Correction des exercices du cours: Structures et Algorithmes Aléatoires
cours 1 du 16 octobre 2009.

1. Vous avez 3 pièces dont une est biaisée: vous savez qu'elle tombe sur pile avec probabilité 2/3. Vous ne savez pas quelle pièce est biaisée mais vous faites un tirage avec chacune des pièces et la première et seconde pièces tombent sur pile tandis que la troisième tombe sur face. Quelle est la probabilité que la première pièce soit celle qui est biaisée?

   - Soit E_i l'événement: la i-ème pièce est celle qui est biaisée et B l'événement: le tirage des trois pièces donne dans l'ordre: pile-pile-face. Avant le tirage, nous n'avons aucune information sur la pièce biaisée donc $\mathbb{P}(E_i) = 1/3$. Nous avons aussi

   $$\mathbb{P}(B|E_1) = \mathbb{P}(B|E_2) = \frac{2}{3}\frac{1}{2}\frac{1}{2} = \frac{1}{6}, \text{ et, } \mathbb{P}(B|E_3) = \frac{1}{12}.$$

   On a donc

   $$\mathbb{P}(E_1|B) = \frac{\mathbb{P}(B|E_1)\mathbb{P}(E_1)}{\sum_{i=1}^{3}\mathbb{P}(B|E_i)\mathbb{P}(E_i)} = \frac{2}{5}.$$

   Après la réalisation du tirage, la vraisemblance pour que la première pièce soit celle qui est biaisée est passée de 1/3 à 2/5.

2. Dans le cas vu en cours de la multiplication matricielle. Sans information sur A, B, C, on fait l'hypothèse a priori que $\mathbb{P}(AB = C) = 1/2$. On fait tourner une fois l'algorithme vu en cours qui nous retourne le résultat: $AB = C$. Avec cette nouvelle information quelle est la probabilité a posteriori que l'identité soit correcte? Et après k itérations de l'algorithme ?

   - Le raisonement est similaire au cas précédent. Soit E l'événement: l'identité est correcte. et F l'événement: l'algorithme retourne $AB = C$. On commence avec $\mathbb{P}(E) = \mathbb{P}(\overline{E}) = 1/2$. De plus, $\mathbb{P}(F|E) = 1$ tandis que $\mathbb{P}(F|\overline{E}) \leq 1/2$ (d'après le résultat vu en cours). On a donc

   $$\mathbb{P}(E|F) = \frac{\mathbb{P}(F|E)\mathbb{P}(E)}{\mathbb{P}(F|E)\mathbb{P}(E) + \mathbb{P}(F|\overline{E})\mathbb{P}(\overline{E})} \geq \frac{2}{3}.$$

   Après le premier test, on suppose maintenant que $\mathbb{P}(E) \geq 2/3$ et $\mathbb{P}(\overline{E}) \leq 1/3$, on a donc apres un second test:

   $$\mathbb{P}(E|F) \geq \frac{2/3}{2/3 + 1/3\dot{1}/2} = \frac{4}{5}.$$

   On trouve après k itérations

   $$\mathbb{P}(E|F) \geq 1 - \frac{1}{2^k + 1}.$$

3. Une coupe dans un graphe est un ensemble d'arêtes qui une fois retiré rend le graphe déconnecté. Une coupe minimale est une coupe de cardinalité minimale. La contraction d'une arête $\{u, v\}$ consiste à rassembler u et v en un seul sommet en éliminant toutes les arêtes entre u et v mais en gardant toutes les autres arêtes du graphe. Le graphe ainsi obtenu peut contenir des arêtes parallèles mais pas de boucle sur un sommet. Vérifier qu'une coupe d'un graphe obtenu après une contraction d'arête est encore une coupe du graphe original. On considère alors l'algorithme qui consiste en $n-2$ étapes où n est le nombre de sommets du graphe et qui à chaque itération, choisit une arête uniformément au hasard et la contracte. A la fin des $n-2$ itérations, le graphe obtenu est un graphe à deux sommets et l'algorithme retourne l'ensemble des arêtes connectant ces deux somments. Montrer que cet algorithme retourne une coupe minimale avec probabilité au moins $2/n(n-1)$.

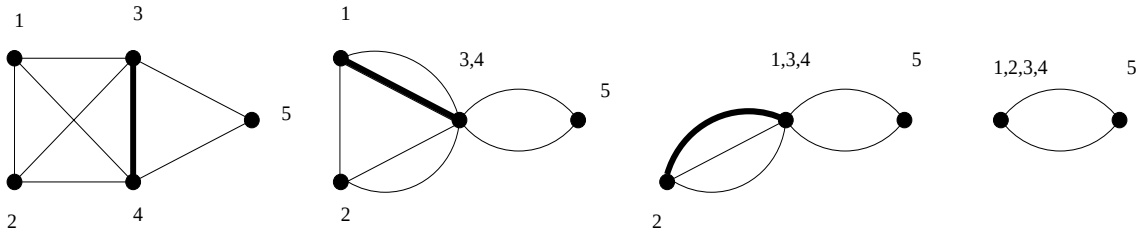

Figure 1: Un exemple où l'algorithme retourne une coupe minimale (à chaque étape, l'arête contractée est marquée).

- Soit k la taille d'une coupe minimale de G et soit C une telle coupe minimale. Nous calculons la probabilité que l'algorithme trouve C.

  Si l'algorithme ne choisit jamais d'arêtes de C lors des $n-2$ itérations, alors il retourne C. Soit E_i l'événement: l'arête contractée lors de la i-ème itération n'est pas dans C et soit $F_i = \cap_{j=1}^{i} E_j$. Nous devons calculer $\mathbb{P}(F_{n-2})$. Nouc commenç ons par:

$$\mathbb{P}(E_1) = \mathbb{P}(F_1) \geq 1 - \frac{2k}{nk},$$

  car chaque sommet a au moins degré k et donc le graphe a au moins $nk/2$ arêtes. En conditionant par F_1, on se retrouve après la première itération avec un graphe de $n-1$ sommets et de coupe minimale de taille k. On a donc

$$\mathbb{P}(E_2|F_1) \geq 1 - \frac{2k}{k(n-1)}.$$

  Finalement, en itérant:

$$\begin{aligned} \mathbb{P}(F_{n-2}) &= \mathbb{P}(E_{n-2}|F_{n-3})\mathbb{P}(F_{n-3}) \\ &\geq \prod_{i=1}^{n-2}\left(1 - \frac{2}{n-i+1}\right) = \frac{2}{n(n-1)}. \end{aligned}$$

4. Montrer qu'il existe une infinité de nombres premiers du type $6n+5$ (résultat utilisé dans la dernière preuve donnée en cours).

   - On suppose par l'absurde qu'ils sont en nombre fini: $\{p_1, p_2, \ldots, p_k\}$ et on considère $p = 6p_1p_2\ldots p_k - 1$. Comme $p \equiv 5\,[6]$, on peut écrire $p = 6n+5$ et si p n'est pas premier alors $p = uv \equiv 5\,[6]$ et donc soit u soit v appartient à $\{p_1, p_2, \ldots, p_k\}$ ce qui est contradictoire.

5. Voici une preuve non probabiliste du théorème vu en cours sur les ensembles dominants d'un graphe G de degré minimal $\delta > 1$. Pour chaque sommet v, on note $C(v)$ l'ensemble constitué de v et de ses voisins. On dit que v couvre $C(v)$. Un ensemble dominant pour le graphe $G = (V, E)$ est un ensemble U tel que $V = \cup_{u\in U} C(u)$. On considère l'algorithme déterministe qui à chaque étape choisit le sommet qui couvre un maximum de sommets qui ne sont pas encore couverts. Soit U_t l'ensemble des sommets choisis depuis le début jusqu'à l'étape t. Soit $r_t = |V \setminus \cup_{u\in U_t} C(u)|$. Montrer qu'à l'étape t, il existe un sommet v qui appartient à au moins $r_t(\delta+1)/n$ ensembles $C(w)$ avec $w \in V \setminus \cup_{u\in U_t} C(u)$. En déduire que $r_{t+1} \leq r_t(1-(\delta+1)/n)$. Retrouver le résultat donné en cours.

   - Soit $n(v)$ le nombre de $C(w)$ avec $w \in V \setminus \cup_{u\in U_t} C(u)$ qui couvre v. On a $\sum_{v\in V} n(v) = \sum_{v\in V\setminus\cup_{u\in U_t} C(u)} |C(v)| \geq r_t(\delta+1)$ et donc il existe $v \in V\setminus U_t$ tel que $n(v) \geq r_t(\delta+1)/n$. En rajoutant ce sommet à U_t, le nombre de sommets non couverts est au plus de $r_t(1-(\delta+1)/n)$, on a donc $r_{t+1} \leq r_t(1-(\delta+1)/n)$. Donc apres $n\ln(\delta+1)/(\delta+1)$ itérations, il restera au plus $n/(\delta+1)$ sommets non couverts que l'on peut alors rajouter pour former un ensemble dominant de taille au plus celle vue en cours.

6. Un hypergraphe est une paire $H = (V, E)$ où V est un ensemble fini de sommets et E est une famille de sous-ensembles de V appelés (hyper)arêtes. Un hypergraphe est n-uniforme si chaque arête contient exactement n sommets. On dit qu'un hypergraphe est 2-coloriable si il existe un 2-coloriage de V tel qu'aucune arête ne soit monochromatique. Soit $m(n)$ le nombre minimal possible d'arêtes d'un hypergraphe n-uniforme qui n'est pas 2-coloriable. Montrer que tout hypergraphe n-uniforme avec moins de 2^{n-1} arêtes est 2-coloriable, donc que $m(n) \geq 2^{n-1}$. On cherche maintenant une borne supérieure sur $m(n)$. On fixe V avec v points, v sera optimisé plus tard. Soit χ un coloriage de V en deux couleurs. Soit $S \subset V$ un ensemble de n points choisi uniformément. Montrer que $\mathbb{P}(S \text{ est monochromatique sous } \chi) \geq p$ avec $p = \frac{2\binom{v/2}{n}}{\binom{v}{n}}$. Soit $S_1, \ldots, S_m$ des ensembles de n points choisis de manière uniforme et indépendante. Pour chaque coloriage χ, soit A_χ l'événement: aucun des S_i n'est monochromatique. Montrer que $\mathbb{P}(\cup_\chi A_\chi) \leq 2^v(1-p)^m$ et donc que $m(n) \leq \lceil \frac{v\ln 2}{p} \rceil$. On doit donc trouver v qui minimise v/p. Pour ceci on utilisera l'approximation:

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2^{1-n}\prod_{i=0}^{n}\frac{v-2i}{v-i} \sim 2^{1-n}e^{-n^2/2v}.$$

   En déduire que $m(n) < (1+o(1))\frac{e\ln 2}{4}n^2 2^n$.

- Démonstration de $m(n) \geq 2^{n-1}$. On colore V de manière aléatoire. Pour chaque arête $e \in E$, soit A_e l'événement: e est monochromatique. On a $\mathbb{P}(A_e) = 2^{1-n}$ et donc

$$\mathbb{P}(\cup_{e\in E} A_e) \leq \sum_{e\in E}\mathbb{P}(A_e) < 1.$$

- On a
$$\mathbb{P}(S \text{ est monochromatique sous } \chi) = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}}.$$
On choisit v pair et donc cette expression est minimisée pour $a = b$.

- Par l'indépendance des S_i, on a
$$\mathbb{P}(A_\chi) \leq (1-p)^m.$$
Il y a 2^v coloriage donc
$$\mathbb{P}(\cup_\chi A_\chi) \leq 2^v(1-p)^m.$$
Quand cette quantité est plus petite que 1, il existe $S_1, \ldots, S_m$ tels que aucun des A_χ ne soit vrai. C'est à dire, $S_1, \ldots, S_m$ ne sont pas deux-coloriables et donc $m(n) \leq m$. Quand $m = \lceil \frac{v \ln 2}{p} \rceil$, on obtient: $2^v(1-p)^m \leq 2^v e^{-pm} \leq 1$.

- L'approximation est valable pour $v >> n^{3/2}$,
$$\frac{v-2i}{v-i} = 1 - \frac{i}{v} + O\left(\frac{i^2}{v^2}\right) = e^{-i/v + O(i^2/v^2)}.$$
La valeur optimale de v est $n^2/2$ et le choix de v pair demandera un changement d'au plus 2 qui sera asymptotiquement négligeable. On obtient donc le résultat souhaité.