculation can be summarized as follows:

- For three instances from practice, results gained by human experts are known. These could be improved or at least be equalled.
- For each of the 45 instances feasible solutions could be found that do not violate the hard constraints.
- For 80% of the instances, solutions without idle periods were generated.
- With regard to the other objective criterion f_{EF}, near optimal values (in terms of a lower bound) have been obtained.
- The calculation times do not exceed 5 minutes for the largest instances with 532 lessons.

All in all, the heuristic seems to be able to cover practical requirements in terms of the solution quality as well as with regard to the calculation time.

References

ABRAMSON, D. (1991). Constructing School Timetables Using Simulated Annealing: Sequential and Parallel Algorithms. Management Science, Vol. 37(1), pp. 98–113.

CARRASCO, M. P. and PATO, M. V. (2001). A Multiobjective Genetic Algorithm for the Class/Teacher Timetabling Problem. In: Burke, E. and Erben, W. (Eds.), Practice and Theory of Automated Timetabling III, LNCS 2079, Springer, pp. 3–17.

DREXL, A. and SALEWSKI, F. (1997). Distribution Requirements and Compactness Constraints in School Timetabling. European Journal of Operational Research, Vol. 102(1), pp. 193–214.

EVEN, S., ITAI, A. and SHAMIR, A. (1976). On the Complexity of Timetabling and Multicommodity Flow Problems. SIAM Journal of Computation, Vol. 5(4), pp. 691–703.

MARTE, M. (2002). Models and Algorithms for School Timetabling – A Constraint-Programming Approach. Dissertation an der Fakultät für Mathematik, Informatik und Statistik der Ludwig-Maximilians-Universität München, Germany.

SCHAERF, A. (1996). Tabu Search Techniques for Large High-School Timetabling Problems. In: Proceedings of the 13th National Conference on Artificial Intelligence (AAAI-96), pp. 363–368. Portland, USA: AAAI Press/MIT Press.

SCHAERF, A. (1999). A Survey of Automated Timetabling. Artificial Intelligence Review, Vol. 13, pp. 87–127.