COURS 15

Version du 27 mars 2025.

Rappel. La dernière fois, nous avons reparlé un peu de la variante où on a le droit de choisir ses plantes d'intérieur, pour souligner que l'analyse que nous avons fait la fois d'avant était correct. J'ai mentionné que Knuth analyse, lui, une variante où chacun a une liste restreinte de personnes avec qui il est prêt à se marier (sans donner davantage de détails). L'idée est différente, car on suppose toujours que tout le monde sera marié. Il n'est pas surprenant que, avec cette limitation, il soit possible qu'il n'y ait aucun système stable possible. Knuth explique comment modifier Gale-Shapley pour déterminer s'il existe un mariage stable sous cette contrainte supplémentaire (mais je préfère passer à autre chose).

Nous avons obtenu une borne supérieure sur le nombre d'offres qui seront fait en moyenne dans l'algorithme Gale-Shapley, si les préférences sont aléatoires. Le déroulement était tout à fait controlé par une suite aléatoire des femmes, chacun prise uniformément au hasard, et l'algorithme se termine au moment que la dernière femme apparaît sur la liste. La borne est donc donné par la longueur attendu d'une liste avec les entrées prises uniformément au hasard, avant qu'on arrive à voir toutes les n entrées possibles.

2.10. **Borne supérieure sur le nombre d'offres qui seront faites en moyenne, *suite*.** Comme rappel, et exemple, considérons le déroulement de l'algorithme avec 4 hommes et 4 femmes, et supposons que chacune des femmes préfèrent $h_1 > h_2 > h_3 > h_4$. Soit la suite aléatoire de femmes la suivante :

$$4, 3, 4, 2, 3, 4, 3, 1$$

h_1 fait une offre à femme f_4. h_2 fait une offre à f_3. h_3 fait une offre à f_4 mais est refusé, donc il fait une offre à f_2. h_4 fait une offre à f_3 mais est refusé, donc il fait une offre à f_4, mais il est refusé, il fait une offre "redondante" à f_3 (c'est une offre qui ne se ferait pas dans Gale–Shapley), avant de faire une offre à f_1.

Nous "déduisons" les préférences suivantes :

$$h_1 : f_4$$
$$h_2 : f_3$$
$$h_3 : f_4 > f_2$$
$$h_4 : f_3 > f_4 > f_1$$

Maintenant, on considère le déroulement si les femmes préfèrent

$$h_2 > h_4 > h_3 > h_1.$$

(J'ai fait ce choix pour une raison particulière.)

h_1 fait une offre à f_4. h_2 fait une offre à f_3. h_3 fait une offre à f_4, qui le préfère, et c'est donc h_1 doit faire une offre à f_2. Maintenant h_4 fait une offre à f_3, rejetée, et en suite à f_4, qui le préfère. Elle rejette h_3, qui doit faire une offre à f_3 rejetée, avant de faire une offre à f_1.

$$h_1 : f_4 > f_2$$
$$h_2 : f_3$$
$$h_3 : f_4 > f_3 > f_1$$
$$h_4 : f_3 > f_4$$

Nous constatons que cette fois, aucune des offres n'était redondante, mais dans les deux cas (indépendamment des préférences des femmes) l'algorithme (avec les offres redondantes) s'est déroulé avec 8 offres, avant l'apparition de la femme finale, ce qui signale que tous les femmes ont reçu une offre, et donc que tout le monde peut se marier.

La chose qu'il nous restait de la dernière fois, était de voir combien de temps il prend en moyenne pour que toutes les femmes apparaissent sur la liste.

Supposons qu'il y a n options, et nous en avons déjà vu i parmi ces options. Alors le temps qu'il nous faudra pour voir l'$i + 1$-ième, c'est

$$
\begin{aligned}
E &= 1 \cdot \frac{n-i}{n} + 2 \cdot \frac{i}{n}\frac{n-i}{n} + 3 \cdot \frac{i}{n}\frac{i}{n}\frac{n-i}{n} + \dots. \\
\frac{i}{n}E &= \qquad\qquad 1 \cdot \frac{i}{n}\frac{n-i}{n} + 2 \cdot \frac{i}{n}\frac{i}{n}\frac{n-i}{n} + \dots
\end{aligned}
$$

et donc

$$
\begin{aligned}
\frac{n-i}{n}E &= 1 \cdot \frac{n-i}{n} + 1 \cdot \frac{i}{n}\frac{n-i}{n} + 1 \cdot \frac{i}{n}\frac{i}{n}\frac{n-i}{n} + \dots \\
E &= 1 + \frac{i}{n} + \frac{i}{n}\frac{i}{n} + \dots \\
\frac{i}{n}E &= \quad + \frac{i}{n} + \frac{i}{n}\frac{i}{n} + \dots
\end{aligned}
$$

et donc

$$
\begin{aligned}
\frac{n-i}{n}E &= 1 \\
E &= \frac{n}{n-i}.
\end{aligned}
$$

On prend maintenant la somme pour $i = 0, \dots, n-1$, et on obtient un temps moyen d'attente de $n(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{1})$. Comme nous avons vu la dernière fois, c'est très proche de l'intégrale $\int_1^{n+1} \frac{n}{x} dx = n \ln(n+1) \approx n \ln n$.

En moyenne, le nombre d'offres que devra faire chaque homme est borné par $\ln n$ — beaucoup moins que le nombre total des femmes.

3. Combinatoire des tableaux

Ma source principale sera le troisième chapitre du livre de Sagan, “The symmetric group,” même si nous n'allons pas vraiment aborder les liens avec le groupe symétrique.

3.1. **Définitions initiales.** Soit $\lambda = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ une partition de n, c'est-à-dire que $\lambda_1 + \lambda_2 + \dots + \lambda_r = n$.

Le diagramme de Ferrers est dessiné avec λ_1 cases sur la première ligne, λ_2 sur la deuxième ligne, et ainsi de suite. Chaque fois, les cases sont jusitifiées à gauche. J'utiliserai la convention française (ou franco-californienne) selon laquelle la première ligne est en bas.

Un tableau de Young standard est un remplissage des n cases par les nombres de 1 à n, avec chaque entrée apparaît une fois, et où les entrées sont croissantes quand on monte dans une colonne où si on lit de gauche à droite dans une ligne.

On peut aussi voir un diagramme de Young standard comme une chaîne de relations de couverture dans le poset de Ferrers : c'est un poset sur les partitions, où $\mu > \lambda$ is les cases de mu contiennent les cases de λ. (D'ailleurs, c'est encore un treillis distributif. Exercice : comment décrire $\vee$ et $\wedge$?)

Un tableau T correspond à la chaîne $\emptyset <$ forme$(T_{\leq 1}) <$ forme$(T_{\leq 2}) \dots$.

Un tableau (de Young) semistandard est un remplissage où les entrées sont strictement croissantes dans les colonnes et faiblement croissantes dans les lignes.

Un tableau croissant est un remplissage où les entrées sont strictement croissantes dans les lignes et les colonnes (mais il peut y avoir des répétitions et des sauts). Un tableau standard est un cas très particulier d'un tableau croissant.

3.2. **Correspondance de Robinson–Schensted.** C'est une correspondance entre les permutations de n et les paires de tableaux standards de la même forme, et de taille n.

La correspondance marche de cette manière : je lis les entrées de π, et je les insère un par un dans un tableau standard, P. À chaque fois, je l'insère dans la première ligne, à l'endroit où il convient pour que la ligne reste croissante. Si c'est à la fin de la ligne, l'insertion a réussi et j'ai fini. Sinon, l'insertion tasse une entrée de la première ligne (la première qui lui est supérieure), qui doit être à son tour insérée dans la deuxième ligne, et ainsi de suite.

Chaque insertion produit un tableau (semi)standard. (Ce n'est pas standard, parce que les entrées ne sont pas les entiers de 1 à n.) Ce n'est pas totalement évident. L'ajout à la première ligne ne pose pas de problème, mais lorsqu'on tasse une entrée et on doit la réinsérer dans la ligne d'après, on pourrait se questionner si ça peut poser problème. L'effet de l'insertion est toujours que l'entrée dans une case particulière diminue, donc la question serait : est-ce que l'insertion peut causer problème avec l'élément sur la ligne d'en dessous. La réponse est non : l'élément sera forcément réinséré faiblement à gauche de sa position dans la ligne d'en dessous, ce qui fait qu'il sera plus grand que l'entrée en dessous.

Cela définit le tableau P (on dit ≪ le tableau d'insertion ≫). Le deuxième tableau enregistre la croissance du tableau P : on met 1 dans la première case, 2 dans la deuxième, et ainsi de suite. Ça donne automatiquement un tableau standard. C'est le ≪ tableau d'enregistrement ≫.

La correspondance de Robinson–Schensted envoie un permutation π à la paire (ordonnée) (P, Q).

Théorème 3.2.1. *La correspondance de Robinson–Schensted est une bijection entre les permutations de n et les paires de tableaux (P, Q) de la même forme, de taille n.*

Démonstration. Pour démontrer le théorème, il suffit de donner une inverse. L'inverse est construite de cette manière. Dans Q, on retrouve n. Sa position indique la dernière case qui a été remplie dans P. Ça veut dire que c'est dans cette case (forcément à la fin de la ligne) que nous avons finalement pu insérer sans tasser un autre entrée. Mais ça ne veut pas dire que c'est cette entrée là qu'on a commencer d'insérer ! Cette entrée-là était tassée de la ligne d'avant. Par quelle entrée ? Par la plus grande entrée dans cette ligne qui lui est inférieure. On continue comme ça, remontant le tableau, inversant de façon graduelle l'insertion qui a fini par ajouter une entrée dans la case n du tableau Q. Lorqu'on arrive à la première ligne, on détermine quelle entrée, actuellement là-dedans, a été insérée. C'est donc la dernière entrée de π.

Nous avons déterminé l'état des lieux avant cette insertion (le n de Q n'est plus présent, l'insertion dans P est inversée), et nous procédons de la même manière pour la case $n-1$ dans Q. Puisque, pour n'importe quelle paire de tableau, on peut trouver l'unique permutation envoyée à la paire par Robinson–Schensted, c'est une bijection. $\square$