

# Marche aléatoire sur un graphe quelconque

BALZAMO Lucas, CARCONE Pierre, SIMONET-DAVIN Sylvain, DETRAZ Gabriel, DODELIN Hugo, DUBOIS Valentin, ISARD MERILLOU Cédrik, NEGRIER-GIRAudeau Thomas, SOWINSKY Julien, PHILLIPOT Benjamin  
(1<sup>ères</sup> et T<sup>ale</sup>)

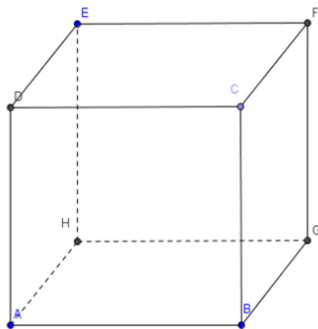
Enseignants : CARCONE Olivier, GRIHON Dominique, GRIHON Pierre, MAIMARAN Sébastien

Chercheur : DEVILLE Robert, Université BORDEAUX 1

## Sujet

Prenons un graphe quelconque. Cela peut être une suite de points, un nuage de points ou encore un carré, un losange... Dans chaque cas, certains points sont reliés entre eux : les points sont reliés à plus ou moins d'autres points du graphe.

[Un graphe est la donnée d'éléments, ici appelés «points» et d'une liste arbitraire de paires de points, ici appelés «segments» ou «liens»]



Au départ, un curseur se trouve sur un des sommets du graphe. Ensuite à chaque tour au hasard, il se déplace vers un des autres points du graphe auquel il est relié. Il ne peut pas rester sur place, sauf dans le cas où le point en relié à lui-même (déplacement avec saut possible).

On cherche à observer la position du curseur lorsque le nombre  $n$  de tours tend vers l'infini.

## Mots-clés

GRAPHE, MARCHE ALÉATOIRE, MATRICE, PROBABILITÉ, TRIANGLE DE PASCAL

Tout au long de notre démarche, nous utiliserons 3 variables :

$V$  : nombre de segments reliés à un point

$D$  : nombre de segments total dans le graphe

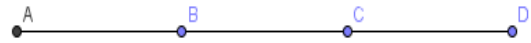
$n$  : numéro du tour où on se place

[par hypothèse, à chaque tour, les déplacements sont équiprobables, sauf en fin d'article où est abordé le cas asymétrique. Le point de départ sera toujours noté A]

## Étude à l'aide de relations de récurrence

Nous avons mis en équation les probabilités d'être en un point d'un graphe donné au tour  $n$ . La probabilité d'être en un point au tour  $n$  va dépendre des probabilités d'être aux points reliés à ce point au tour  $n-1$ . Nous obtenons donc des relations de récurrence.

• Pour la ligne :

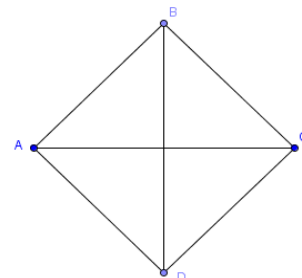


Lorsque l'on est sur le point A, on ne peut qu'aller sur B. Une fois sur B, on peut aller sur A ou C, d'où une probabilité d'une chance sur deux. En procédant ainsi pour chaque point, on obtient une suite de formules :

- $P(A_n) = \frac{1}{2} P(B_{n-1})$
- $P(B_n) = P(A_{n-1}) + \frac{1}{2} P(C_{n-1})$
- $P(C_n) = P(D_{n-1}) + \frac{1}{2} P(B_{n-1})$
- $P(D_n) = \frac{1}{2} P(C_{n-1})$

De même, nous avons sélectionné plusieurs graphes et avons trouvé les formules analogues pour chacun.

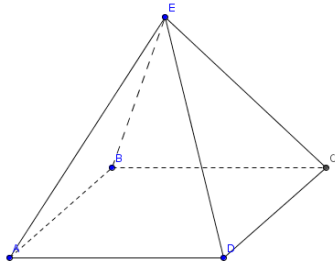
• Pour le carré doté de ses diagonales, chaque point est relié à trois autres points, ce qui nous donne à chaque déplacement des probabilités de  $\frac{1}{3}$  :



D'où les formules suivantes :

$$\begin{aligned} P(A_n) &= \frac{1}{3} (P(B_{n-1}) + P(C_{n-1}) + P(D_{n-1})) \\ P(B_n) &= \frac{1}{3} (P(A_{n-1}) + P(C_{n-1}) + P(D_{n-1})) \\ P(C_n) &= \frac{1}{3} (P(A_{n-1}) + P(B_{n-1}) + P(D_{n-1})) \\ P(D_n) &= \frac{1}{3} (P(A_{n-1}) + P(B_{n-1}) + P(C_{n-1})) \end{aligned}$$

- Nous avons ensuite pris l'exemple d'une figure en volume, en l'occurrence une **pyramide à base carrée** :



Chaque point est relié à trois autres points, excepté le sommet E, qui est relié à quatre points.

En raisonnant de manière analogue, on obtient cette suite de formules :

$$P(A_n) = 1/3(P(B_{n-1}) + P(D_{n-1})) + 1/4 P(E_{n-1})$$

$$P(B_n) = 1/3(P(A_{n-1}) + P(C_{n-1})) + 1/4 P(E_{n-1})$$

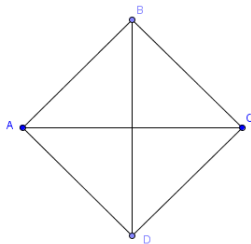
$$P(C_n) = 1/3(P(B_{n-1}) + P(D_{n-1})) + 1/4 P(E_{n-1})$$

$$P(D_n) = 1/3(P(A_{n-1}) + P(C_{n-1})) + 1/4 P(E_{n-1})$$

$$P(E_n) = 1/3(P(A_{n-1}) + P(B_{n-1}) + P(C_{n-1}) + P(D_{n-1}))$$

- Nous avons fait de même pour une suite de losanges, un cube, une droite, un losange à une diagonale...

À fin de pouvoir observer l'évolution des probabilités de présence du curseur sur un point, il est nécessaire de pouvoir simuler de façon aisée les sauts du curseur. Le tableur nous permet, à l'aide des formules de récurrence trouvées ci-dessus, d'observer ces évolutions. On observe une stabilisation des probabilités. Ainsi, pour une configuration où tous les points sont reliés, la probabilité d'arriver à un point pour un nombre  $n$  important de tours semble être toujours la même pour tous les points. Prenons l'exemple d'un carré à 2 diagonales :



A	B	C	D
1	0	0	0
0	0,33333333	0,33333333	0,33333333
0,33333333	0,22222222	0,22222222	0,22222222
0,22222222	0,25925926	0,25925926	0,25925926
0,25925926	0,24691358	0,24691358	0,24691358
.....	.....	.....	.....
0,25000002	0,24999999	0,24999999	0,24999999
0,24999999	0,25	0,25	0,25
0,25	0,25	0,25	0,25

- Nous avons fait de même pour les autres figures (droite, cube, losange à 1 diagonale, losange, suite de losanges, pyramide).

Dans certains cas, un point ne peut être atteint à chaque tour. Dans le cas de la droite par exemple, le point B ne peut pas être atteint aux tours pairs lorsque le départ est en A].



A	B	C	D
1	0	0	0
0	1	0	0
0,5	0	0,5	0
0	0,75	0	0,25
.....	.....	.....	.....
0	0,66667175	0	0,33332825
0,33333588	0	0,66666412	0

⓪ Cependant, bien que les probabilités se stabilisent toujours, elles ne respectent pas forcément une éventuelle alternance pair/impair.

La stabilisation dépend de « l'accessibilité du point », autrement-dit du nombre de liaisons qu'il possède avec d'autres points, mais aussi de « l'accessibilité » et du nombre de liens que possèdent ces autres points.

Suite aux observations faites sur ces cas de figure, on a remarqué la présence d'une convergence des probabilités vers des valeurs spécifiques à chaque point et nous avons essayé d'établir une conjecture de formule générale de la limite des probabilités qui s'avère être la suivante :

**Conjecture**

$$\lim_{n \rightarrow \infty} P(A_n) = \frac{V_A}{2D}$$

où  $P(A_n)$  désigne la probabilité d'être sur le point au tour  $n$ ,  $V_A$  le nombre de points auquel A est relié et D le nombre total de segments du graphe.

⓪ Cependant, dans la plupart des cas, un point ne peut être atteint avec équiprobabilité à chaque tour. On obtient alors des limites différentes en fonction du tour.

[En fait, ce qui est conjecturé est l'existence, pour chaque point, de centaines de valeurs, en nombre fini, telles que la probabilité de présence au tour  $n$  soit, pour  $n$  grand, proche de l'une de ces valeurs. De plus, ces valeurs de stabilisation apparaîtraient périodiquement, lorsque le nombre de tour augmente. La conjecture énoncée ci-dessus ne concerne que le cas où il n'y a qu'une valeur de stabilisation non nulle.]

## Formule dans le cas « tous points reliés »

On s'intéresse à un graphe où tous les points sont reliés. On note  $D$  le nombre de segments du graphe (qui correspondent à des liens symétriques entre les points) et  $N$  le nombre de points. On a donc  $V_A = N-1$  et

$$D = \frac{N(N-1)}{2}$$

En regardant pour le triangle, on a conjecturé puis démontré par récurrence la formule suivante [Théorème]

$$P(A_n) = \frac{1}{3} + \frac{2}{3} \left( \frac{-1}{2} \right)^{n-1}$$

si on part de A (soit  $P(A_1) = 1$ )

• Par analogie, nous avons conjecturé dans le cas général puis démontré par récurrence la formule suivante :

[Théorème]

$$P(A_n) = \frac{1}{N} + \frac{N-1}{N} \left( -\frac{1}{N-1} \right)^{n-1}$$

[Preuve].

La formule est vraie si  $n = 1$

On a  $P(A_{n+1}) = \frac{1}{N-1} (1 - P(A_n))$  puisque le point

A est relié aux  $N-1$  autres points avec équiprobabilité d'y aller.

Donc

$$\begin{aligned} P(A_{n+1}) &= \frac{1}{N-1} \left( 1 - \frac{1}{N} - \frac{N-1}{N} \left( -\frac{1}{N-1} \right)^{n-1} \right) \\ &= \frac{1}{N} - \frac{N-1}{N} \frac{1}{N-1} \left( -\frac{1}{N-1} \right)^{n-1} \\ &= \frac{1}{N} + \frac{N-1}{N} \left( -\frac{1}{N-1} \right)^n \end{aligned}$$

La formule est donc vraie pour tout  $n$  non nul.

## Limite

On voit que si  $n \rightarrow \infty$  on a  $P(A_n) \rightarrow \frac{1}{N}$

On a conjecturé plus haut que

$$\lim_{n \rightarrow \infty} P(A_n) = \frac{V_A}{2D}$$

Cette formule se trouve être vérifiée dans le cas « tous points reliés ». En effet, dans ce cas,

$V_A = N-1, D = \frac{N(N-1)}{2}$ , donc la limite de  $P(A_n)$  est bien  $\frac{V_A}{2D}$ .

## Étude à l'aide de matrices

Reprenons le triangle avec liens symétriques

Les trois relations de récurrence peuvent se traduire sous la forme matricielle suivante :

$$\begin{pmatrix} P(A_{n+1}) \\ P(B_{n+1}) \\ P(C_{n+1}) \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} * \begin{pmatrix} P(A_n) \\ P(B_n) \\ P(C_n) \end{pmatrix}$$

Soit sous forme simplifiée  $X_{n+1} = MX_n$ .

Nous avons reconnu une suite « géométrique » comme en cours et par analogie nous pensons que  $X_n = M^n X_0$ .

Mais nous ne savons pas calculer le produit de matrices, alors nous avons essayé de voir le cas du carré de  $M$  pour une matrice à 2 lignes et deux colonnes.

Soit  $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  alors

Par convention :

$$P(A_{n+1}) = a * P(A_n) + b * P(B_n) \quad \text{et} \quad P(B_{n+1}) = c * P(A_n) + d * P(B_n)$$

$$\Leftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} * \begin{pmatrix} P(A_n) \\ P(B_n) \end{pmatrix} = \begin{pmatrix} P(A_{n+1}) \\ P(B_{n+1}) \end{pmatrix}$$

En revenant à l'écriture de récurrence, on a :

$$P(A_{n+2}) = a * P(A_{n+1}) + b * P(B_{n+1})$$

$$P(B_{n+2}) = c * P(A_{n+1}) + d * P(B_{n+1})$$

$$\Leftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} * \begin{pmatrix} P(A_{n+1}) \\ P(B_{n+1}) \end{pmatrix} = \begin{pmatrix} P(A_{n+2}) \\ P(B_{n+2}) \end{pmatrix}$$

Par substitution,

$$P(A_{n+2}) = a * (a * P(A_n) + b * P(B_n)) + b * (c * P(A_n) + d * P(B_n))$$

$$\Leftrightarrow P(A_{n+2}) = a^2 * P(A_n) + ab * P(B_n) + bc * P(A_n) + bd * P(B_n)$$

$$\Leftrightarrow P(A_{n+2}) = a^2 * P(A_n) + ab * P(B_n) + bc * P(A_n) + bd * P(B_n)$$

$$\Leftrightarrow P(A_{n+2}) = P(A_n)(a^2 + bc) + P(B_n)(ab + bd)$$

$$P(B_{n+2}) = c * (a * P(A_n) + b * P(B_n)) + d * (c * P(A_n) + d * P(B_n))$$

$$\Leftrightarrow P(B_{n+2}) = P(A_n)(ac + dc) + P(B_n)(cb + d^2)$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} * \begin{pmatrix} a & b \\ c & d \end{pmatrix} * \begin{pmatrix} P(A_n) \\ P(B_n) \end{pmatrix} = \begin{pmatrix} P(A_{n+2}) \\ P(B_{n+2}) \end{pmatrix}$$

$$\text{Par substitution, } \begin{pmatrix} a & b \\ c & d \end{pmatrix} * \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & ab + bd \\ ac + dc & cb + d^2 \end{pmatrix}$$

[Autrement dit, le produit de deux matrices  $M$  et  $M'$  peut simplement se calculer en multipliant  $M$  par chaque colonne de  $M'$  puis en mettant ensemble les colonnes obtenues]

Pour connaître la probabilité d'être sur un point au tour  $n$ , il nous faut donc calculer  $M^n$ , ce qui est possible avec une calculatrice.

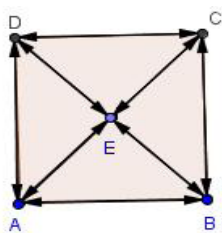
Dans le cas du triangle on aurait par exemple :

$$M^{10} \approx \begin{pmatrix} 0,334 & 0,333 & 0,333 \\ 0,333 & 0,334 & 0,333 \\ 0,333 & 0,333 & 0,334 \end{pmatrix}$$

Il semble bien y avoir une convergence vers une matrice formée de  $1/3$ .

Ce qui a été démontré par récurrence plus haut.

### • Cas de la pyramide



$$\begin{pmatrix} P_{An+1} \\ P_{Bn+1} \\ P_{Cn+1} \\ P_{Dn+1} \\ P_{En+1} \end{pmatrix} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/4 \\ 1/3 & 0 & 1/3 & 0 & 1/4 \\ 0 & 1/3 & 0 & 1/3 & 1/4 \\ 1/3 & 0 & 1/3 & 0 & 1/4 \\ 1/3 & 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} * \begin{pmatrix} P_{An} \\ P_{Bn} \\ P_{Cn} \\ P_{Dn} \\ P_{En} \end{pmatrix}$$

On calcule la puissance 10 de cette matrice :

$$\begin{pmatrix} 0,19 & 0,183 & 0,19 & 0,183 & 0,187 \\ 0,183 & 0,19 & 0,183 & 0,19 & 0,187 \\ 0,19 & 0,183 & 0,19 & 0,183 & 0,187 \\ 0,183 & 0,19 & 0,183 & 0,191 & 0,187 \\ 0,25 & 0,25 & 0,25 & 0,25 & 0,25 \end{pmatrix}$$

La dernière ligne correspond au point E dont la probabilité tend manifestement vers  $1/4$  ; c'est lui qui a le plus de liens « entrant ». [...]

### Étude de la marche sur un triangle à l'aide du Triangle de Pascal

On considère ici un triangle ABC direct avec une probabilité  $p$  d'effectuer un déplacement dans le sens direct et une probabilité  $q=1-p$  d'en effectuer un dans le sens indirect.

$$\text{Soit : } P(A_{n+1}) = p.P(C_n) + q.P(B_n)$$

$$P(B_{n+1}) = p.P(A_n) + q.P(C_n)$$

$$P(C_{n+1}) = p.P(A_n) + q.P(B_n)$$

Si on s'intéresse au cas des 4 premiers sauts, en partant de A, on obtient :

$$P(A_0)=1 \quad P(B_0)=0 \quad P(C_0)=0$$

$$P(A_1)=0 \quad P(B_1)=p \quad P(C_1)=q$$

$$P(A_2)=2pq \quad P(B_2)=q^2 \quad P(C_2)=p^2$$

$$P(A_3)=q^3+p^3 \quad P(B_3)=3p^2q \quad P(C_3)=3pq^2$$

$$P(A_4)=6p^2q^2 \quad P(B_4)=4pq^3+p^4$$

$$P(C_4)=4p^3q+q^4$$

On observe l'apparition des termes du triangle de Pascal dans le développement de  $(p+q)^n$ .

Sauts(n)\termes	$p^n$	$p^{n-1}q$	$p^{n-2}q^2$	$p^{n-3}q^3$	$p^{n-4}q^4$	$p^{n-5}q^5$	...	$q^n$
0	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		

Tableau représentatif du triangle de Pascal

En prenant la colonne  $p^n$  comme la colonne 0, nous allons exprimer les probabilités  $P(A)$ ,  $P(B)$  et  $P(C)$  en fonction de leur colonne :

$$P(A_0) \Rightarrow \text{colonne 0}$$

$$P(B_1) \Rightarrow 0 \quad P(C_1) \Rightarrow 1$$

$$P(A_2) \Rightarrow 1 \quad P(B_2) \Rightarrow 2 \quad P(C_2) \Rightarrow 0$$

$$P(A_3) \Rightarrow 0+3 \quad P(B_3) \Rightarrow 1 \quad P(C_3) \Rightarrow 2$$

$$P(A_4) \Rightarrow 2 \quad P(B_4) \Rightarrow 0+3 \quad P(C_4) \Rightarrow 1+4$$

$$\text{Par exemple } P(C_4) = q^4 + 4pq^3$$

Lorsque l'on va du sommet A au sommet B, du sommet B au sommet C ou du sommet C au sommet A, cela revient à multiplier la probabilité d'être sur le sommet précédent par  $p$ .

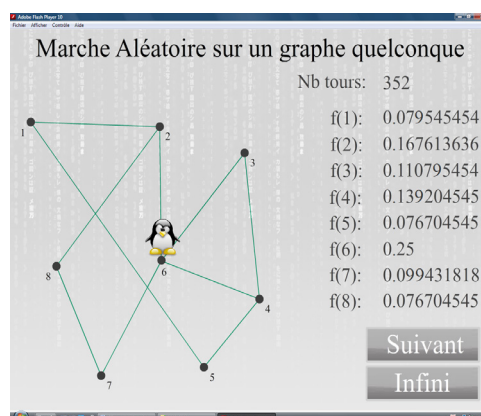
Sur le tableau, cela revient à rester sur la même colonne.

Lorsque l'on va du sommet A au sommet C, du sommet C au sommet B ou du sommet B au sommet A, cela revient à multiplier la probabilité d'être sur le sommet précédent par  $q$ .

Sur le tableau, cela revient à se placer sur la colonne qui est à droite.

### Animation

Pour finir, nous avons fait une animation sur laquelle on voit un pingouin se déplacer sur un graphe. On calcule simultanément les fréquences de passage sur les différents points depuis le départ. Cela permet de voir apparaître les convergences évoquées plus haut.



\*\*\*