Three-phase Curriculum-Based University Course Timetabling with Student Assignment

Elmar Steiner[1], Ulrich Pferschy[1], and Andrea Schaerf[2]

[1] Department of Operations and Information Systems, University of Graz, Austria
{elmar.steiner, ulrich.pferschy}@uni-graz.at
[2] Dipartimento Politecnico di Ingegneria e Architettura, University of Udine, Italy
andrea.schaerf@uniud.it

Abstract. We consider a complex university timetabling problem arising in a four-year study program of teacher education where every student has to choose two subjects. Since any combination of two subjects is feasible, the goal of designing a collision-free timetable for every student seems to be unreachable. However, the task becomes more tractable because for most courses several parallel groups are offered, i.e. sectioning of students is possible. Further difficulties arise from the highly individual progress of students who often follow neither the prescribed term of each course nor the prescribed ordering of courses. Under these and other conditions an optimized timetable should be determined and adjusted to the estimated student numbers and their past achievements. After moving main lectures into a regular time grid with minimal changes concerning the previously existing plan, the task of finding a timetable for all lectures with parallel groups is modeled as an integer linear program (ILP). Later, students with their actual demands are allocated a non-overlapping set of courses that is relevant and feasible for their individual study situation. This part can be handled by an assignment-type model followed by a round-robin allocation of remaining capacities.

Keywords: Course Timetabling · Student Sectioning · ILP Model

1 Introduction

A timetabling problem generally consists of assigning a set of activities to resources such that a set of complex constraints is fulfilled, which varies depending on the given problem. Whereas these constraints are usually considered to be *hard*, desirable characteristics of the timetable are introduced as *soft constraints* into the objective function. The goal is to find a feasible assignment while minimizing the weighted sum of the penalties representing these violations.
There is a wide range of real-world applications, including *university timetabling*, where different categories of specific problems are distinguished: *Examination timetabling* (ETT), *post-enrollment course timetabling* (PE-CTT) and *curriculum-based course timetabling* (CB-CTT). Both of the latter two deal with the assignment of courses to time periods and usually rooms, however, there are certain

differences. In PE-CTT the timetable is established after the enrolment of students, thus taking into account that students are enrolled in various combinations of events and somehow incorporating these selections. CB-CTT determines a timetable based on a curriculum of study programs, that is, a list of courses to be taken by a group of students. A recent survey on all the available formulations of educational timetabling is provided by Ceschia et al. [4].
These timetabling problems form only one side of the issues that operations research has to address in education, see Johnes [8] for an overview. Real-world educational planning scenarios often simultaneously comprise components of various problems, depending on the stage of planning and area of application (such as elementary or tertiary education). Therefore the correct choice of methods depends on the planning characteristics (such as information availability or choice of planning entities) and combined approaches seem appropriate.
One of the additional issues is the group of *student sectioning problems* (see e.g. [12]), where students are assigned to particular sections of a course satisfying constraints such as room or section capacity and avoiding conflicts in students' timetables due to overlaps. Quite often this is considered a *sub-problem* of course timetabling. That is, after deriving an adequate timetable, one seeks an optimal assignment of students to classes avoiding conflicts and taking into account students' needs/requests and other soft constraints such as preferences or daily workload.
The timetabling problem we are dealing with is a complex scenario involving several non-standard properties. Its main decision problem can be categorized as a variant of a sectioning problem.

After giving a general description of the problem in Section 2 we point to some related literature in Section 3. The mathematical models introduced for solving our timetabling problem will be presented in Section 4. Our approach consists of three phases: Phase 1 shifts important "main lectures" from their historical starting times into the regular time grid which serves as a basis of all our plans. In Phase 2 a complex ILP model is set up which determines in one optimization step the time periods of all courses (many of them consisting of several parallel groups) and also assigns individual sets of relevant courses in collision-free time periods to groups of students with identical properties. Since this step has to be carried out many months before the start of an academic year, these groups of students are only estimations of future student demand. The final course assignment of students is done at a later time in Phase 3 by matching actual students to estimated groups. Both Phases 1 and 3 employ generalizations of the linear assignment problem with additional conflict constraints. Computational experiments in Section 5 illustrate the potential and limitations of the large ILP model.

2 Problem description

We were asked by the central administration of the University of Graz, Austria, to develop an automated solution approach for a complex timetabling task arising in the teacher education study program which involves roughly 4,200 students. From an educational planning perspective, it can be considered a *multi-phase scheduling problem*. More precisely it consists primarily of a *timetabling* step producing as output the day of the week and starting time of every course (which stays constant over the whole term). This step consists of two phases: In Phase 1, starting times of courses with a larger audience (and no parallel groups) are moved from their historical starting times into the regular time grid prescribed by the university. Since these moves cause major disturbances and negative side effects, their total time deviation will be minimized under non-collision constraints. In the second phase the starting times for all courses with parallel groups are computed from scratch (also obeying the regular time grid). To assure feasible timetables for individual students this phase is coupled with *the sectioning* of *projected students*. Later, in the third phase, the planning problem asks for an *assignment* of pre-registered actual students which assigns to each student a set of courses that are feasible, relevant, and non-overlapping for the respective student.

The reason for the coupling of methods as well as separation into distinct phases is the structure of the given planning process at our university. This procedure essentially covers the capacity planning of the number of sections and the establishment of yearly timetables for the teacher education study program. The study program of teacher education requires the choice of *two subjects* (such as English and chemistry), thus the planning entails the coordination of all involved departments and those of their provided courses, which are part of the program's curriculum.

While the definition of the number of sections is based on enrollment predictions and the establishment of the timetable (first and second phase) needs to be carried out in March for the two terms of the academic year to come (starting in October), the actual enrollments (required for the third phase) only come to be known in September (for the winter term) resp. February (for the summer term), when an adaption and especially a re-scheduling of courses is not permitted anymore. So there is a crucial *temporal* delay between planning and information receipt. Moreover, since departments currently schedule courses autonomously and only distinctive overlapping time conflicts (e.g. of prominent courses) are resolved bilaterally, the majority of students of any combination of subjects will face time conflicts in their weekly timetable. Nevertheless, the final schedule has been ascertained to be of major influence in regards to the study conditions and therefore students' performance in completing their studies.

To improve the situation of studying, we seek to support both planning processes by deriving models to optimize the weekly timetable using the respective information given during phases, resulting - opposed to conventional approaches - in a *combination* of sectioning and timetabling. While the true target is being free of conflicts, we also seek a compact timetable for the individual student and didactic practicability. The target groups for facilitation are especially those students who exhibit non-standard study progresses, such that courses are not completed in the term that is recommended in the curriculum. For that reason and as opposed to many curriculum-based models the method needs to comprise the capability to avoid time conflicts between courses that are not nominally taken in the same year. The resulting term-overlapping constraints are determined from given historical data that describes to which extent courses are completed earlier or (more likely) later than recommended. Courses that are prone to be taken later than in the prescribed term will be called *displaced courses*.

The input structure of the problem primarily consists of a set of courses. These consist on one hand of so-called *main courses* (lectures) C^m, which usually cater to a larger audience and do not have parallel groups. However, many of them have a historically established starting time, possibly outside the regular time grid. Then there are standard *courses* (exercise classes, seminars, etc.) C, where each course $c \in C$ has a limited capacity $cap(c)$ and therefore a certain number of *sections* (parallel groups) are offered. As stated the number of these groups is determined by a separate planning process (involving also financial considerations) and is provided as an input to the timetabling task. For each course resp. parallel-group a lecturer is given. Equivalently, for each lecturer $l \in L$, its set of courses $C(l)$ is known. Following the most common teaching mode, we assume that each course or group is given by exactly one lecturer, although team teaching or shared courses may well occur in practice. However, our model could be easily extended to accommodate more than one lecturer per course.

Each course is part of the curriculum of exactly one subject and is prescribed for exactly one term. Besides that, there are also some general main courses which are part of every curriculum (any combination of subjects) and have to be taken by all students.

Note that – different from many existing timetabling applications – rooms are not considered in our planning task. This is because rooms are shared with the programs of the other ca. 30,000 students of the university. Therefore, an automated allocation of *rooms* would have far-reaching consequences for the decentral planning process of the whole university. However, rooms currently do not pose a major problem to the planers because the majority of courses have either very specific requirements with regards to rooms (such as chemistry labs) or none at all. While the former use a room that is tightly coupled to the department and usually shared in a limited and well-practiced manner, the former can use any room on the campus.

3 Related work

In the literature, there exist several strategies of section management, depending on the concrete problem where they are applied. The problem has been tackled either as a separate problem or integrated into the timetabling procedure. Aubin and Ferland [1] for example iteratively adjusted both timetable and section assignment given an initial timetable. Banks et al. [3] propose a rather simultaneous approach, where they assign sections of courses to time periods and iteratively add constraints, each representing a student course selection, to satisfy as many choices as possible.

A very comprehensive approach is the one by Müller and Murray [12], who propose a multi-phase sectioning strategy, considering various information stages in educational planning and using different (heuristic) algorithms for each phase. In detail, they identify three different approaches to sectioning that they call *Initial sectioning*, *Batch sectioning*, and *Online sectioning*, which differ based on the time when it takes place and the information available at that time. The same approach has been further extended and specifically applied to a Faculty of Education by Müller and Rudová [13].

An integrated timetabling and sectioning problem has been recently proposed for the fourth International Timetabling Competition (ITC-2019) [14]. The ITC-2019 problem consists of sectioning students into classes based on course enrollments and then assigning classes to available periods and rooms. Courses may have a complex structure of classes, with one or more configurations, further divided into subparts and the parent-child relationship between classes. The other remarkable feature is that the timetable may differ from week to week, instead of replicating the same weekly timetable for the whole semester.

It is worth mentioning that for such a complex timetabling/sectioning problem as the ITC-2019 one, the solution techniques based on MIP solvers turned out to be very competitive. Indeed, the MIP formulation by Holm et al. [7] won the competition and produced the best solution for the majority of the instances. Most of the remaining best solutions have been obtained by a local search approach, which did not enter the competition as it was proposed by one of the organizers (i.e., Müller [11]).

Other successful applications of MIP models to timetabling problems are the works by Lach and Lübbecke [10] and by Bagger et al. [2], that worked on the CB-CTT problem obtaining both good solutions and tight lower bounds. Another complex, real-world sectioning problem has been proposed by Esmaeilbeigi et al. [6] for a military school. In their problem, a lesson has a multiphase structure, such that each phase may require different resources and is taken by different students.

Finally, complexity analysis of the student sectioning problem has been carried out by Dostert et al. [5] and Schindt [15], identifying the cases in which the problem is polynomial and those in which the problem is NP-hard.

4 The Mathematical Models

As pointed out in Section 2 our planning problem consists of three separate phases. In the following, we describe our solution approach for each of them. The aim of Phases 1 and 2 is a complete timetable for all courses.

Each course $a \in C^m \cup C$ belongs to exactly one subject $f(a)$ and is prescribed for a certain term $n(a)$ with $n(a) \in \{1, 2, \dots, 8\}$, corresponding to winter and summer semesters of a four year program. As an exception, there is a small subset of main courses which is prescribed for all subjects (educational theory, etc.).

Furthermore, each regular course $a \in C$ comprises $g(a) \geq 1$ sections (parallel groups). Thus, the timetable consists of a starting time for each main course $a \in C^m$ and for each section of every course $a \in C$.

4.1 Phase 1: Alignment of main lectures

The feasible time periods of the timetable consist of a day of the week and a starting time. The latter is set by university rules to a fixed time grid starting at 8:15 and continuing with a sequence of 90 minutes of lecture time and 15 minutes breaks. This allows for seven time slots per day, i.e. a set P consisting of 35 time periods per week as shown below. We only assign lectures of 90 minutes and exclude from consideration the small number of lectures with deviating duration.

start	Mon	Tue	Wed	Thu	Fri
08:15					
10:00					
11:45					
13:30					
15:15					
17:00					
18:45					

In Phase 1 the *main lectures* C^m (which do not have parallel sections) will be aligned to the given time grid. A majority of main lectures already follow this prescribed time frame, but a non-negligible minority deviates from the time grid. Since the main lectures in general hardly change their time and room over the years and some of them are also part of other curricula outside our planning task, it makes sense to change their starting times as little as possible. Therefore, we set their starting times by solving a version of the linear assignment problem with additional conflict constraints. Thereby we match main lectures to time periods $p \in P$ of the given time grid with the additional restrictions that main courses belonging to the same term $t \in T$ must not overlap. This non-collision condition is imposed independently from the subject since it should be possible to study any combination of subjects without overlaps in the main courses as there exists no alternative for them. The special main courses which are prescribed for all subjects cannot overlap with any other main course of the same term. Function $C^m(t)$ returns the set of courses that belong to term $t \in T$. Additionally, lecturers $l \in L$ cannot be assigned to more than one course at the same time, both in the winter and the summer semester. Function $C^m(l)$ then returns the set of courses that are given by lecturer l. The single binary decision variable $y_{cp} = 1$ if main course c is assigned to time period p, and 0 otherwise.
As a linear objective function, we consider the *distance* $\Delta(c, p)$ between the current time slot of course c (i.e. as in the previous year) and the new time period $p \in P$. If both times are on the same day, $\Delta(c, p)$ describes the absolute difference in minutes between the current starting time and the beginning of period p. Otherwise, i.e. if the course is moved to a different day, we assume a penalty

value ρ equal to three times the maximum intra-day distance (independently from the new day). The resulting assignment-type integer linear program is as follows:

$$\min_{y} \quad \sum_{c \in C^m} \sum_{p \in P} y_{cp} \cdot \Delta(c,p) \tag{1a}$$

$$\text{s.t.} \quad \sum_{p \in P} y_{cp} = 1, \quad \forall c \in C^m, \tag{1b}$$

$$\sum_{c \in C^m(t)} y_{cp} \leq 1, \quad \forall p \in P,\ t \in T, \tag{1c}$$

$$\sum_{c \in C^m(l)} y_{cp} \leq 1, \quad \forall p \in P,\ l \in L, \tag{1d}$$

$$y_{cp} \in \{0,1\} \tag{1e}$$

The results are listed in Table 1, together with the computation time. It turned out that in the winter and summer semester 69% and 47.4%, respectively, of all main lectures had to be adjusted. The assignment of the remaining courses in C^m already followed the time grid and was not changed. Notably, the solutions do not comprise any alignment to another working day. Furthermore, the amount of rescheduled lectures varies substantially for different terms, such that earlier terms show more displacements.

Table 1: Results of main lecture adjustments.

semester	$\|C^m\|$	adjusted courses	av. adj.	max. adj.
winter	71	49	1h48m	7h
summer	38	18	2h51m	7h

It should be noted that Phase 1 is relevant mostly for the introductory year of the new planning tool, or when additional subjects are integrated into the planning process. Once all main courses are aligned with the time grid, Phase 1 will be used only for assigning new main courses and for handling exceptions such as enforced changes.

4.2 Phase 2: Timetabling-Sectioning

Phase 2 is the most complex part. It considers the computation of starting times for all freely assignable courses, many of them being offered with parallel groups, which necessitates the sectioning of the estimated student cohorts.

The demand structure of the planning task is captured by sets of students each of them enrolled in two subjects $\{f_1, f_2\}$ (of equal importance), where each f_i is chosen arbitrarily from a set F of 28 offered subjects. In our Central European setting the progress of a student does not follow a strict yearly pattern, e.g. a student cannot be immediately identified as a third-year student, which is quite different from many international university systems. To place a student s in a certain term, we count the sum $ECTS_s$ of ECTS credits reached and assign the student to term $t_s = \lceil \frac{ECTS_s}{30} \rceil$, since 30 ECTS are the usual workload assigned for one term. Our experience tells us that most students proceed equally fast in their two subjects. Thus, we do not distinguish the progress in the two subjects. To scale the optimization problem of Phase 2 to a more tractable size, we combine sets of Δ (e.g. $\Delta = 5$) identical students, i.e. students with the same pair of subjects and in the same assumed term, to a so-called *student quantum*. The set of all quanta is denoted by I. Each quantum $i \in I$ is placed in the same term t_i as the corresponding students. Although the Δ students represented by a single quantum may well differ in the precise set of courses they have passed already, this simplification should be acceptable because the students will have another one or two terms to increase their credits before the next assignment phase. Thus, even students with a currently identical track record may well differ in their state at the beginning of the next term. For this reason, it is also pointless to include a full precedence check in the selection of courses for a quantum.

In the following, we describe the generation of the courses C_i assigned to each quantum $i \in I$. All courses prescribed for term t will be denoted as *regular* courses $C_r(t)$. As stated, some of these regular courses shall be assigned to conflict-free time slots with courses of previous terms - so-called *displaced courses*. As described further below we will rate courses as displaced according to historical exam data. Depending on the proportion of students taking such a course $c_r \in C_r(t)$ late, the sections of c_r will be split in two parts: One part remains in c_r to be done in term t. The remaining sections comprise a newly generated displaced course $c_d \in C_d(t+2)$ to be assigned with delay for term $t+2$. The quantum capacity of a regular or displaced course c denoted by q_c is given by the number of sections times the capacity of a section (parallel-group) $cap(c)$ scaled by the quantum size Δ.
To connect the course supply with demand, every quantum $i \in I$ with term t_i is assigned a set of courses C_i taken from the *relevant* courses $\bar{C}_i \subset (C_r(t_i) \cup C_d(t_i))$ that are required to be completed in the upcoming term. This set $\bar{C}_i$ consists of all courses, which - given quantum i assumed term t_i and combination of subjects - need to be completed according to the curriculum.
The generation of the quanta's course sets is described in Algorithm 1. Starting with the highest term t (i.e. $t = 8$) for each student quantum i at first and as long as the quantum capacity of the course is not met, all *displaced courses* $c_d \in C_d(t)$ are assigned to C_i. Secondly, regular courses $c_r \in C_r(t)$ of term t are assigned, however only if C_i does not contain any displaced predecessor course of c_r. In both cases, the algorithm stops as soon as at least 20 ECTS are reached. Note that the prescribed workload for a student in a term amounts to 30 ECTS. The chosen discrepancy serves as slack for the matching of real enrolled students in the second phase when a difference between quanta's course lists and the actual requirements of real students seems inevitable. This also helps to reach

Algorithm 1 Generation of courses C_i for quantum $i \in I$

```
n terms
displaced courses $C_d(t)$, regular courses $C_r(t)$, relevant courses $\bar{C}_i$, $C_i = \{\}$  $\forall i \in I$,
course capacities $q_c$, quantum size $\Delta$
term $t = n$
while $t \geq 1$ do
    for $i \in I$ do
        for $c_d \in C_d(t) \cap \bar{C}_i$ do
            if $q_c \geq \Delta$ then
                $C_i \leftarrow C_i \cup \{c_d\}$
                $q_c \leftarrow q_c - \Delta$
                if $ECTS(C_i) \geq 20$ then
                    break
                end if
            end if
        end for
        for $c_r \in C_r(t) \cap \bar{C}_i$ do
            if $ECTS(C_i) \geq 20$ then
                break
            end if
            if $C_i$ does not contain precedence of $c_r$ then
                $C_i \leftarrow C_i \cup \{c_r\}$
            end if
        end for
    end for
    $t \leftarrow t - 1$
end while
```

feasibility. Moreover, the university administration would like to keep some degree of freedom for the students to select additional courses on their own thus making it easier to accept a centralized course assignment regime.

From a different angle, this slack also reflects the special situation of main courses C^m which are not considered in C_i. These main courses are seen as crucial parts of each subject and therefore should be available for every student without collisions in the respective term. Thus, we will not consider their ECTS in the workload of the current term.

As mentioned above, the definition of the number of displaced sections is based on historical examination data. The examination data provides the set of students $S(a)$ that have completed this course and for each $s \in S(a)$ we are given the term $p(a, s)$ in which student s has passed the course a. Note that under the flexible rules of our university $p(a, s)$ may well differ from the prescribed term $n(a)$.

Based on this data we compute a *lateness value* $L(a, s)$ which represents the delay of student s in passing course a relative to other courses. Therefore, we count the courses (weighted by their ECTS) prescribed for later terms which s has taken in the same term or earlier than a, and the course prescribed for the same term as a but taken in an earlier term than a. The total number of these "preponed" courses serves as a lateness value $L(a,s)$ and is compared to a predefined threshold T to label course a as *passed late* by student s. Formally, we have:

$$L1(a,s) := \sum_{b\in C} ECTS(b) \text{ with } n(b) > n(a) \text{ and } p(b,s) \leq p(a,s)$$
$$L2(a,s) := \sum_{b\in C} ECTS(b) \text{ with } n(b) = n(a) \text{ and } p(b,s) < p(a,s)$$
$$L(a,s) := L1(a,s) + L2(a,s)$$

Note that this definition also yields meaningful values for "slowly progressing" students which pass *all* their courses later than prescribed. If $L(a,s) > T$ then a is passed late by s. Computing the lateness over all students $s \in S(a)$, we determine the delay factor of course a as

$$del(a) = \frac{|\{s \in S(a) \mid L(a,s) > T\}|}{|S(a)|}.$$

To align the timetable with the actual progress of students we split the $g(a)$ sections of the regular course a as follows. Defining $g_1(a) := \lfloor g(a) \cdot del(a) \rfloor$, we introduce a new "auxiliary" course a' with $g_1(a)$ sections and prescribed for the successive year, i.e. for term $n(a') := n(a) + 2$. The original course a remains at term $n(a)$ but its sections are reduced to $g(a) := g(a) - g_1(a)$. In this way, some course sections are offered in line with the study schedule of slower progressing students.

The ILP-Model The optimization step in essence seeks to assign *sections* $g \in G(c)$ of all courses $c \in C$ to time periods $p \in P$ of a recurring working week. As is the case for many timetabling problems, the main goal of the planning task is reaching a feasible solution, while the actual objective function is of secondary importance. In our planning problem, the university administration did not specify a particular goal or quality criterion for the timetable. However, our discussions with student representatives and teachers exhibited clear preferences not dissimilar from goals observed in classical university timetabling tasks. Our objective function consists of two parts: The first part aims at avoiding pairs of lectures with long breaks in between for a student quantum. Considering travel times and missing facilities for spending free time this represents the desire of having courses in a single time block. The second part takes into account pedagogical as well as group dynamic aspects. It considers each session of a course and aims at minimizing the number of different secondary subjects followed by the student quanta of this session. Indeed, it would often be preferred to have a more homogeneous student body in a lecture, possibly all enrolled in the same or only two different other subjects (besides the subject of the course).
The main decision variable $y_{cgp} = 1$ if section g of course c is assigned to $p \in P$, and 0 otherwise. Likewise a *student quantum* $i \in I$ is assigned to a section g

of one of it's compulsory courses $c \in C_i$ if variable $x_{icg} = 1$, analogously for $c \in C^m$ with $G(c) = 1$. In addition, we will introduce auxiliary variables v_{ip} for every quantum $i \in I$ and time period $p \in P$ expressing continuity of assigned time slots for quantum i, and d_{cgf} to measure the heterogeneity of the quanta assigned to a section g of course $c \in C$.

The *model's input* technically comprises:

- a set of student quanta I and quantum size Δ (students per quantum),
- overall available time periods P,
- first and last time periods of a day $FP,\ LP \subset P$,
- courses $c \in C$ with capacities $cap(c)$ and number of parallel sections $G(c)$,
- main courses C^m and for every $c_m \in C^m$ the time period $p(c_m)$
- for each quantum $i \in I$ the two subjects $f_1(i)$, $f_2(i)$, and the required courses $C_i \subseteq C$ and $C^m(i) \subseteq C^m$ [3]
- set of lecturers L and the courses $C(l) \subset C \cup C^m$ taught by each lecturer $l \in L$
- subjects F, subject $q(c) \in F$ of course c
- threshold Π limiting the number of sections taking place at the same period

The model is defined as follows:

$$\min_{d,v} \quad \alpha \cdot \sum_{p \in P} \sum_{i \in I} v_{ip} + \beta \cdot \sum_{f \in F} \sum_{c \in C} \sum_{g=1}^{G(c)} d_{cgf} \tag{2a}$$

$$\text{s.t.} \quad \sum_{g=1}^{G(c)} x_{icg} = 1, \quad \forall i \in I, c \in C_i \ \cup \ C^m(i), \tag{2b}$$

$$\sum_{p \in P} y_{cgp} \leq 1, \quad \forall c \in C \cup C^m,\ g \in \{1, \ldots, G(c)\}, \tag{2c}$$

$$y_{c_m 1 p(c_m)} = 1, \quad \forall c_m \in C^m, \tag{2d}$$

$$\sum_{c \in C(l)} \sum_{g=1}^{G(c)} y_{cgp} \leq 1, \quad \forall l \in L,\ p \in P, \tag{2e}$$

$$\sum_{i \in I} x_{icg} \cdot \Delta \leq cap(c),\ \forall c \in C,\ g \in \{1, \ldots, G(c)\}, \tag{2f}$$

$$x_{ic'g_1} + x_{ic''g_2} + y_{c'g_1p} + y_{c''g_2p} \leq 3 \quad \forall i \in I,\ c' \neq c'' \in C_i \ \cup C^m(i)\ , \quad p \in P,\ g_1 \in \{1, \ldots, G(c')\},\ g_2 \in \{1, \ldots, G(c'')\} \tag{2g}$$

$$\sum_{g=1}^{G(c)} \sum_{c \in C_i \cup C^m(i)} (y_{cgp} - y_{cg(p-1)} - y_{cg(p+1)}) \leq v_{ip} \quad \forall i \in I\ , \quad p \in P - \{FP \cup LP\} \tag{2h}$$

[3] More precisely, $C^m(i)$ contains the main lectures that belong to the term that student quantum i is assumed to be enrolled in according to her/his accomplished ECTS (for the two subjects chosen by i).

$$\sum_{g=1}^{G(c)} \sum_{c \in C_i \cup C^m(i)} (y_{cgp} - y_{cg(p+1)}) \leq v_{ip} \quad \forall i \in I, p \in FP, \tag{2i}$$

$$\sum_{g=1}^{G(c)} \sum_{c \in C_i \cup C^m(i)} (y_{cgp} - y_{cg(p-1)}) \leq v_{ip} \quad \forall i \in I, p \in LP, \tag{2j}$$

$$\sum_{g=1}^{G(c)} y_{cgp} \leq \Pi \quad \forall p \in P,\ c \in C \cup C^m, \tag{2k}$$

$$\sum_{\substack{\{i \in I: \\ f_1(i)=f \vee f_2(i)=f\}}} x_{icg} \leq d_{cgf} \left\lceil \frac{cap(c)}{\Delta} \right\rceil \ \forall c \in C, g \in \{1, \ldots, G(c)\}, \ f \in F \setminus \{q(c)\} \tag{2l}$$

$$x_{icg}, y_{cgp} \in \{0, 1\}, \tag{2m}$$

$$v_{ip}, d_{cgf} \in \mathbb{N} \tag{2n}$$

As described above the objective 2a is twofold: (A) minimizing individual timetable compactness via minimizing auxiliary variable v_{ip}, which counts the number of free periods in-between assigned lectures for each student quantum. (B) The auxiliary variable d_{cgf} represents the overall number of second subjects (curricula) in a section of a compulsory course c. Minimizing this variable results in a higher homogeneity of student quanta per section and is hoped to be didactically advantageous, since course contents can be brought into line with the second subject. The coefficients α and β allow a linear combination of the two parts and should be chosen in collaboration with the decision-makers. Constraint 2b ensures that student quantum i is assigned exactly once to a section of a mandatory course. The set of required courses for a student i, however, comprises some of the main courses C^m - that were assigned to periods in Phase 1 - and some smaller ones, such as labs or seminars C - therefore $C_i \cup C^m(i)$. Via constraint 2c sections can take place at most once (if the planning of the number of sections is fairly reliably, 2c can be written with equality). However, some slots are already taken by the main courses (with one section, $g = 1$) as assigned in the preceding Phase 1 (2d). Constraint 2e avoids that a section g of course c in the set of all courses $C(l)$ that a teacher l is giving is assigned to the same period. Given the quantum size Δ, constraint 2f ensures that the capacities of the regular courses C are not exceeded. Constraint 2g essentially avoids collisions: Whenever a student quantum i is assigned both to a section 1 and 2, these cannot take place in the same time period. If they do, the student cannot be assigned to both of them.
Constraint 2h is used to activate the auxiliary variable v_{ip}: For all time periods, except those at the beginning and the end of each working day (sets FP and LP), it is verified whether the preceding and following time slot is also taken by a section that student quantum i has to follow. If not, then there exists an isolated lecture for student i at time period p and v_{ip} is set to one. Constraints 2i and 2j account for isolated lectures at the end and the beginning of the day. Constraint 2k bounds by a *threshold* Π the number of sections that may take place at the same time period p. Although the collision avoidance can be expected to

imply a certain spread of sections over the weekly time grid, it appears necessary to impose an explicit threshold, since rooms are not explicitly considered in the model. The more evenly sections are distributed over the week, the easier it will be to find rooms, whose number of naturally restricted. Constraint 2l finally adjusts the auxiliary variable d_{cgf}, which is used in the objective function to reduce the number of different second subjects of quanta in the same section. For all sections, the enrolled subjects of assigned students are compared to all curricula (except the one the course belongs to) and counted.

4.3 Phase 3: Student assignment

Phase 3 takes place several months after Phases 1 and 2, shortly before the beginning of a new term. In this phase actual, enrolled students $s \in S$ with their updated records of courses passed are matched to *quanta* i (multiples of students) resulting from Phase 2. Although a student's s required or relevant courses C_s depend on her/his study record and curriculum-related prerequisites - that is, completion of specific courses to enroll in others - the course lists of quanta C_i are estimated based on ECTS of students one or two terms in the past. Consequently, discrepancies between the required courses of a student and the course list of the quantum the student is assigned to will be inevitable.

$$\max \quad \sum_{s \in S} \sum_{i \in I} w_{si} x_{si} \tag{3a}$$

$$\text{s.t.} \quad \sum_{i \in I} x_{si} = 1, \quad \forall s \in S, \tag{3b}$$

$$\sum_{s \in S} x_{si} \leq \Delta, \quad \forall i \in I, \tag{3c}$$

$$x_{si} \geq 0 \tag{3d}$$

As formulated above we seek a *maximum weight perfect matching* on a complete bipartite graph by solving a variant of the classical linear *assignment problem*. Associated to each 'assignment' is a variable x_{si} such that $x_{si} = 1$ iff student s is assigned to quantum i, and 0 otherwise. The weight w_{si} in the objective function 3a represents the *degree of fit* and is defined as the cardinality of the intersection of the quanta's and the actual students' course set, $w_{si} := |C_s \cap C_i|$. Naturally, students will be only allocated the courses in $C_s \cap C_i$. Constraint 3b ensures that each student s is assigned to exactly one quantum i. Via constraint 3c at most Δ students can be assigned to one quantum. It is well known that the above mathematical program can be solved as a linear program and the integrality of x_{si} is given by default.

For ease of computation, we will apply the optimization model repeatedly on smaller parts of the data, since only matching of students and quanta that belong to the *same subject combination* appears purposeful. Therefore input data $S_{\{f_1,f_2\}}$ and $I_{\{f_1,f_2\}}$ is divided accordingly as well as corresponding weight matrices $W_{\{f_1,f_2\}}$.

Allocating residual course places Recalling that the courses C_i assigned to a quantum i are usually just above 20 ECTS and that possibly not all these courses are relevant for a student s matched to quantum i, it can be expected that a sizable number of course places remain free after the above assignment phase. These places will be allocated by a *Round Robin* procedure to any students left with less than 20 ECTS assigned courses.
Students are sorted as follows: To facilitate timely graduation, at first, all students requiring at most 30 ECTS for completing their program (not counting the courses allocated in the assignment phase) are selected and sorted in increasing order of missing ECTS. All other students are appended to this sequence and sorted in increasing order of ECTS received in the above assignment phase, which reflects a max-min fairness criterion.
Considering students in this sequence, we take the first student and assign her/him a section of a course that is not fully booked, is feasible for the student w.r.t. the study program, and that does not overlap with any previously allocated courses. Among these, a course is randomly selected from those which are prescribed for the earliest term. If no allocation is possible, the student is removed from the sequence, otherwise, the student is reinserted according to the sorting criteria. One can also choose to remove students (except those close to graduation) from the sequence once their workload exceeds 20 ECTS or another bound set by the university administration.

5 Computational Insights

5.1 Data from Graz University

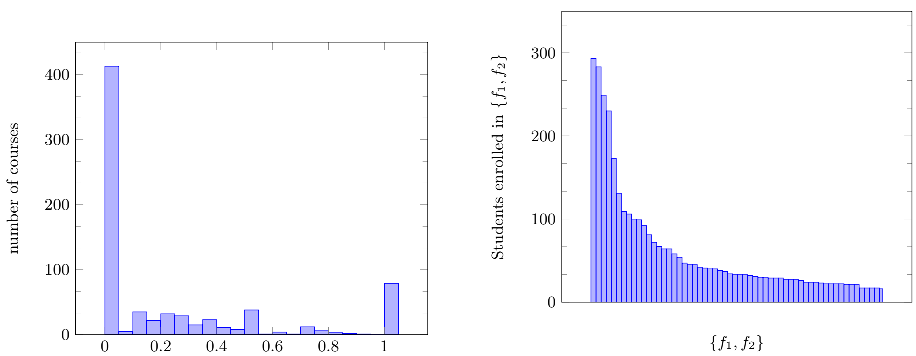
We applied our model to a subset of the courses at the University of Graz, specifically those which form the curriculum of the teacher education study program. As stated, the curriculum requires the choice of two out of 28 possible different subjects (e.g. English and German). Notice that the choice of subject pairs is not at all evenly distributed as depicted in Figure 1b. On the contrary, all combinations of the most prominent eight subjects account for more than 50% of the students, which have therefore been the focus of the study.

Table 2: Scope of the study.

Subjects	8
Courses	767
Sections	1,454
Students	2,240
Quantum size Δ	5
Quanta	454

As summarized in Table 2, the supply side consists of 1454 different sections belonging to 767 courses and we seek to assign 454 quanta to them. The courses belong to one of 8 subjects (English, German, History, Geography, Chemistry, Physics, Mathematics, Biology), constituting more than half of all students enrolled in teacher education. The quanta are established using historical anonymized examination data of all students enrolled in the study program, whereas courses and number of sections are taken from a different source representing the curricula.

The examination data further serves as the basis for the derivation of the *delay factor* of a course (as outlined in Section 4.2). As depicted in Figure 1a based on ECTS (not weekly hours) over 40% of the courses are passed late using a threshold of $T = 30$ and 79 courses are preponed entirely. Consequently, taking into account *delayed courses* C_d reflects actual student behavior in reality. The surprisingly high fraction of delayed courses indicates that (i) overlaps of courses may indeed be a reason for slowed-down progress (as often claimed by students, but sometimes questioned but other involved persons) and (ii) an optimized timetable should take the delay of courses into account for increasing study performance.



(a) Histogram of delay values. (b) Distribution of chosen subject pairs.

Fig. 1: Information derived from examination data.

5.2 Preliminary Results

We conducted our tests on a PC with processor Intel Core i5-9500 with 3.00GHz and 32GB RAM. The data processing and preparatory computation steps as well as the mathematical models have been implemented in Python and solved using the Gurobi solver (version 9.0.0). The weighting factors α and β are both set to 1, though this setting will be the subject of further discussions with all stakeholders.

Initially, a *feasibility check* concerning the overall section capacities is carried out. Based on the provided data we compare for each course the number of available places with the required places resulting from the given number of students for

each term and subject combination. In case of a shortfall, additional sections are introduced such that at least a general coverage is possible, without considering conflict freeness.
At the time being, we can give results for Phase 1 and 2, while Phase 3 could not be carried out yet due to a lack of data. However, the first two phases are considered to be the essential part, since the final output, and the timetable's applicability respectively is largely based on the quality reached in Phase 2, building upon the output of Phase 1. Moreover, from a computational point of view, Phase 2 poses the major obstacle, while Phases 1 and 3 will not be a hurdle for practical solvability. Within the given scope we solve the current test case yielding 537,330 ILP variables (where 469,743 are binary) and reach preliminary results as follows.

Table 3: Results of the first two phases.

subjects	Π	time[s]	obj. part 1	obj. part 2	gap[%]
8	20	8.5	632	2,307	0.00
8	15	9.4	573	2,611	0.00
8	10	10.0	651	2,778	0.00
8	5	16.8	733	2,941	0.00

Examining the results of the case and analyzing different parameter settings we can make the following observations:

- The test case concerning the scope in Table 2 can be solved to optimality in a surprisingly short running time despite the considerable size of the ILP model.
- Reducing the parameter Π, which is limiting the number of sections at the same time period, results in decreasing solution quality and at some point, the instance is not feasible anymore (e.g. $\Pi = 4$ for the combinations of 8 subjects). Table 3 shows for different Π the number of integer and binary variables the ILP-model of the test case comprises, the objective function value, and the elapsed computation time (in seconds).
- Extending the scope, however, to a larger set of subject combinations may result in infeasibility of the problem - even with the addition of just one single subject (i.e. with all the associated pairs of subjects). If the problem remains feasible, different numbers of variables are generated for different added subjects. Also, the new solution values vary a lot as well as the computation times (see examples in Table 4).
  An obvious reason for this behavior is that the amount of enrolled students and the course structure is subject-related. A thorough examination of this relationship is subject to further investigation.
- In general it is a surprise, that computation time is not an issue, while feasibility is.

Table 4

added subject	time[s]	obj.	gap[%]	int. variables (binary)
Informatics	9.0	4,003	0.00	589,515 (485,538)
French	34.85	3,924	0.00	580,658 (477,011)
Nutrition	51.81	3,978	0.00	581,493 (477,696)
Sports	-	infeasible	-	635,892 (530,025)

For the practical application of our solution approach to the full planning problem with 28 subjects we currently see five measures for dealing with the inherent infeasibility issue:

(i) Increase the number of options by adding more sections.
(ii) Increase the capacity $cap(c)$ of each section of course c.
(iii) Restrict the optimization model to a subset of the most frequently chosen subjects (up to 10 or 12) and add less popular subjects by a manual process (basically as it is done now).
(iv) Omit pairs of subjects from consideration which are chosen by a very small number of students.
(v) Reduce the number of ECTS for which collision-free courses are provided by the planning system for every student.

While measures (iii) and (iv) will be unavoidable and easily accepted, there is an interesting cost/quality trade-off involved in the decision for (i) and (ii): The former causes additional costs (assuming that external teachers are available) whereas the latter comes for free but diminishes teaching quality. Thus, it will be very interesting for the decision-makers to be informed about the effect of employing certain levels of measures (i) and (ii). In particular, it will be interesting to identify a suitable subset of crucial courses for which these measures should be applied to reach feasibility. Measure (v) is easy to implement and does not incur any direct cost, but it compromises the original goal of this project.

6 Conclusions

In this paper, we developed a solution approach for a complex university timetabling task arising at the University of Graz, Austria. The main features which make our problem different from standard university timetabling instances are the following:

1. Each student is enrolled in two different subjects selected arbitrarily from a wide range of available subjects.
2. Student's progress does not follow a strict term pattern but may exhibit highly irregular behavior, including gaps and deviations from the prescribed ordering of courses.

3. Timetable planning is done at an early stage when the data of actual student numbers and their progress status, as it is required for finally assigning courses to students, is still subject to major changes.
4. Most courses are offered in several parallel groups, which requires the sectioning of students to reach a conflict-free timetable for a highly heterogeneous set of students.

Our solution approach consists of three phases. Two of them solve variants of the linear assignment problem, extended by conflict constraints. The main planning task (Phase 2) is performed by an intricate integer linear program (ILP). In this way, we managed to determine feasible timetables offering conflict-free course allocations for the projected student body. The data for the final allocation of individual students are still missing, but the handling of the main computational hurdle, namely the solution of a complex ILP, can be illustrated by our computational results. These reveal in particular that the choice of subjects given their course structure and amount of enrolled students per term has a non-trivial impact on solvability, computation time, and solution quality and is of interest for further investigation.

We expect to employ the full solution approach in practice for the planning task in the next year. It will also be interesting to investigate additional options for the objective function since different stakeholders have different ideas about the appropriate quality measure of a timetable.

In the future, we could also try to rate the difficulty of the problem according to the 'Complexity' as introduced in [9].

Acknowledgments This study was partially funded by the University of Graz under the Field of Excellence "COLIBRI".

References

1. Aubin, J., Ferland, J.A.: A large scale timetabling problem. Computers & Operations Research **16**(1), 67–77 (1989)
2. Bagger, N.C.F., Sørensen, M., Stidsen, T.R.: Dantzig–Wolfe decomposition of the daily course pattern formulation for curriculum-based course timetabling. European Journal of Operational Research **272**(2), 430–446 (2019)
3. Banks, D., van Beek, P., Meisels, A.: A heuristic incremental modeling approach to course timetabling. In: Carbonell, J.G., Siekmann, J., Goos, G., Hartmanis, J., van Leeuwen, J., Mercer, R.E., Neufeld, E. (eds.) Advances in Artificial Intelligence, Lecture Notes in Computer Science, vol. 1418, pp. 16–29. Springer (1998)
4. Ceschia, S., Di Gaspero, L., Schaerf, A.: Educational timetabling: Problems, benchmarks, and state-of-the-art results. arXiv preprint arXiv:2201.07525 (2022)
5. Dostert, M., Politz, A., Schmitz, H.: A complexity analysis and an algorithmic approach to student sectioning in existing timetables. Journal of Scheduling **19**(3), 285–293 (2016)
6. Esmaeilbeigi, R., Mak-Hau, V., Yearwood, J., Nguyen, V.: The multiphase course timetabling problem. European Journal of Operational Research **300**(3), 1098–1119 (2021)

7. Holm, D.S., Mikkelsen, R.Ø., Sørensen, M., Stidsen, T.J.: A graph-based MIP formulation of the international timetabling competition 2019. Journal of Scheduling pp. 1–24 (2022)
8. Johnes, J.: Operational research in education. European Journal of Operational Research **243**(3), 683–696 (2015)
9. de La Rosa-Rivera, F., Nunez-Varela, J.I., Puente-Montejano, C.A., Nava-Muñoz, S.E.: Measuring the complexity of university timetabling instances. Journal of Scheduling **24**(1), 103–121 (2021)
10. Lach, G., Lübbecke, M.E.: Curriculum based course timetabling: new solutions to Udine benchmark instances. Annals of Operations Research **194**(1), 255–272 (2012)
11. Müller, T.: ITC 2019: Preliminary results using the UniTime solver. In: Proceedings of the 13th International Conference on the Practice and Theory of Automated Timetabling (PATAT), Volume III (2022)
12. Müller, T., Murray, K.: Comprehensive approach to student sectioning. Annals of Operations Research **181**(1), 249–269 (2010)
13. Müller, T., Rudová, H.: Real-life curriculum-based timetabling with elective courses and course sections. Annals of Operations Research **239**(1), 153–170 (2016)
14. Müller, T., Rudová, H., Müllerová, Z.: University course timetabling and International Timetabling Competition 2019. In: Proceedings of the 12th International Conference on the Practice and Theory of Automated Timetabling (PATAT-2018). pp. 5–31 (2018)
15. Schindl, D.: Optimal student sectioning on mandatory courses with various sections numbers. Annals of Operations Research **275**(1), 209–221 (2019)