Confection d'un horaire des activités de formation pour un programme d'étude

Sarah Morency-Lemieux et Viviane Gascon
{sarah.morency-lemieux|viviane.gascon}@uqtr.ca

Département de mathématiques et d'informatique, École de gestion

Problématique

Le projet consiste à confectionner l'horaire des activités de formation pour une école spécialisée. Le programme d'étude comporte 179 activités étalées sur une durée de 15 semaines pour une cohorte de 72 étudiants. Comme spécifications particulières, nous pouvons noter :

- Taille variée des groupes des activités
- Durée variée des activités
- Activités préalables
- Ressources disponibles en nombre limité (enseignants, locaux, matériels)
- Convention collective des enseignants

Objectif : Générer un horaire de formation tout en minimisant le recours aux enseignants externes.

Présentation de CPLEX

IMB ILOG CPLEX Optimization est un logiciel qui permet de résoudre des problèmes d'optimisation de programmation linéaire, de programmation en nombres entiers en utilisant diverses approches de résolution dont la programmation par contraintes. Dans le cadre de ce projet, c'est ce logiciel qui a été utilisé pour résoudre le problème.

Paramètres

A	ensemble de toutes les activités
$D_1 \subset A$	ensemble des activités qui durent 1 période (90 minutes)
$D_2 \subset A$	ensemble des activités qui durent 2 périodes (180 minutes)
J	ensemble des jours de planification (75 jours ou 69 jours si on considère les jours fériés)
P	ensemble des périodes disponibles durant une journée ; $P = \{1, ..., 6\}$, soit 2 périodes le matin, 2 l'après-midi et 2 le soir. Les activités de 90 minutes (1 période) peuvent débuter à n'importe quelle période de la journée.
P'	ensemble contenant le dénombrement de toutes les périodes ; $P' = \{1, .., 450\}$
AR	ensemble des couples d'activités (a_1, a_2) pour lesquelles il y a une relation de préséance (a_1 est préalable à a_2).
$ecartmin_{a_1a_2}$	nombre minimal de périodes entre le début de l'activité a_1 et de l'activité a_2
G_a	ensemble des groupes possibles de l'activité a ; $G_a = \{1, .., n\}$ où n est le nombre de groupes de l'activité, soit 1 (72 étudiants), 2 (36 étudiants), 4 (18 étudiants), 6 (12 étudiants), 8 (9 étudiants), 12 (6 étudiants).
G	ensemble généralisant les groupes possibles pour toutes les activités ; $G = \{1, ..., 8\}$
$G_{a_1a_2}$	$G_{a_1a_2} = \{(g_1, g_2) : g_1 \in G_{a_1}, g_2 \in G_{a_2}$ et g_1, g_2 sont *reliés*$\}$ Deux groupes sont *reliés* lorsqu'ils ont des étudiants en commun.

Constraint programming

Principe

- *Filtrage* : réduction du domaine des variables en supprimant les valeurs qui ne peuvent appartenir à la solution pour un sous-problème
- *Propagation* : la modification du domaine d'une variable peut entrainer la réduction du domaine d'autres variables
- *Recherche de solutions* : trouver une solution par l'affectation de valeurs

Variable de décision

$$x_{ag} = p \qquad \text{où } p \in P', a \in A, G \in G$$

En d'autres mots, la signification de la variable est : l'activité a du groupe g débute à la période globale p.

Remarque : la contrainte (1) est maintenant implicite par la définition de la variable : chaque groupe de chaque activité se verra attribuer une période de début.

Cette approche de résolution offre d'excellentes performances : le temps de résolution est moins d'une minute avec les contraintes similaires au modèle mathématique. Il est donc intéressant d'ajouter les enseignants au modèle. Il faut tenir compte de :

- Chaque activité doit être enseignée par le nombre d'enseignants compétents requis
- Un enseignant ne peut enseigner plus d'une activité par période
- Modification de la fonction objectif : minimiser le recours aux enseignants externes

Le temps de résolution est d'environ 6 minutes.

Modèle mathématique

Approche initiale

Construction de l'horaire des activités en laissant de côté les enseignants. L'idée est d'obtenir rapidement une solution réalisable et ensuite d'améliorer l'horaire en prenant en considération les contraintes reliées aux enseignants (approche heuristique).

Ce problème peut être modélisé comme un *problème d'optimisation en nombre entiers avec variables binaires*.

Variable de décision

$$x_{agjp} = \begin{cases} 1 & \text{si l'activité } a \text{ du groupe } g \text{ est enseignée le jour } j \text{ à la période } p \\ 0 & \text{sinon} \end{cases} \qquad \text{où } j \in J, p \in P, a \in A, g \in G_a$$

Exemples de contraintes

Chaque activité doit être offerte à tous les groupes à exactement d_a périodes, où d_a est la durée de l'activité en nombre de périodes.

$$\sum_{j \in J} \sum_{p \in P} x_{agjp} = d_a \qquad \forall a \in A, g \in G_a \tag{1}$$

Il faut aussi s'assurer que les périodes soient consécutives (si $n > 1$). Par exemple, pour une activité dont la durée est de 2 périodes, sachant qu'une telle activité peut commencer aux périodes 1, 3 ou 5, il faut ajouter la contrainte :

$$x_{agjp} = x_{agj(p+1)} \qquad \forall j \in J, p \in \{1, 3, 5\}, a \in D_2, g \in G_a \tag{2}$$

Uniquement en considérant la contrainte (1) et les différents cas de (2), le nombre de contraintes générées est estimé à 3253 contraintes. Ce problème est donc complexe et de grande taille : beaucoup de variables et de contraintes.

Redéfinition de la variable

Pour réduire le nombre de variables, on a redéfini la variable en considérant seulement les jours et les périodes durant lesquels l'activité peut débuter. On a ensuite adapté le modèle en conséquence, ce qui permet d'alléger certaines contraintes. Le nombre de variables est passé de 313 384 à 119 507.

$$x_{agjp} = \begin{cases} 1 & \text{si l'activité } a \text{ du groupe } g \text{ débute le jour } j \text{ à la période } p \\ 0 & \text{sinon} \end{cases} \qquad \text{où } a \in A, g \in G_a$$

et où $j \in J_a, p \in P_a$, la fenêtre de temps de l'activité a

Exemple de modification des contraintes

La contrainte (1) devient :

$$\sum_{j \in J_a} \sum_{p \in P_a} x_{agjp} = 1 \qquad \forall a \in A, g \in G_a$$

Le nombre de contraintes est réduit à environ 358 contraintes (la contrainte (2) n'étant plus nécessaire). En plus, elles sont moins lourdes (moins de termes).

Complexité du problème

Ce problème est aussi complexe, parce qu'il est plus difficile de formuler les contraintes avec des groupes de taille variée. Par exemple, les contraintes modélisant les *relations de préséance* doivent être générées seulement entre les groupes ayant des étudiants en commun.

Exemple : Si les activités a_1 et a_2 durent 90 minutes (une période), l'activité a_2 doit commencer au moins un nombre $ecartmin$ de périodes plus tard que la fin de l'activité a_1.

$$\sum_{j_1 \in J} \sum_{p_1 \in P} (6(j_1 - 1) + p_1)\, x_{a_1g_1j_1p_1} + ecartmin_{a_1a_2} \leq \sum_{j_2 \in J} \sum_{p_2 \in P} (6(j_2 - 1) + p_2)\, x_{a_2g_2j_2p_2} \qquad \forall (a_1, a_2) \in AR, a_1, a_2 \in D_1, (g_1, g_2) \in G_{a_1a_2} \tag{3}$$

Une contrainte essentielle est de s'assurer qu'*un étudiant suive au maximum une activité par période*. Cela signifie que les groupes ayant des étudiants en commun ne peuvent être placés à la même période. C'est la contrainte qui forme l'horaire, c'est-à-dire qui force l'étalement des activités tout au long des 15 semaines.

$$x_{a_1g_1jp} + x_{a_2g_2jp} \leq 1 \qquad \forall (g_1, g_2) \in G_{a_1a_2} \tag{4}$$

Ce sont des contraintes à deux termes, donc légères et faciles à évaluer, mais le nombre de contraintes générées est de l'ordre des millions.

D'où, même si la taille du problème a été réduite par la nouvelle définition de la variable, aucun horaire n'est encore trouvé après deux jours de résolution avec le logiciel CPLEX (notons que c'est sans les contraintes reliées aux enseignants).

Étant donné que l'ajout de la contrainte (4) augmente considérablement la taille du problème, un compromis a été de générer cette contrainte uniquement pour les combinaisons de groupe de 18, 36 et 72 étudiants, tout en ajoutant une contrainte supplémentaire qui limite le nombre d'étudiants par période à 72 (taille de la cohorte). D'autres contraintes ne sont pas prises en considération, notamment les contraintes sur les ressources et les locaux. Bien que le temps de résolution a été réduit à 4h, il n'est pas raisonnable pour la qualité de l'horaire produite : fréquence élevée d'étudiants suivant plus d'une activité en même temps.

Analyse des résultats

Tableau 1 : Comparaison des deux approches

	Modèle mathématique	Constraint programming
Nombre de variables	119 507	1424
Nombre de contraintes	> 5 050 432	8847
Temps de résolution (minutes)	> 240	0,58

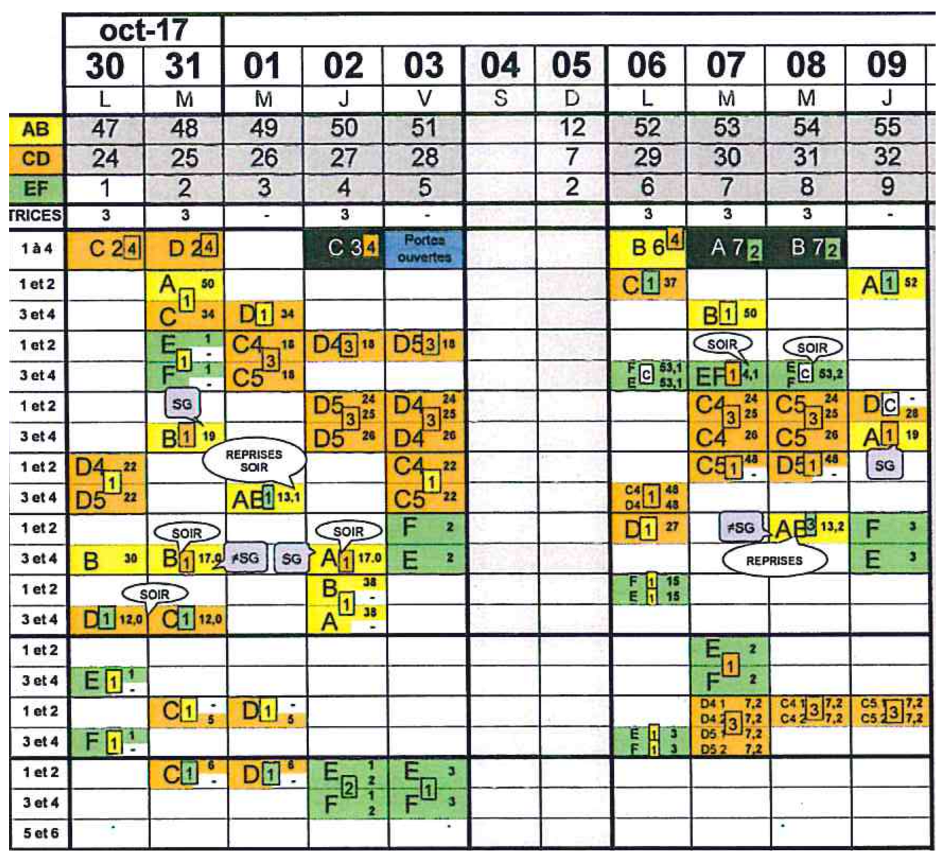

Figure 1 : Horaire construit à la main

Conclusion

Le temps de résolution est très élevé avec la résolution à partir du modèle mathématique. Cela pourrait s'expliquer par la fonction objectif qui ne permet pas suffisamment de restreindre l'espace de recherche. L'approche par constraint programming semble être une alternative avantageuse dans les problèmes d'ordonnancement et de planification de grande taille.

La prochaine étape est de retravailler l'horaire généré pour respecter les règles de convention collective (par exemple, un enseignant peut travailler maximum 4 périodes par jour).