Un esquema basado en programación lineal entera para la programación de horarios de clases en la Universidad Torcuato Di Tella

Carolina Ferreiro[0009-0008-4170-6719]
Javier Marenco[0000-0003-2694-4758]

Escuela de Negocios, Universidad Torcuato Di Tella, Av. Figueroa Alcorta 7350
(1428) Ciudad de Buenos Aires, Argentina
carolina.ferreiro@mail.utdt.edu, javier.marenco@utdt.edu

Resumen En este trabajo consideramos el problema de programación de horarios de las clases de grado en la Universidad Torcuato Di Tella. Dado el conjunto de comisiones a dictar, la disponibilidad de docentes y aulas, y diversas restricciones sobre la programación, el problema consiste en determinar en qué horario se debe dictar cada clase de cada comisión intentando minimizar los "tiempos muertos" para docentes y estudiantes. Una característica importante del problema consiste en que muchas materias son compartidas entre varias (o todas) las carreras ofrecidas por la universidad, de modo tal que se debe proponer un esquema de cursada que sea factible para todos los grupos de estudiantes. Proponemos en este trabajo un enfoque de dos fases para este problema junto con un esquema secuencial para resolver consecutivamente sub-instancias de la instancia original. Analizamos los resultados sobre los datos de la universidad y comentamos los pasos que se están dando para implementar esta herramienta en la universidad.

Keywords: timetabling · programación entera · resolución secuencial

En este trabajo consideramos el siguiente problema, que forma parte de las actividades de planificación académica de la Universidad Torcuato Di Tella.

Consideramos una semana de cinco *días* (llamamos $D = \{1, \ldots, 5\}$ al conjunto de días), y cada día está dividido en siete *bandas horarias* (llamamos $B = \{1, \ldots, 7\}$ al conjunto de bandas horarias). Llamamos *módulo* a la combinación de día y banda horaria, de modo tal que la semana tiene 35 módulos totales, representados por el conjunto $H = D \times B$.

Tenemos un conjunto M de *materias*, y cada materia $m \in M$ está compuesta por un conjunto S_m de *comisiones* a dictar. Llamamos $S = \cup_{m \in M} S_m$ al conjunto de todas las comisiones. A su vez, cada comisión $s \in S$ está compuesta por un conjunto C_s de *clases*. Particionamos el conjunto de clases de cada comisión en un conjunto de *clases teóricas* y un conjunto de *clases prácticas*. Llamamos $C = \cup_{s \in S} C_s$ al conjunto de todas las clases. Cada comisión $s \in S$ tiene un conjunto de *módulos factibles* $H_s \subseteq H$ en los cuales se pueden programar sus clases, y un *cupo máximo* $\text{cupo}_s \in \mathbb{Z}_+$ de estudiantes que pueden cursar en la comisión. Además, cada clase $c \in C$ tiene asociado un conjunto T_c de *docentes*. Cada docente $t \in T_c$ tiene un conjunto de *módulos factibles* $H_t \subseteq H$ en los cuales puede dictar clases, y una cantidad máxima $\max_t \in \mathbb{Z}_+$ de módulos horarios que puede dictar por día. Además, para cada comisión tenemos parámetros que especifican si todas las clases de la comisión deben dictarse en forma consecutiva en el mismo día o, en caso negativo, la mínima cantidad de días entre dos clases teóricas/prácticas consecutivas.

Tenemos además un conjunto P de *paquetes* de estudiantes, que representan grupos de estudiantes con requerimientos de cursada similares. Para cada paquete $p \in P$, tenemos la cantidad $\text{cant}_p \in \mathbb{Z}_+$ de estudiantes en el paquete y un conjunto $M_p \subseteq M$ de materias que deben cursar los estudiantes del paquete. Tenemos además la cantidad máxima $\max_p \in \mathbb{Z}_+$ de módulos horarios que los estudiantes del paquete pueden tomar en cada día. Finalmente, tenemos un conjunto A de aulas, y cada aula $a \in A$ tiene una capacidad máxima de estudiantes $\text{cap}_a \in \mathbb{Z}_+$ y un conjunto $H_a \subseteq H$ de módulos en los que el aula está disponible.

El problema consiste en determinar (a) un módulo horario $\tau_c \in H$ y un aula $\rho_c \in A$ para cada clase $c \in C$, y (b) una partición $G_1^p \cup \cdots \cup G_{k_p}^p$ en *grupos* del conjunto $\{1, \ldots, \text{cant}_p\}$ de estudiantes de cada paquete $p \in P$ junto con una asignación de comisiones $\alpha_{ip} \subseteq S$ para $i = 1, \ldots, k_p$, de modo tal que se cumplan las siguientes restricciones.

1. Dos clases de una misma comisión no pueden recibir el mismo módulo horario (es decir, $\tau_c \neq \tau_{c'}$ para todo $c, c' \in C_s$, $c \neq c'$, para toda comisión $s \in S$).
2. Dos clases programadas en el mismo módulo horario no pueden recibir la misma aula (es decir, si $\tau_c = \tau_{c'}$ para $c, c' \in C$, $c \neq c'$, entonces $\rho_c \neq \rho_{c'}$).
3. Para cada comisión $s \in S$ y cada clase $c \in C_s$, el módulo horario τ_c asignado a la clase c debe cumplir $\tau_c \in H_s$ y $\tau_c \in H_t$ para todo docente $t \in T_c$ (es decir, se debe programar cada clase en un módulo permitido para la clase y el docente).
4. Los módulos horarios de cada comisión deben cumplir las restricciones de dictarse en forma consecutiva (si corresponde) o, en su defecto, de días mínimos entre dos clases teóricas/prácticas consecutivas.
5. Cada docente $t \in T$ debe tener asignados a lo sumo $\max_t$ módulos diarios de clase. Esta restricción es elástica.
6. Para cada comisión $s \in S$ y cada clase $c \in C_s$, el aula ρ_c asignada a la clase c debe cumplir $\text{cupo}_s \leq \text{cap}_{\rho_c}$. Esta restricción es elástica.
7. Para cada paquete $p \in P$ y cada $i \in \{1, \ldots, k_p\}$, el conjunto α_{ip} de comisiones asignado al *grupo* G_i^p del paquete p debe contener exactamente una comisión de cada materia de M_p.
8. Para cada paquete $p \in P$ y cada $i \in \{1, \ldots, k_p\}$, las clases de las comisiones de α_{ip} no se pueden superponer (es decir, no puede haber dos clases $c, c' \in \cup_{s \in \alpha_{ip}} C_s$, $c \neq c'$, con $\tau_c = \tau_{c'}$).
9. Para cada paquete $p \in P$, los estudiantes del paquete p no tienen asignados más de $\max_p$ módulos diarios de clase. Esta restricción es elástica.

El objetivo de la asignación es minimizar los "tiempos muertos" (es decir, módulos horarios sin clases entre módulos con clase en un mismo día) tanto para docentes como para estudiantes de un mismo paquete.

Proponemos en este trabajo dos modelos de programación linel entera para este problema, que se resuelven secuencialmente. El modelo de **Fase I** utiliza los siguientes conjuntos adicionales a los ya presentados.

- $G_p = \{1, \ldots, k_p\}$: Conjunto de grupos asociados con el paquete p, para $p \in P$ y una constante k_p seleccionada arbitrariamente. Definimos además $G = \cup_{p \in P} G_p$.
- $M_g \subseteq M$: Conjunto de materias que deben cursar los estudiantes del grupo g, para $g \in G$, es decir, $M_g = \{m \in M : \text{existe } p \in P \text{ con } g \in G_p \text{ y } m \in M_p\}$.
- $S_g \subseteq S$: Conjunto de comisiones compatibles con el grupo g, es decir, $S_g = \{s \in S : \text{existe } m \in M_g \text{ con } s \in S_m\}$.

Además, utiliza las siguientes variables.

- $z_{gs} \in \{0,1\}$: El grupo g toma la comisión s, para $g \in G$, $s \in S_g$.
- $w_g \in \mathbb{R}_+$: Cantidad de estudiantes en el grupo g, para $g \in G$.
- $w_{gs} \in \mathbb{R}_+$: Cantidad de estudiantes del grupo g asignados a la comisión s, para $g \in G$, $s \in S_g$.
- $wc_s \in \mathbb{R}_+$: Violación de cupo máximo de estudiantes que pueden cursar en la comisión s, para $s \in S$.
- $maxvc \in \mathbb{R}_+$: Máxima violación de cupo entre todas las comisiones.

Con estas definiciones, el modelo de Fase I se define del siguiente modo.

$$\text{mín} \quad maxvc \tag{1}$$

$$\sum_{s \in S_m} z_{gs} = 1 \qquad \forall g \in G,\ \forall m \in M_g \tag{2}$$

$$w_{gs} \leq \text{cant}_p z_{gs} \qquad \forall g \in G,\ \forall p \in P,\ \forall s \in S_g \tag{3}$$

$$\sum_{s \in S_g} w_{gs} = w_g \qquad \forall g \in G \tag{4}$$

$$\sum_{g \in G: s \in S_g} w_{gs} \leq \text{cupo}_s + wc_s \qquad \forall s \in S \tag{5}$$

$$wc_s \leq maxvc \qquad \forall s \in S \tag{6}$$

$$\sum_{g \in G_p} w_g = \text{cant}_p \qquad \forall p \in P \tag{7}$$

$$z_{gs} \in \{0,1\} \qquad \forall g \in G,\ \forall s \in S_g \tag{8}$$

$$w_g \geq 0 \qquad \forall g \in G \tag{9}$$

$$w_{gs} \geq 0 \qquad \forall g \in G,\ \forall s \in S_g \tag{10}$$
$$wc_s \geq 0 \qquad \forall s \in S \tag{11}$$
$$maxvc \geq 0 \tag{12}$$

La función objetivo (1) busca minimizar la violación máxima de cupo de las comisiones, garantizando la distribución equitativa de los estudiantes entre las comisiones de cada materia. La restricción (2) asegura que cada grupo sea asignado exactamente a una comisión de cada materia de M_p. La restricción (3) establece una relación consistente entre las variables binarias y continuas. Al ligar las variables w_{gs} y z_{gs}, esta restricción asegura que si un grupo g fue asignado a una comisión s, es decir, si $z_{gs} = 1$, entonces $w_{gs} > 0$ o, específicamente, $w_{gs} = w_g$ como establece la restricción (4). Esta restricción asegura que todos los estudiantes de cada grupo g puedan cursar en las comisiones asignadas al grupo al que pertenecen. La restricción (5) garantiza que la cantidad total de estudiantes entre todos los grupos asignados a una comisión s no exceda el cupo máximo de estudiantes que pueden cursar en la comisión, más un término de penalización que contabiliza cualquier violación de cupo máximo de la comisión. La restricción (6) introduce un límite superior al incumplimiento del cupo máximo de cada comisión. Juntas, las restricciones (5) y (6), buscan evitar la sobreasignación de estudiantes a una comisión específica, manteniendo una asignación equilibrada y factible. La restricción (7) asegura que, para cada paquete p, la suma de los estudiantes de todos los grupos pertenecientes a p sea igual a la cantidad total de estudiantes en p. Finalmente, (8) a (12) aseguran que cada variable tome el dominio adecuado.

El modelo de **Fase II** toma como entrada los grupos de estudiantes formados por el modelo de Fase I y diseña los horarios. Llamamos un *esquema* de cursada a la especificación de los horarios completos de las clases teóricas o prácticas de una sección. Este modelo tiene una etapa de preprocesamiento en la que se generan todos los esquemas posibles para cada sección. Consideramos los siguientes conjuntos adicionales.

- Q_s^T: Conjunto de esquemas correspondientes a las *clases teóricas* de la comisión s, para $s \in S$.
- Q_s^P: Conjunto de esquemas correspondientes a las *clases prácticas* de la comisión s, para $s \in S$.
- $Q = \cup_{s \in S}(Q_s^T \cup Q_s^P)$, conjunto de todos los esquemas.
- R: Conjunto de docentes asignados a más de una comisión.
- U: Conjunto de capacidades de las aulas.
- $Q_{db} \subset Q$: Conjunto de esquemas que tienen una clase en el día d y en la banda b, para $d \in D$, $b \in B$.

- $Q_{sdb} \subset Q$: Conjunto de esquemas de la comisión s que tienen una clase en el día d y en la banda horaria b, para $s \in S$, $d \in D$, $b \in B$.
- $Q_{rd} \subset Q$: Conjunto de esquemas correspondientes al docente r en el día d, para $r \in R$, $d \in D$.
- $Q_{rdb} \subset Q$: Conjunto de esquemas correspondientes al docente r en el día d y en la banda horaria b, para $r \in R$, $d \in D$, $b \in B$.
- $Q_{gd} \subset Q$: Conjunto de esquemas de las comisiones asignadas al grupo g en la Fase I que tienen una clase en el día d, para $g \in G$, $d \in D$.
- $Q_{gdb} \subset Q$: Conjunto de esquemas de las comisiones asignadas al grupo g en la Fase I que tienen una clase en el día d y en la banda horaria b, para $g \in G$, $d \in D$, $b \in B$.

Además, se tienen los siguientes parámetros.

- ideal_i: Cantidad ideal de bandas horarias en del esquema i, para $i \in Q$.
- banded_i: Cantidad de bandas horarias dentro del turno seleccionado para el esquema i, para $i \in Q$.
- cap_i: Capacidad de la comisión asociada al esquema i, para $i \in Q$.
- cant_u: Cantidad de aulas con capacidad u, para $u \in U$.
- primera_{id}: Primera banda horaria del esquema i en el día d, para $i \in Q$, $d \in D$.
- ultima_{id}: Última banda horaria del esquema i en el día d, para $i \in Q$, $d \in D$.

El modelo utiliza las siguientes variables.

- $x_i \in \{0,1\}$: El esquema i es elegido, para $i \in Q$.
- $w_{sdb} \in \mathbb{R}_+$: Exceso de clases asignadas a la comisión s en el día d en la banda horaria b, para $s \in S$, $d \in D$, $b \in B$.
- $w_{rdb} \in \mathbb{R}_+$: Exceso de clases asignadas al docente r en el día d en la banda horaria b, para $r \in R$, $d \in D$, $b \in B$.
- $w_{gdb} \in \mathbb{R}_+$: Exceso de clases asignadas al grupo g en el día d en la banda horaria b, para $g \in G$, $d \in D$, $b \in B$.
- $w_{udb} \in \mathbb{R}_+$: Violación de la capacidad máxima u de estudiantes en el día d en la banda horaria b, para $u \in U$, $d \in D$, $b \in B$.
- $mw \in \mathbb{R}_+$: Máxima violación de la capacidad de un aula.

- sd_{rd}, ed_{rd}: Primera y última (+ 2) bandas horarias programadas para el docente r en el día d, para $r \in R$, $d \in D$.

- sg_{gd}, eg_{gd}: Primera y última (+ 2) bandas horarias programadas para el grupo g en el día d, para $g \in G$, $d \in D$.

Finalmente, definimos las siguientes penalizaciones para la función objetivo.

- $\alpha = 0{,}0001$ si se selecciona la función objetivo con bandas, y $\alpha = 0$ en caso contrario.

- $\beta = -100$: Penalización por violaciones de restricciones elásticas.

- $\gamma = 0{,}01$: Penalización por violaciones de cantidad máxima de módulos horarios que un docente puede dictar por día.

- $\delta = 0{,}01$ si se penalizan las violaciones de cantidad máxima de módulos horarios que un grupo puede cursar por día en la función objetivo, y $\delta = 0$ en caso contrario.

- $\nu = -10000$: Penalización por violaciones de la capacidad de aulas.

En este contexto, el modelo de Fase II es el siguiente.

$$\begin{aligned}\text{máx} \sum_{i \in Q} (\text{ideal}_i + \alpha \text{banded}_i) x_i + \beta \sum_{s \in S} \sum_{d \in D} \sum_{b \in B} w_{sdb} \\ + \beta \sum_{r \in R} \sum_{d \in D} \sum_{b \in B} w_{rdb} + \beta \sum_{g \in G} \sum_{d \in D} \sum_{b \in B} w_{gdb} \\ - \gamma \sum_{r \in R} \sum_{d \in D} (ed_{rd} - sd_{rd}) - \delta \sum_{g \in G} \sum_{d \in D} (eg_{gd} - sg_{gd}) + \nu mw \end{aligned} \tag{13}$$

$$\sum_{i \in Q_s^T} x_i = 1 \qquad \forall s \in S : Q_s^T \neq \emptyset \tag{14}$$

$$\sum_{i \in Q_s^P} x_i = 1 \qquad \forall s \in S : Q_s^P \neq \emptyset \tag{15}$$

$$\sum_{i \in Q_{sdb}} x_i \leq 1 + w_{sdb} \qquad \forall s \in S,\ \forall d \in D,\ \forall b \in B \tag{16}$$

$$\sum_{i \in Q_{rdb}} x_i \leq 1 + w_{rdb} \qquad \forall r \in R,\ \forall d \in D,\ \forall b \in B \tag{17}$$

$$\sum_{i \in Q_{gdb}} x_i \leq 1 + w_{gdb} \qquad \forall g \in G,\ \forall d \in D,\ \forall b \in B \tag{18}$$

$$\sum_{i \in Q_{db} : \text{cap}_i \geq u} x_i \leq \text{cant}_u + w_{udb} \qquad \forall u \in U,\ \forall d \in D,\ \forall b \in B \tag{19}$$

$$w_{udb} \leq mw \qquad \forall u \in U,\ \forall d \in D,\ \forall b \in B \tag{20}$$

$$\sum_{i \in Q_{rd}} \text{primera}_{id} x_i \geq sd_{rd} \qquad \forall r \in R,\ \forall d \in D \tag{21}$$

$$\sum_{i \in Q_{rd}} (\text{ultima}_{id} + 2)x_i \leq ed_{rd} \qquad \forall r \in R,\ \forall d \in D \tag{22}$$

$$sd_{rd} \leq ed_{rd} \qquad \forall r \in R,\ \forall d \in D \tag{23}$$

$$\sum_{i \in Q_{gd}} \text{primera}_{id} x_i \geq sg_{gd} \qquad \forall g \in G,\ \forall d \in D \tag{24}$$

$$\sum_{i \in Q_{gd}} (\text{ultima}_{id} + 2)x_i \leq eg_{gd} \qquad \forall g \in G,\ \forall d \in D \tag{25}$$

$$sg_{gd} \leq eg_{gd} \qquad \forall g \in G,\ \forall d \in D \tag{26}$$

$$x_i \in \{0,1\} \qquad \forall i \in Q \tag{27}$$

$$w_{sdb} \geq 0 \qquad \forall s \in S,\ \forall d \in D,\ \forall b \in B \tag{28}$$

$$w_{rdb} \geq 0 \qquad \forall r \in R,\ \forall d \in D,\ \forall b \in B \tag{29}$$

$$w_{gdb} \geq 0 \qquad \forall g \in G,\ \forall d \in D,\ \forall b \in B \tag{30}$$

$$w_{udb} \geq 0 \qquad \forall u \in U,\ \forall d \in D,\ \forall b \in B \tag{31}$$

$$sd_{rd}, ed_{rd} \geq 0 \qquad \forall r \in R,\ \forall d \in D \tag{32}$$

$$sg_{gd}, eg_{gd} \geq 0 \qquad \forall g \in G,\ \forall d \in D \tag{33}$$

$$mw \geq 0 \tag{34}$$

Este modelo propone como objetivo principal (13) maximizar la selección de esquemas ideales para cada comisión, minimizando a su vez el incumplimiento de las capacidades de las aulas, las superposiciones de clases y los tiempos muertos tanto para docentes como para estudiantes de cada grupo. Las restricciones (14) y (15) garantizan que cada comisión reciba un único esquema para sus clases teóricas y prácticas, asegurando una asignación completa. Las restricciones (16), (17) y (18) introducen las variables de penalización w_{sdb}, w_{rdb} y w_{gdb}, que contabilizan las asignaciones de clases superpuestas para cada comisión, docente, y grupo, respectivamente. La restricción (19) asegura que el esquema i seleccionado sea asignado a un aula con capacidad suficiente en la medida que sea posible. La variable de violación asociada, w_{udb}, es la encargada de penalizar cualquier incumplimiento de la capacidad de las aulas con el objetivo de alentar al modelo a encontrar soluciones que respeten las capacidades de las aulas. La restricción (20) limita la violación máxima de la capacidad de las aulas, evitando desviaciones excesivas. Las restricciones (21) a (23) buscan minimizar los tiempos muertos (es decir, los módulos horarios sin clases entre módulos con clase en un mismo día) para docentes, e igulmente las restricciones (24) a (26) para estudiantes de un mismo grupo. La adición de dos unidades en las restricciones (22) y (25) tiene el propósito de penalizar tanto la cantidad de horas o módulos horarios en la universidad como la cantidad de días de cursada, impulsando al modelo a encontrar una solución que equilibre la cantidad de clases por día, minimizando los tiempos muertos, y la cantidad de días de cursada. Por último, las restricciones (27) a (34) garantizan que las variables de decisión tomen valores apropiados.

Dado que los tiempos de ejecución de estos modelos son relativamente altos para los solvers actuales, proponemos también un esquema secuencial para resolver este problema, que consiste en resolver consecutivamente sub-instancias de la instancia original y acumular los resultados para construir una solución completa. Analizamos los resultados sobre los datos de la universidad, mostrando que este esquema permite encontrar soluciones de buena calidad en tiempos de ejecución aceptables. Además, comentamos los pasos que se están dando para implementar esta herramienta en la universidad, incluyendo el desarrollo de una herramienta de visualización y edición de los resultados y la integración con los sistemas de la universidad. Existen en la literatura trabajos que resuelven problemas similares para instituciones educativas (ver, por ejemplo, [1], [2], [3] y [4]).

Agradecimientos. Queremos expresar nuestro especial agradecimiento a Ariana Bernachea, Juan José Cruces, Hernán Langé y Agustín Socías por el apoyo y el acompañamiento institucional durante el desarrollo de este proyecto.

Referencias

1. Christou, I., Vagianou E., y Vardoulias, G.: Planning Courses for Student Success at the American College of Greece. INFORMS Journal on Applied Analytics, to appear.
2. Kassa, B.: Implementing a Class-Scheduling System at the College of Business and Economics of Bahir Dar University, Ethiopia. Interfaces **45** 203–215 (2015).
3. Miranda, J.: A Course Scheduling System for the Executive Education Unit at the Universidad de Chile. Interfaces **40**, 196–207 (2010).
4. Miranda, J., Rey , P., y Robles J.: udpSkeduler: A Web architecture based decision support system for course and classroom scheduling. Decis. Support Syst. **52** 505–513 (2012).