School Timetable Creation with the Use of Integer Linear Programming

Vít Fanta
Thursday, 9:15
Cybernetics and Robotics
fantavit@fel.cvut.cz

This report aims to describe the solution of the semestral project for the Combinatorial Algorithms (RM35KOA) course. The topic was chosen by the author and was approved by the lab teacher.

I. Assignment

In the present day, numerous software solutions exist to provide schools of all levels of the education system to address the issue of timetabling, e.g. [1], [2] or [3]. The last entry in particular, [3] is popular in Czechia due to its interface's localization in Czech, and the ability to integrate other essential parts of administration software a school may need. However, the author witnessed that the timetabling software [3] was unable to find a solution which would be even remotely usable, all despite the claims in the company's videos. Therefore, the task often has to be performed manually, although some tool-assistance is provided by the software (preventing timetable collisions). Such attitude is also employed at small educational institutions.

The task at hand is to implement an Integer Linear Programming (ILP) solution for the problem of timetable creation for a fictional primary school. The solution should take various spatial (e.g. limited amount of classrooms) and time (e.g. ensure that teachers do not have a collision in their timetable) limitations into account, and should be easily extendable should any additional constraints arise. The problem is NP-hard, because there exists a polynomial reduction to ILP formulation.

II. Related works

The first experiments consisting of using ILP formulations to solve timetabling problems dates to the 1960s, [4]. Although the solution was functional, the article mentions computing limitations and resolves to several simplifications such as omitting some classes from the problem formulation. Recently, an ILP solver for a generalized timetabling problem was proposed in [5]. Often, a multiple-phase optimization is performed to achieve results in less time-consuming manner (as in this report). One example of such attitude is presented in [6] where a genetic algorithm is employed to mutate an already found solution in order to obtain another one which is better with respect to an objective function. This objective function can be of arbitrary form, because the evolution step is not part of the ILP. Despite linear programming methods, constraint programming approaches can also be used to tackle the timetabling problem. One such strategy is presented in [7], although there are several simplifications compared to the formulation presented in this report (lack of curriculum to be fulfilled, lack of room assignment). Nevertheless, constrained programming can be used to formulate the timetabling task, maybe in a more straightforward way than the one presented below, due to its ability to express and handle more complex constraints.

III. Problem solution

A. Design

The proposed solution consists of two subsequent ILP optimization problems, with the second ILP directly using the results of the previous level of optimization. The detailed description of each of the two phases is given below. For the sake of clarity, let the following notation elements be described:

$p \in \{1, 2, \ldots, P\}$ This set of indices enumerates individual subjects (*předměty*) taught at the school.
$u \in \{1, 2, \ldots, U\}$ This set of indices enumerates individual teachers (*učitelé*) teaching at the school.
$t \in \{1, 2, \ldots, T\}$ This set of indices enumerates individual classes (*třídy*) of pupils.
$d \in \{1, 2, 3, 4, 5\}$ This set of indices enumerates weekdays (*dny*).
$h \in \{1, 2, \ldots, H\}$ This set of indices enumerates lessons (*hodiny*) throughout a day.
$m \in \{1, 2, \ldots, M\}$ This set of indices enumerates classrooms (*místnosti*) available in the school building(s).
$h_O \in \{o_1, o_2, \ldots, o_k\}$ This set of indices enumerates lessons during which pupils can have lunch (*oběd*) in the canteen.

1) First Phase: The goal of the first phase ILP is to assign individual teachers to classes and subjects in such a way that the curriculum of each of the classes is fulfilled (i.e. no subject is left without a teacher), each teacher is assigned a total number of lessons in a reasonable interval (this is discussed later), and ensure that each teacher is qualified to teach all of the subject he has been assigned (i.e. approbation).

The implementation uses the following given parameters:

$$\mathbb{N}_0 \ni \text{Curriculum}_{t,p} = \text{the number of lessons of subject } p \text{ required for class } t, \tag{1}$$

$$\{0,1\} \ni \text{Approbation}_{u,p} = \begin{cases} 1, & \text{if the teacher } u \text{ is qualified to teach subject } p, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

$$\mathbb{N} \ni \text{Upper}_u = \text{the upper bound on the number of lessons teacher } u \text{ can teach,} \tag{3}$$

$$\mathbb{N} \ni \text{Lower}_u = \text{the lower bound on the number of lessons teacher } u \text{ can teach.} \tag{4}$$

The variables are

$$\{0,1\} \ni \text{Teaches}_{u,t,p} = \begin{cases} 1, & \text{if the teacher } u \text{ is assigned to teach subject } p \text{ in class } t, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

$$\mathbb{N} \ni \text{Lessons}_u = \text{the total number of lessons the teacher } u \text{ teaches.} \tag{6}$$

The optimization is ran without a criterion (any feasible solution is admitted), but is subjected to the following constraints:

$$\forall u : \sum_{t,p} \text{Curriculum}_{t,p} \cdot \text{Teaches}_{u,t,p} = \text{Lessons}_u \qquad \text{Bind the variables together.} \tag{7}$$

$$\forall u : \text{Lower}_u \leq \text{Lessons}_u \leq \text{Upper}_u \qquad \text{Apply the bounds.} \tag{8}$$

$$\forall u \forall t \forall p : \text{Approbation}_{u,p} \geq \text{Teaches}_{u,t,p} \qquad \text{Ensure the proper qualification.} \tag{9}$$

$$\text{If } \text{Curriculum}_{t,p} > 0 : \sum_{u} \text{Teaches}_{u,t,p} = 1 \qquad \text{Ensure that a teacher is assigned to every taught class and subject.} \tag{10}$$

2) Second Phase: The second phase uses the precomputed result of the first phase to put all lessons in an spatio-temporal slot. The constraints are specified to avoid any collisions of teachers' timetable (a teacher teaching two lessons simultaneously), classes' timetable (each class has at most one lesson every hour), classrooms' timetable (there can only be at most a single lesson taught in a classroom), while also taking notice of less obvious constraints, such as limited capacity of the school canteen.

The variables of the second phase model are:

$$\{0,1\} \ni \text{Timetable}_{t,p,m,d,h} = \begin{cases} 1, & \text{if the teacher } u \text{ teaches subject } p \text{ for class } t \text{ on day } d \text{ during lesson } h, \\ 0, & \text{otherwise,} \end{cases} \tag{11}$$

$$\{0,1\} \ni \text{Lunches}_{t,d,h} = \begin{cases} 1, & \text{if the class } t \text{ has lunch during lesson } h \text{ on day } d, \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

$$\{0,1\} \ni \text{Afternoon}_{t,d} = \begin{cases} 1, & \text{if class } t \text{ has afternoon education (lessons after lunch) on day } d, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The optimization task is of the form

$$\min \sum_{t,p,m,d,h} h \cdot \text{Timetable}_{t,p,m,d,h}, \tag{14}$$

subject to the following constraints:

$$\forall m \forall d \forall h : \sum_{t,p} \text{Timetable}_{t,p,m,d,h} \leq 1 \qquad \text{Each room can host at most a single lesson.} \tag{15}$$

$$\forall d \forall h \forall u \forall u', u \neq u', \text{whenever } \text{Teaches}_{p,u,t} = 1 \text{ and } \text{Teaches}_{p',u,t'} = 1 :$$

$$\left(\sum_{m} \text{Timetable}_{t,p,m,d,h} = 1 \right) \gg \left(\sum_{m} \text{Timetable}_{t',p',m,d,h} = 0 \right)$$

$$\text{Ensure no time collisions for teachers.} \tag{16}$$

$$\forall t \forall p : \sum_{m,d,h} \mathrm{Timetable}_{t,p,m,d,h} = \mathrm{Curriculum}_{t,p} \qquad \text{Adhere to the curriculum.} \tag{17}$$

$$\forall t \forall d : \sum_{m,d} \mathrm{Timetable}_{t,p,m,d,1} = 1 \qquad \text{The education always starts in the morning.} \tag{18}$$

$$\forall t \forall d \forall h : \sum_{p,m} \mathrm{Timetable}_{t,p,m,d,h} \leq 1 \qquad \text{Avoid time conflicts for pupils.} \tag{19}$$

$$\forall t \forall d : \sum_{h} \mathrm{Lunches}_{t,d,h} = 1 \qquad \text{One lunch per day.} \tag{20}$$

$$h \notin h_O \forall t \forall d : \mathrm{Lunches}_{t,d,h} = 0 \qquad \text{Specify which lessons are available for lunch.} \tag{21}$$

$$\forall t \forall d \forall h_O : (\mathrm{Lunches}_{t,d,h_O} = 1) \gg \left(\sum_{p,m} \mathrm{Timetable}_{t,p,m,d,h_O} = 0 \right) \qquad \text{No lesson and lunch simultaneously.} \tag{22}$$

$$\forall d \forall h_O : \sum_{t} \mathrm{Lunches}_{t,d,h_O} \leq 3 \qquad \text{Limited canteen capacity.} \tag{23}$$

B. Implementation Details

A careful reader could notice the strange syntax of constraints (16) and (22). This is abuse of Gurobi's indicator constraint, as stated in [8]. The constraint in right set of parentheses is active only if the condition in the left parentheses is fulfilled. However, the constraints cannot be arbitrary. The left constraint must contain a binary comparison with a single binary variable and the right constraint must be linear. To adhere to these requirements, a variable

$$\{0,1\} \ni \mathrm{mSumsTimetable}_{t,p,d,h} = \sum_{m} \mathrm{Timetable}_{t,p,m,d,h} \tag{24}$$

was created to make formulation of (16) possible.

A caveat is hidden in the definition of the model variable $\mathrm{Afternoon}_{t,d}$, (13). The problem is nor the education timetable nor the time of the lunch break are known in advance which makes the definition of this variable more complicated. Two more supplementary variables were created:

$$\{0,1\} \ni \mathrm{LunchesHelper}_{t,d,h} = \begin{cases} 1, & \text{if the class } t \text{ had lunch at } h \text{ or before on day } d, \\ 0, & \text{otherwise,} \end{cases} \tag{25}$$

$$\mathbb{N}_0 \ni \mathrm{AfternoonSums}_{t,d} = \text{number of lessons taken after lunch by class } t \text{ on day } d. \tag{26}$$

These were bound to the existing variables through the following constraints:

$$h = 2, \ldots, H, \forall t \forall d : \mathrm{LunchesHelper}_{t,d,h-1} \leq \mathrm{LunchesHelper}_{t,d,h}, \tag{27}$$

$$h = 2, \ldots, H, \forall t \forall d : \left(\mathrm{LunchesHelper}_{t,d,h-1} = 0\right) \gg \left(\mathrm{Lunches}_{t,d,h} = \mathrm{LunchesHelper}_{t,d,h}\right), \tag{28}$$

$$\forall t \forall d : (\mathrm{Afternoon}_{t,d} = 0) \gg (\mathrm{AfternoonSums}_{t,d} = 0), \tag{29}$$

$$\forall t \forall d : (\mathrm{Afternoon}_{t,d} = 1) \gg (\mathrm{AfternoonSums}_{t,d} \geq 1), \tag{30}$$

$$\forall t \forall d : \sum_{p,m,h} \mathrm{Timetable}_{t,p,m,d,h} \cdot \mathrm{LunchesHelper}_{t,d,h} = \mathrm{AfternoonSums}_{t,d}. \tag{31}$$

Although the constraint (31) is quadratic, the Gurobi solver is able to transform it to SOS (special ordered set) constraints and solve the complete system using linear methods [9] which is confirmed by the solver's output during the optimization process.

IV. TESTING

Suitable testing data were found in the curriculum specification for the primary school Lesní in Liberec which is available online at [10], [11]. The fictional school consists of 9 classes, each in different grade. At first, the approbation parameter was ignored (all of the teachers could teach all of the subjects) and their number was set to 11. This was done in order to properly test the functionality of the two optimization tasks without encountering potential infeasibility due to lack of qualified teachers in certain subjects.

The resulting timetable of this simplified task was then consulted with a real life teacher. As a consequence, a number of additional requirements have emerged.

- Until grade 4, all of the subjects, except for foreign languages, are taught by a single teacher (home-room teacher). This was implemented by excluding first four teachers from the first phase ILP and directly assigning them to their respective class.

- In some cases, the optimization would result in poor distribution of lessons throughout the week for a particular subject. Therefore, two additional constraints were added to the second phase ILP:

$$\text{If Curriculum}_{t,p} \leq 5, \forall t \forall p \forall d : \sum_{m,h} \text{Timetable}_{t,p,m,d,h} \leq 1, \tag{32}$$

$$\text{If Curriculum}_{t,p} \geq 6, \forall t \forall p \forall d : \sum_{m,h} \text{Timetable}_{t,p,m,d,h} \geq 1. \tag{33}$$

- A class should not have more than 3 afternoon education during the week:

$$\forall t : \sum_{d} \text{Afternoon}_{t,d} \leq 3. \tag{34}$$

- It is an unwritten rule that no afternoon education should take place on Friday:

$$\sum_{t} \text{Afternoon}_{t,5} = 0. \tag{35}$$

- Certain lessons can only be taught in certain classrooms, e.g. physical education. Let PE denote the subject index of physical education and A the room index of the gymnasium:

$$m \neq \text{A} : \sum_{t,d,h} \text{Timetable}_{t,\text{PE},m,d,h} = 0. \tag{36}$$

After these alterations, the teachers' approbations were finally specified as seen in table IV. The teaching hours vary between 14 hours to 22 hours per week, where the 22 hour teaching schedule corresponds to a full-time employed teacher. This spread is justified as often not all teachers at a school take on this amount of workload. The approbations were selected iteratively in order to satisfy the curriculum requirements for all classes.

Teacher designation	Approbation
T5	Physics, chemistry, math, elective
T6	Geology, biology
T7	English, Czech
T8	Math, civics
T9	Music, English
T10	History, civics
T11	Czech
T12	Czech, German
T13	Arts, crafts, elective, vocation
T14	IT, PE, health

TABLE I. APPROBATIONS OF INDIVIDUAL TEACHERS. TEACHERS 1 TO 4 ARE OMITTED BECAUSE THEY ARE HOME-ROOM TEACHERS FOR THE FIRST 4 GRADES.

The whole optimization process was ran on a laptop equipped with Intel's i7-9750H 6-core CPU clocked at $2.6\,\text{GHz}$ and $16\,\text{GB}$ of RAM. The total time of computation is approximately $35\,\text{s}$ (averaged over 5 runs) of which the first phase optimization lasted $1\,\text{s}$. Half of the computation time is spent on obtaining the certificate of optimality and no further improvements are found. It is also worth noting that before the inclusion of the additional constraints specified in this section, the whole optimization process was completed under $3\,\text{s}$.

A. Results

The resulting computed timetable for the nine classes is presented in figure 1. Because of the model's nature to implement all constraints as hard constraints, all of the desired requirements are met. The colouring is designed to improve readability: There are different colours for natural sciences, social sciences, languages, physical education and the most dominant subject (Czech and math) have their own colour. The subject name is specified in the top of each tile, the teacher index in the bottom right, and the room number in the bottom left. It can be seen that the first four grades are taught all subjects by their home-room teacher, except for the language lessons.

V. CONCLUSION

In this report, an integer linear programming solver for a timetabling problem for a primary school with authentic curriculum and 9 classes has been presented. The ILP formulation consists of several multidimensional integer (mostly binary) variables and a great number of constraints ensure generation of a usable timetable. The set of constraints on the defined variables can be easily augmented to adhere to other requirements, should any arise.

The main drawback of the presented model is its inability to allow divisions of a class to several sections. This is a necessary feature to allow construction of timetables for schools in higher levels of education. Its incorporation could be added via

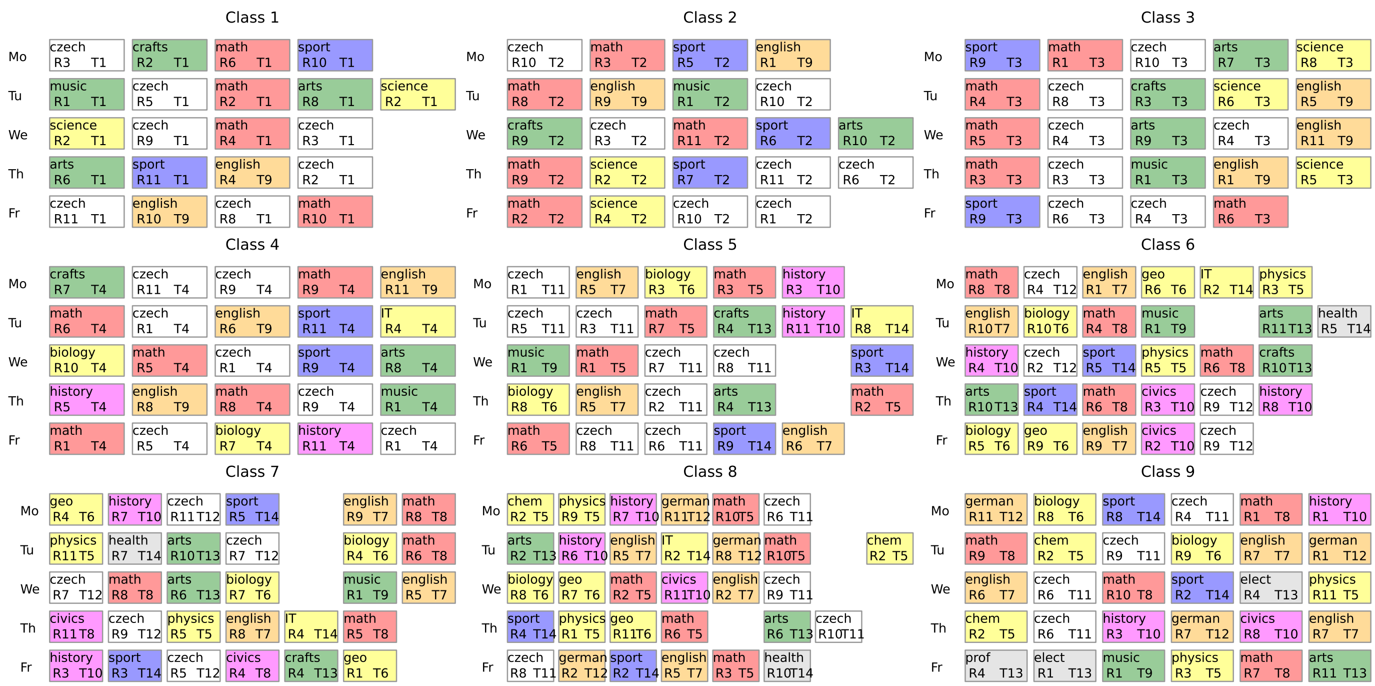

Fig. 1. The resulting timetable computed by the 2-phase ILP.

considering the individual class sections instead of the whole classes and subject these variables to additional constraints. Further modifications could allow the model to handle any input of the XHSTT format which is a standardized format for general timetabling problems, [12].

Finally, I would like to thank Mgr. Jan Voženílek, my mom (a teacher) and Ing. Antonín Novák for consultations.

References

[1] "TimeTabler — Easy and Efficient School Timetabling Software," https://www.timetabler.com/.

[2] "aSc Rozvrhy," https://www.asctimetables.com/.

[3] "Bakaláři – mezi Školou a rodinou — Bakaláři," https://www.bakalari.cz/Static/rozvrh.

[4] N. L. Lawrie, "An integer linear programming model of a school timetabling problem," *The Computer Journal*, vol. 12, no. 4, pp. 307–316, Apr. 1969.

[5] S. Kristiansen, M. Sørensen, and T. R. Stidsen, "Integer programming for the generalized high school timetabling problem," *Journal of Scheduling*, vol. 18, no. 4, pp. 377–392, Aug. 2015.

[6] S. Ribic and S. Konjicija, "Evolution strategy to make an objective function in two-phase ILP timetabling," in *2011 19thTelecommunications Forum (TELFOR) Proceedings of Papers*. Belgrade, Serbia: IEEE, Nov. 2011, pp. 1486–1489.

[7] C. Valouxis and E. Housos, "Constraint programming approach for school timetabling," *Computers & Operations Research*, vol. 30, no. 10, pp. 1555–1572, Sep. 2003.

[8] "TempConstr," https://www.gurobi.com/documentation/9.5/refman/py_tempconstr.html#pythonclass:TempConstr.

[9] "Is Gurobi able to handle a decision variable in the index/as a boundary in a sum? – Gurobi Support Portal," https://support.gurobi.com/hc/en-us/community/posts/360055448071-Is-Gurobi-able-to-handle-a-decision-variable-in-the-index-as-a-boundary-in-a-sum-.

[10] "Školní kurikulum ZŠ Lesní v Liberci," https://www.zslesni.cz/zakladni-skola-liberec-lesni-57512-prispevkova-organizace/kurikulum.html.

[11] M. Ulvrová, "Školní vzdělávací program: Škola pro život," 31\08\2021.

[12] G. Post, J. H. Kingston, S. Ahmadi, S. Daskalaki, C. Gogos, J. Kyngas, C. Nurmi, N. Musliu, N. Pillay, H. Santos, and A. Schaerf, "XHSTT: An XML archive for high school timetabling problems in different countries," *Annals of Operations Research*, vol. 218, no. 1, pp. 295–301, Jul. 2014.