

Répartition des PIFE 5A

Rendu n ° 3

Johan Brunet, Tristan Rivière

IG4 -G2,

Polytech Montpellier

25 octobre 2017

1 Prélude

1.1 Contexte & Objectifs

Au cours de leur formation, les étudiants en Informatique et Gestion de Polytech sont amenés à participer à un projet industriel par groupe de 2 à 3 élèves.

Une liste de projets leur est proposée et les élèves doivent constituer des groupes et choisir un sujet. Chaque sujet ne peut être choisi que par un seul groupe, il ne peut y avoir que 18 groupes car il n'y a que 18 projets, et chaque élève ne peut appartenir qu'à un seul groupe.

Nous chercherons ici une méthode de constitution des groupes qui soit à la fois équitable, satisfaisante, stable, non manipulable et implémentable.

1.2 Sujet

L'objectif de ce papier est de formaliser les caractéristiques auxquelles cette méthode devra pouvoir répondre.

2 Modélisation

2.1 Données

Promo Soit $Promo_n$ l'ensemble des élèves d'une même promo que l'on souhaite répartir tel que :

$$Promo_n = \{eleve_1, \dots, eleve_n\}, n = \text{taille de la promo 5A}$$

Projets Soit $Projets_p$ l'ensemble des projets tel que :

$$Projets_p = \{projet_1, \dots, projet_p\}, p = \text{nombre de projets maximum}$$

Répartition Soit $R(n, p)$ une partition $\sigma(n)$ de $Promo_n$ tel que :

$$|\sigma| \leq 18, \forall g \in \sigma, 2 \leq |g| \leq 3$$

Si il y a $n \leq 36$ élèves, on a deux cas de figure :

- si n pair, on a $\frac{n}{2}$ binômes : $\forall x \in \{1, \dots, \frac{n}{2}\}, g_x = \{e_i, e_j\}, \forall i, j \in n, i \neq j$
- si n impair, on a $\lfloor \frac{n}{2} \rfloor^1$ groupes, dont $\lfloor \frac{n}{2} \rfloor - 1$ binômes et 1 trinôme :
 $\forall x \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}, g_x = \{e_i, e_j\}, \forall i, j \in n, i \neq j$
 soit $i = \lfloor \frac{n}{2} \rfloor, g_i = \{e_k, e_l, e_m\}, \forall k, l, m \in n \setminus \{i, j\}, k \neq l \neq m$

Si il y a $n > 36$ élèves, on a $n - 36$ trinômes et $18 - (n - 36)$ binômes :
 $\forall x \in \{1, \dots, i\} \mid i = n - 36, g_x = \{e_j, e_k, e_l\}, \forall j, k, l \in n, j \neq k \neq l$
 et $\forall y \in \{i + 1, \dots, n\} \mid n = 18, g_y = \{e_f, e_h\}, \forall f, h \in n \setminus \{j, k, l\}, f \neq h$

Méthode Une méthode de répartition est une correspondance

$$\begin{aligned} &Meth\{(Promo_n, Projets_p) \mid n \in \{1, \dots, 50\}, p \in \text{Entier}^*\} \\ &\rightarrow \{R(n, p) \mid n \in \{1, \dots, 50\}, p \in \text{Entier}^*\} \end{aligned}$$

Mentions Soit M l'ensemble des mentions tel que :

$$M = \{TB, B, AB, P, I, AR\}$$

Les mentions TB, B, AB, P, I, AR traduisent respectivement les préférences "Très bien", "Bien", "Assez bien", "Passable", "Insuffisant", "A rejeter".

Les mentions suivent la relation d'ordre suivante sur M :

$$TB > B > AB > P > I > AR$$

2.2 Méthode

Fonctions On définit une fonction $PREF_{projet}(e_i, p_j)$ qui associe à un élève ses préférences de projets, ainsi qu'une fonction $PREF_{eleve}(e_i, e_j)$ qui associe à un élève ses préférences concernant les autres élèves. Ces préférences sont données par les mentions attribuées (voir le paragraphe "Mentions"). On cherchera à maximiser l'une ou l'autre des fonctions afin d'avoir une répartition équitable et satisfaisante pour le plus d'élèves possibles. On a donc :

$$\begin{cases} PREF_{projet}(e_i, p_j) : E \times P \longrightarrow M \\ PREF_{eleve}(e_i, e_j) : E \times E \longrightarrow M \end{cases}$$

1. $\lfloor \frac{n}{2} \rfloor$ est la partie entière de $\frac{n}{2}$

Bulletins Les préférences seront données par les étudiants à l'aide de bulletins, présentés comme suit :

Bulletin de préférence élève-élève :

e_i	TB	B	AB	P	I	AR
e_1	X					
\vdots
e_i						
\vdots
e_j					X	
\vdots
e_n		X				

On a donc les préférences élève-élève suivantes :

$$\begin{cases} PREF_{eleve}(e_i, e_1) = TB \\ PREF_{eleve}(e_i, e_j) = I \\ PREF_{eleve}(e_i, e_n) = B \end{cases}$$

Bulletin de préférence élève-projet :

e_i	TB	B	AB	P	I	AR
p_1		X				
p_2	X					
p_3				X		
\vdots
p_n						X

On a donc les préférences élève-projet suivantes :

$$\begin{cases} PREF_{projet}(e_i, p_1) = B \\ PREF_{projet}(e_i, p_2) = TB \\ PREF_{projet}(e_i, p_3) = P \\ PREF_{projet}(e_i, p_n) = AR \end{cases}$$

On peut alors définir des matrices de préférences.

Matrice de préférences élèves-élèves :

	e_1	...	e_i	...	e_j	...	e_n
e_1		...	TB	...	AB	...	B
\vdots
e_i	P	I	...	TB
\vdots
e_j	B	...	B	AR
\vdots
e_n	TB	...	AB	...	B	...	

Matrice de préférences élèves-projets :

	p_1	p_2	p_3	...	p_n
e_1	B	TB	I	...	B
\vdots
e_i	B	AB	TB	...	AR
\vdots
e_n	P	P	TB	...	B

Résultat le résultat de notre méthode nous donnera une Répartition $R(n,p)$ représenté par une matrice binaire :

$$\sigma \in M_{n,p}(\{0,1\})$$

$\Omega_{ij} = 1$ si le projet j est affecté au groupe de l'élève i .
 $\Omega_{(n,p)}(e_{leve_i}) = projet_j$

3 Dénombrements

Cette partie servira à dénombrer les nombres des combinaisons possibles pour les répartitions élèves-élèves ainsi que pour les répartitions groupes-projets.

Répartition Soit $R(n,p)$ une partition $\sigma(n)$ de $Promo_n$ tel que :

$$|\sigma| \leq 18, \forall g \in \sigma, 2 \leq |g| \leq 3$$

Si il y a $n \leq 36$ élèves, on a deux cas de figure :

- si n pair, on a $\frac{n}{2}$ binômes : $\forall x \in \{1, \dots, \frac{n}{2}\}, g_x = \{e_i, e_j\}, \forall i, j \in n, i \neq j$
- si n impair, on a $\lfloor \frac{n}{2} \rfloor^2$ groupes, dont $\lfloor \frac{n}{2} \rfloor - 1$ binômes et 1 trinôme :
 $\forall x \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}, g_x = \{e_i, e_j\}, \forall i, j \in n, i \neq j$
soit $i = \lfloor \frac{n}{2} \rfloor, g_i = \{e_k, e_l, e_m\}, \forall k, l, m \in n \setminus \{i, j\}, k \neq l \neq m$

Si il y a $n > 36$ élèves, on a $n - 36$ trinômes et $18 - (n - 36)$ binômes :
 $\forall x \in \{1, \dots, i\} \mid i = n - 36, g_x = \{e_j, e_k, e_l\}, \forall j, k, l \in n, j \neq k \neq l$
et $\forall y \in \{i + 1, \dots, n\} \mid n = 18, g_y = \{e_f, e_h\}, \forall f, h \in n \setminus \{j, k, l\}, f \neq h$

Méthode Une méthode de répartition est une correspondance

$$\begin{aligned} &Meth\{(Promo_n, Projets_p) \mid n \in \{2, \dots, 54\}, p \in Entier^*\} \\ &\rightarrow \{R(n, p) \mid n \in \{2, \dots, 54\}, p \in Entier^*\} \end{aligned}$$

Nombre de répartitions Le nombre de partitions de $G(n)$, noté $|G(n)|$ est le nombre de combinaisons possibles pour la répartition des élèves dans des groupes. Le nombre pour un nombre maximal de 54 élèves on a que des trinômes, le nombre de partitions est donc :

$$C_{54}^3 = 24804$$

Le nombre de répartitions possibles $R(n, p)$, noté $|R(n, p)|$ est donné en multipliant le nombre de partitions de $G(n)$ (nombre de groupes que l'on peut former) par le nombre de répartitions possibles des projets. Or pour chaque projet p , on peut l'associer avec

2. $\lfloor \frac{n}{2} \rfloor$ est la partie entière de $\frac{n}{2}$

chacune des partitions de $G(n)$. Le nombre de répartitions des projets possibles est donc $p! = 18!$.

Ce qui nous donne :

$$|R(n, p)| = |G(n)| * p! = 24804 * 18! \approx 1,588 * 10^{20}$$

Transitivité Soient $g_{ij} \in G(n)$ et $g_{jk} \in G(n)$.

Si $g_{ij} = 1$ (i est associé à j) et $g_{jk} = 1$ (j associé à k), alors on a $g_{ik} = 1$. De la même manière, Si $g_{ij} = 0$ (i n'est pas associé à j) et $g_{jk} = 0$, alors on a $g_{ik} = 0$.

4 Propositions

4.1 Constat sur le présent

Afin de construire notre méthode, nous avons tout d'abord cherché des méthodes mises en oeuvre dans le passé afin de mettre en évidence leurs faiblesses et ainsi savoir quels pièges éviter lors de la construction de celle-ci.

Cas 1 L'une des idées précédemment mise en avant était celle visant à la minimisation des mentions AR, I et P, ou bien à la maximisation du nombre de mentions TB et B. Une telle méthode a récemment été employée pour la répartition d'élèves d'IG4 dans différentes équipes d'un projet. Chaque étudiant définissait ses préférences à l'aide d'un bulletin similaire à celui employé ici, puis l'algorithme détermine la répartition ainsi :

On plaçait un maximum de personnes dans leur choix TB, puis on regardait si on pouvait placer les élèves restant sur l'une de leurs mentions bien. On effectuait ensuite un maximum d'échanges visant à faire passer les élèves n'ayant eu qu'une mention Bien vers une mention TB sans pour autant faire passer un autre élève de TB vers B. Si cette méthode semble au premier abord viable, une étude en profondeur de celle-ci nous permet de mettre en avant l'importance de construire une méthode équitable, et non manipulable.

Nous avons remarqué les problèmes suivants : - Si un bulletin ne comprend qu'une seule mention TB et un ensemble de mentions AR, alors le choix associé à cette mention TB sera forcément accordé car il fait monter le nombre de mentions TB ou B (et ceci au détriment d'un autre élève qui pourra se retrouver sur son second choix si il est associé à B au minimum. On a alors un soucis de manipulabilité. - Dans la même idée, on peut décider de ne mettre aucune mention TB mais seulement des mention Bien au maximum. L'algorithme va alors vouloir accomplir ses vœux car ils sont le signe que l'élève n'a pas obtenu ses premiers vœux !

A l'observation de ces problèmes, nous en déduisons l'importance que l'algorithme ne cherche pas la maximisation des meilleures mention (ou la minimisation des moins bonnes)).

Cas 2 Une idée mise en avant pour résoudre le soucis de la manipulation était de mettre en place un traitement des bultin en amont afin de les normaliser. On définirait ainsi un bulletin idéal constitué par exemple de 2 mention TB, 4B, 150 AS, 15 P, 5I, et 2AR.

De la on transformerait les bulletins afin de leur faire prendre la forme de ce bulletin idéal en transformant les surplus de chaque mention haute en une mention inférieur, et inversement les surplus des mention inférieur en des mentions supérieur. Une personne ayant alors pris la peine de nuancer ses choix ne verra son bulletin que très peu modifié. Une personne ayant décidé au contraire de ne mettre que des mentions TB et AR se verra son bulletin fortement changés sans pour autant modifier la satisfaction de celui-ci résultante à l'algorithme. Si il avais au départ 5TB, même si après transformation 3 d'entres elles se retrouvent avec la mention Bien, cela n'aura pas d'effet sur sa satisfaction car il à lui même choisi qu'il affectionnait autant ces 5 personnes. On supprimerait ainsi uniquement le risque de manipulation de la méthode. Néanmoins un tel système présente deux gros problèmes. Tout d'abord, prenons l'exemple de quelqu'un n'ayant mis que des TB : l'algorithme va tout de même chercher à le mettre avec les personnes qui sont rester dans des mention hautes suite à la transformation. Pourtant cette personne n'avait au départ aucune préférence et on aurait surement pu constituer de meilleurs groupes sans cette transformation

On ne peux donc pas non plus chercher à faire une méthode qui traite les bulletins en amont afin d'éviter les manipulations.

Conclusion Suite à ces observations, nous pouvons ainsi éliminer un certains nombre de pistes et proposer une idée de méthode évitant certains pièges...

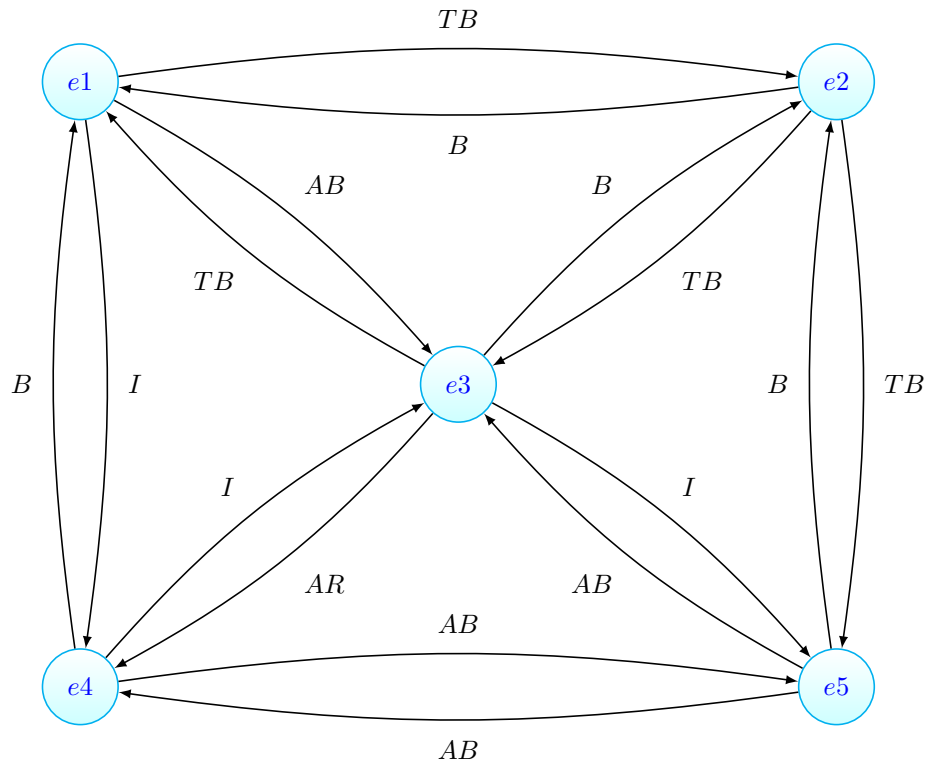
5 Algorithme

5.1 Initialisation

On considère la matrice de préférences élèves-élèves suivante :

	<i>e1</i>	<i>e2</i>	<i>e3</i>	<i>e4</i>	<i>e5</i>
<i>e1</i>	—	<i>TB</i>	<i>AB</i>	<i>I</i>	<i>B</i>
<i>e2</i>	<i>B</i>	—	<i>TB</i>	<i>B</i>	<i>TB</i>
<i>e3</i>	<i>TB</i>	<i>B</i>	—	<i>AR</i>	<i>I</i>
<i>e4</i>	<i>B</i>	<i>TB</i>	<i>I</i>	—	<i>AB</i>
<i>e5</i>	<i>TB</i>	<i>B</i>	<i>AB</i>	<i>AB</i>	—

Pour la phase d'initialisation, nous représentons les préférences sous forme de graphe. Chaque élève correspond à un sommet, et chaque mention correspond à une arête. On obtient alors un graphe similaire à celui ci-dessous (toutes les arrêtes n'ont pas été représentées pour des raisons de lisibilité).



A partir de ce graphe, nous essayons de former des groupes selon les critères suivants :

- le premier sommet du graphe correspond au premier élève dans l'ordre alphabétique
- on recherche pour ce sommet les autres sommets avec lesquels il partage deux arrêtes portant la valeur "TB"
- lorsque deux tels sommets sont trouvés, ils sont associés et les élèves correspondants sont placés dans le même groupe
- ensuite on continue à appliquer cette méthode pour les sommets restants.

Lorsqu'il n'y a plus de sommets partageant deux arrêtes "TB", on change de critère de sélection. Ceux-ci sont les suivants :

- TB - TB
- TB - B
- B - B
- TB - AB
- B - AB
- AB - AB

Dans le cas où l'on n'aurait pas traité tous les sommets après avoir utilisé le critère "AB - AB", on continue avec une mention inférieure.

Lorsque tous les sommets ont été traités, des groupes sont constitués avec les associations de sommets faites au fur et à mesure des étapes. Ce premier algorithme

permet de constituer une première répartition des groupes d'élèves, mais qui ne sera pas forcément optimale.

5.2 Optimisation

De là, on essaye de modifier cette répartition afin de maximiser la satisfaction globale des élèves. On définira la satisfaction global par une satisfaction maximale pour tous les groupes. Par exemple si un groupe G1 a pour satisfaction totale $S1 = \{2TB, 1B\}$ mais qu'un autre groupe G2 a pour satisfaction $S2 = \{1TB, 2AB\}$, on cherche à augmenter la satisfaction du G2. Dans le cas où un "AB" du G2 deviendrait un "B" dans le G1 et le "B" du G1 deviendrait un "TB" dans le G2, la nouvelle satisfaction par groupe serait : $S1 = \{2TB, 1B\}$ et $S2 = \{2TB, 1AB\}$. Pour cette répartition on obtient une meilleure satisfaction totale.

L'algorithme s'arrête lorsque lorsqu'une permutation la satisfaction est égale ou inférieure à la satisfaction précédente.