TD 14 – Random walks (corrigé)

Exercice 1. *Cover time in graphs*

Given a finite, undirected non-bipartite and connected graph $G = (V, E)$, recall that the *cover time* of G is the maximum over all vertices $v \in V$ of the expected time to visit all of the nodes in the graph by a random walk starting from v.

1. Recall that $h_{v,u}$ is the expected number of steps to reach u from v and $h_{u,u} = \frac{2|E|}{d(u)}$. Show that

$$\sum_{w \in N(u)} (1 + h_{w,u}) = 2|E|.$$

☞ Let $N(u)$ be the set of neighbors of vertex u in G. We compute $h_{u,u}$ in two different ways :

$$h_{u,u} = \frac{2|E|}{d(u)}$$

and

$$h_{u,u} = \frac{1}{d(u)} \sum_{w \in N(u)} (1 + h_{w,u}).$$

Hence

$$2|E| = \sum_{w \in N(u)} (1 + h_{w,u}).$$

2. Let T be a *spanning tree* of G (i.e. T is a tree with vertex set V). Show that there is a *tour* (i.e. a walk with the same starting and ending points) passing each edge of T exactly twice, once for each direction.

☞ Following Depth-first Search.

3. Let $v_0, v_1, ..., v_{2|V|-2} = v_0$ be the sequence of vertices of such tour. Prove that

$$\sum_{i=0}^{2|V|-3} h_{v_i, v_{i+1}} < 4|V| \times |E|.$$

☞ From Q1, we have $h_{v,u} < 2|E|$ for any edge uv. The result trivially follows.

4. Conclude that the cover time of G is upper-bounded by $4|V| \times |E|$.

☞ $\sum_{i=0}^{2|V|-3} h_{v_i, v_{i+1}}$ is the expected time to go through the tour, i.e the expected time for any vertex v_i visit all nodes in the graph in a strict order. Hence it is longer than the expected time for any vertex v_i to visit all nodes in the graph without such order restriction, i.e. the cover time of G is at most $\sum_{i=0}^{2|V|-3} h_{v_i, v_{i+1}} < 4|V| \times |E|$.

Exercice 2. *$s - t$ connectivity in graph*

Suppose we are given a finite, undirected and non-bipartite graph $G = (V, E)$ and two vertices s and t in G. Let $n = |V|$ and $m = |E|$. We want to determine if there is a path connecting s and t. This is easily done in linear time using a standard breadth-first search or depth-first search. Such algorithms, however, require $\Omega(n)$ space. Here we develop a randomized algorithm that works with only $O(\log n)$ bits of memory. This could be even less than the number of bits required to write the path between s and t. The algorithm is simple :

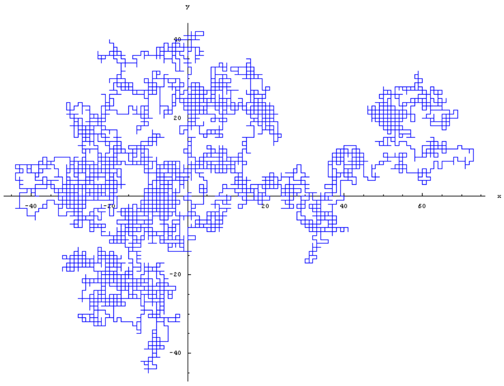
Perform a random walk from s on G. If the walk reaches t within $4n^3$ steps, return that there is a path. Otherwise, return that there is no path.

1. Show that the algorithm returns the correct answer with probability at least $1/2$, and it only makes errors by returning that there is no path from s to t when there is such a path.

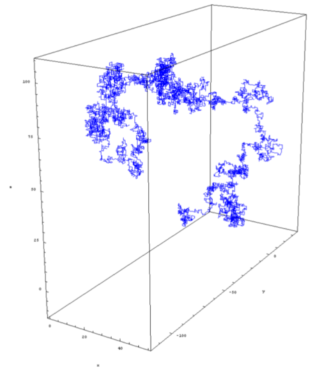
☞ If there is no path then the algorithm returns the correct answer. If there is a path, the algorithm errs if it does not find the path within $4n^3$ steps of the walk. The expected time to reach t from s (if there is a path) is bounded from above by the cover time of their shared component, which by Lemma 7.16 is at most $4nm < 2n^3$. By Markov's inequality, the probability that a walk takes more than $4n^3$ steps to reach s from t is at most $1/2$.

Exercice 3. *Random walks on $\mathbb{Z}^d$*

Consider the simple random walk over $\mathcal{Z}^d$ with probability $1/2d$ to jump towards any of the $2d$ neighbors in the grid. The walk is clearly irreducible.



A random walk in $\mathcal{Z}^2$
(10000 steps, Wikipedia)



A random walk in $\mathcal{Z}^3$
(10000 steps, Wikipedia)

1. For $d = 1$, whether it is recurrent ? positive recurrent ?

☞ [Déjà fait dans les TDs avant]

On pose $\overrightarrow{X_k} = (X_k, \dots, X_{k+l})$ et $\overrightarrow{i_k} = (i_k, \dots, i_{k+l})$.

On remarque que

$$\begin{aligned}
\mathbb{P}\left(\overrightarrow{X_{k+1}} = \overrightarrow{i_{k+1}} | \overrightarrow{X_k} = \overrightarrow{i_k}\right) &= \mathbb{P}\left(X_{k+l+1} = i_{k+l+1}, \dots, X_{k+1} = i_{k+1} | X_{k+l} = i_{k+l}, \dots, X_k = i_k\right) \\
&= \mathbb{P}\left(X_{k+l+1} = i_{k+l+1} | X_{k+l} = i_{k+l}, \dots, X_k = i_k\right) \\
&= \mathbb{P}\left(X_{k+l+1} = i_{k+l+1} | X_{k+l} = i_{k+l}, \dots, X_0 = i_0\right) \\
&= \mathbb{P}\left(\overrightarrow{X_{k+1}} = \overrightarrow{i_{k+1}} | \overrightarrow{X_k} = \overrightarrow{i_k}, \dots, \overrightarrow{X_0} = \overrightarrow{i_0}\right)
\end{aligned}$$

Grâce à l'étude des $\left(\overrightarrow{X_t}\right)_{t\in\mathbb{N}}$, on se ramène à une chaine de MARKOV.

Cas $d = 1$

Soit S_n la position de la marche au moment n et T_0 l'instant de premier retour en 0.

La probabilité $\mathbb{P}(T_0 < +\infty)$ de retour en 0 n'est rien d'autre que la somme $\sum_{n=1}^{+\infty} \mathbb{P}(T_0 = 2n)$. Et cette dernière somme est télescopique : on a pour tout entier m

$$\begin{aligned}
\sum_{n=1}^{m} \mathbb{P}(T_0 = 2n) = \sum_{n=1}^{m} \mathbb{P}(S_{2n-2} = 0) - \mathbb{P}(S_{2n} = 0) \\
&= \mathbb{P}(S_0 = 0) - \mathbb{P}(S_{2m} = 0) \\
&= 1 - \frac{\binom{2m}{m}}{2^{2m}}
\end{aligned}$$

Or lorsque m tend vers $+\infty$, le rapport $\frac{\binom{2m}{m}}{2^{2m}}$ tend vers 0 ce qui se justifie avec la formule de STIRLING :

$$\begin{aligned}
\frac{\binom{2m}{m}}{2^{2m}} &= \frac{(2m)!}{2^{2m}(m!)^2} \\
&\sim \frac{(2m)^{2m} e^{-2m} \sqrt{4\pi m}}{2^{2m}(m^m e^{-m} \sqrt{2\pi m})^2} \\
&= \frac{1}{\sqrt{\pi m}}
\end{aligned}$$

En particulier, le rapport tend vers 0 donc en passant à la limite :

$$\sum_{n=1}^{+\infty} \mathbb{P}(T_0 = 2n) = 1$$

2. For $d = 2$, whether it is recurrent ? positive recurrent ?

Hint : Consider decomposing the walk into two independent walks.

☞

Cas d=2

On change de coordonnées. On suppose qu'on était dans la base orthonormée $\mathbb{B} = (\overrightarrow{i}, \overrightarrow{j})$ et on fixe deux nouveaux vecteurs :

- $\overrightarrow{u} = \frac{1}{2}\left(\overrightarrow{i} + \overrightarrow{j}\right)$
- $\overrightarrow{v} = \frac{1}{2}\left(\overrightarrow{j} - \overrightarrow{i}\right)$

Et on appelle $\mathbb{B}' = (\overrightarrow{u}, \overrightarrow{v})$.

Aussi un point de coordonnées (x, y) dans la base $\mathbb{B}$ a des coordonnées $(i + j, j - i)$ dans la base $\mathbb{B}'$.

Les 4 mouvements possibles deviennent donc :

- $(+1, 0) \to (+1, -1)$
- $(-1, 0) \to (-1, +1)$
- $(0, +1) \to (+1, +1)$
- $(0, -1) \to (-1, -1)$

On obtient donc un produit cartésien de deux variables aléatoires indépendantes. On a donc deux marches aléatoires sur $\mathcal{Z}$ indépendantes. C'est donc une marche aléatoire récurrente.

3. In the case $d = 3$, for every n, show that

$$\mathbb{P}(S_{2n} = 0) = \sum_{r+s+t=n} \binom{2n}{n} \binom{n}{r, s, t}^2 \frac{1}{6^{2n}},$$

where S_i is the location of the walk at time i.

☞

Cas d=3

Le nombre d'étape pour revenir à l'origine est $2n$ pour un certain $n \in \mathbb{N}$. Dans le cas de $\mathcal{Z}^3$, de tels chemins doivent aller r fois vers le haut, r fois vers le bas, t fois à gauche, t fois à droite, s fois devant et s fois vers l'arrière et tel que $r + s + t = n$. Alors

$$\begin{aligned}
\mathbb{P}(S_{2n} = 0) &= \sum_{r+s+t=n} \frac{(2n)!}{(r!s!t!)^2} \frac{1}{6^{2n}} \\
&= \sum_{r,s,t} \frac{(2n)!}{(r!s!t!)^2} \frac{(n!)^2}{(n!)^2} \frac{1}{6^{2n}} \\
&= \binom{2n}{n} \sum_{r,s,t} \frac{(n!)^2}{(r!s!t!)^2} \frac{1}{6^{2n}}
\end{aligned}$$

4. Show that

$$\sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0) < \infty$$

and conclude for the case $d = 3$.

☞ Donc en remarquant que $\sum_{r,s,t} \binom{n}{rst} \frac{1}{3^n} = 1$. Pour montrer que $\sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0) < \infty$, on considère le cas le plus "défavorable" : $n = 3m$ (ie. $r = s = t = m$). On a clairement $\mathbb{P}(S_{2n} = 0) \leqslant \mathbb{P}(S_{6m} = 0)$. Alors,

$$\mathbb{P}(S_{6m} = 0) = \sum_m \binom{2n}{n} \binom{n}{m^3} \frac{1}{3^n} \frac{1}{2^{2n}}$$

On sait que $\binom{2n}{n} \frac{1}{2^{2n}} \sim \frac{1}{\sqrt{\pi n}}$ quand $n \to \infty$ et $\binom{n}{m^3} \sim \frac{3\sqrt{3}}{2n\pi}$ (conséquence immédiate de la formule de Stirling).

$$\begin{aligned}
\sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0) &\leqslant \sum_{n=0}^{\infty} \mathbb{P}(S_{6m} = 0) \\
&\sim \frac{3\sqrt{3}}{2\pi^{\frac{3}{2}}} \sum_{n=1}^{\infty} \left(\frac{1}{n}\right)^{\frac{3}{2}} \\
&< \infty.
\end{aligned}$$

Cette marche est donc transitoire.

Exercice 4. *Cat and mouse*

A cat and mouse each independently take a random walk on a connected, undirected, non-bipartite graph G. They start at the same time on different nodes, and each makes one transition at each time step. The cat eats the mouse if they are ever at the same node at some time step. Let n and m denote, respectively, the number of vertices and edges of G.

1. Show an upper bound of $\mathcal{O}(m^2 n)$ on the expected time before the cat eats the mouse. (Hint : Consider a Markov chain whose states are the ordered pairs (a, b), where a is the position of the cat and b is the position of the mouse.)

☞ Following the hint, we formulate a new Markov chain with n^2 states of the form $(i, j) \in [1, n]^2$. Each node (i, j) in the new chain is connected to $N(i)N(j)$ neighbors, where $N(i)$ denotes the number of neighbors of state i in the old Markov chain. Hence, the number of edges in the new chain comes to

$$2|E| = \sum_i \sum_j N((i,j)) = \sum_i \sum_j N(i)N(j) = \Big(\sum_i N(i)\Big)\Big(\sum_j N(j)\Big) = 4m^2$$

By Lemma 15 (If $(u, v) \in E$, then $h_{v,u} < 2|E|$), if edge exists between nodes $u = (i_1, j_1)$ and $v = (i_2, j_2)$, then $h_{u,v} \leq 2|E| = 4m^2$.
In order to obtain the $\mathcal{O}(m^2 n)$ upper bound, we need to show that for any node (i, j), there exists a path of length $\mathcal{O}(m^2 n)$ connecting it to some node of the form (v, v). In fact, we show that there exists a length $\mathcal{O}(n)$ path between (i, j) and (i, i). Since the graph is undirected, the cat can always go back to node i in two steps. At the same time, because the graph is connected, there exists a path of length $k < n$ from j to i. If k is even, then the mouse will run into the cat. If k is odd, then the mouse will get to node i when the cat is away. But since the chain is non-bipartite, there must be a path of odd length from i back to itself; let the mouse follow this path, and it will run into the cat on the next return to i. Thus the total length of this path from (i, j) to (i, i) is at most $3n$. Each edge on this path requires at most $4m^2$ steps, thus the desired upper bound on the time to collision is $\mathcal{O}(m^2 n)$ steps.