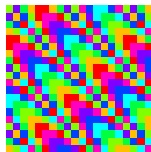


LA MATRICE CACHÉE DE GOOGLE

Michel Rigo (Université de Liège)

<http://www.discmath.ulg.ac.be/>



Nombre de pages sur internet $\sim 10.000.000.000 = 10^{10}$



EXEMPLE

Résultats **1 - 10** sur un total
d'environ **9 560 000** pour **matrice**

Résultats **1 - 10** sur un total
d'environ **279 000 000** pour **football**

QUESTION

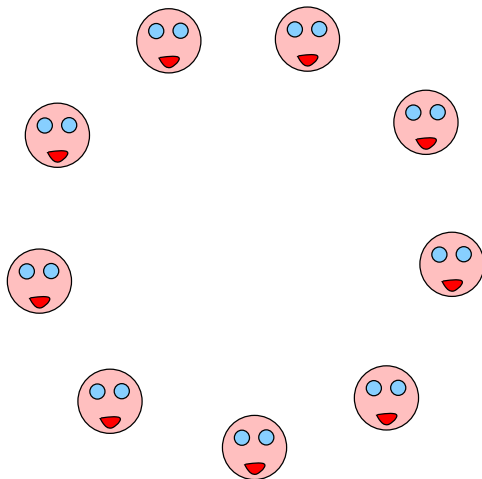
Comment **classer** ces pages pour mettre en évidence les plus intéressantes ?



Larry Page, Sergei Brin

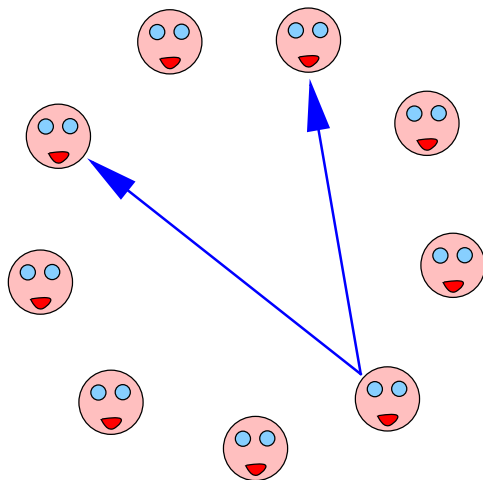
COMMENT CLASSER ?

Qui est/sont le/les meilleur(s) étudiant(s) en **sport** ?



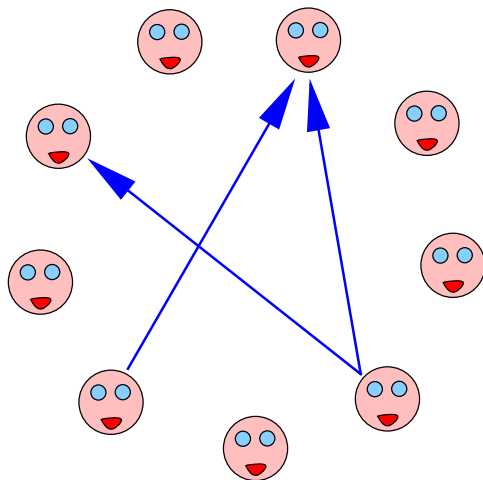
COMMENT CLASSER ?

Qui est/sont le/les meilleur(s) étudiant(s) en sport ?



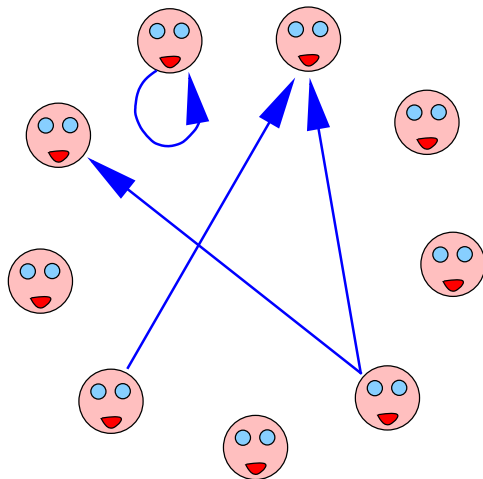
COMMENT CLASSER ?

Qui est/sont le/les meilleur(s) étudiant(s) en sport ?



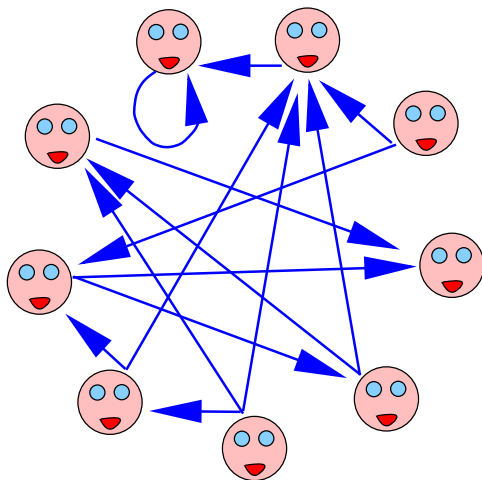
COMMENT CLASSER ?

Qui est/sont le/les meilleur(s) étudiant(s) en sport ?



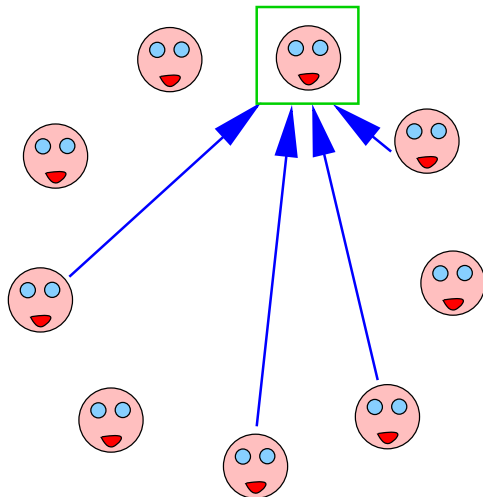
COMMENT CLASSER ?

Qui est/sont le/les meilleur(s) étudiant(s) en sport ?



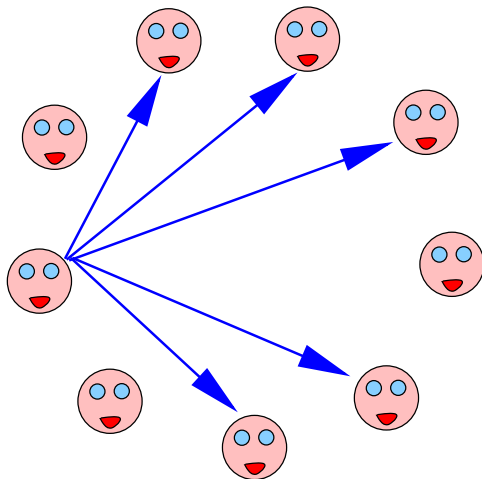
PREMIÈRE OBSERVATION

- *plus le soutien est grand, plus important sera le score*



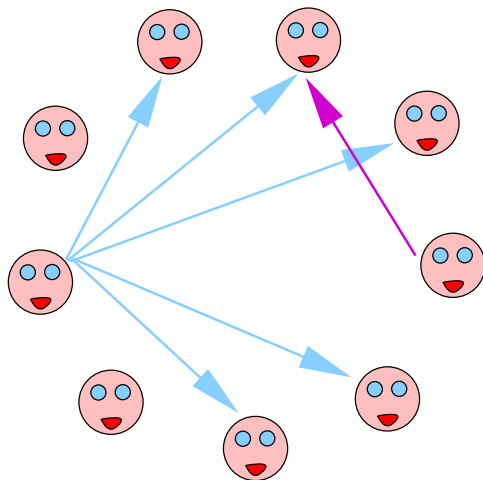
SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes !*



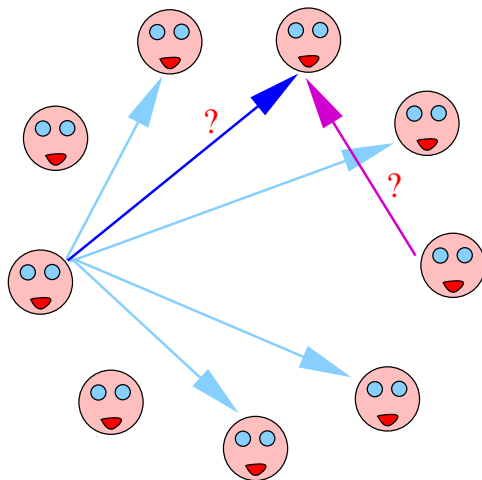
SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes !*



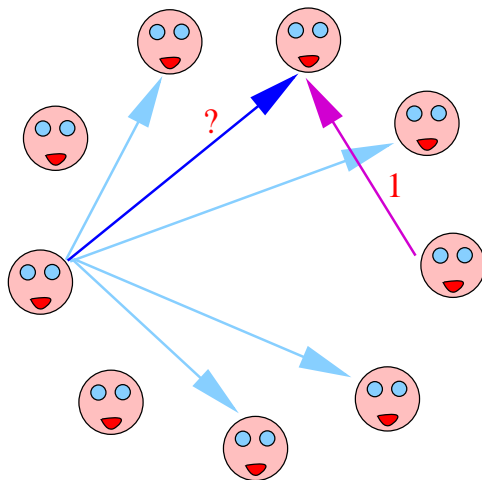
SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes !*



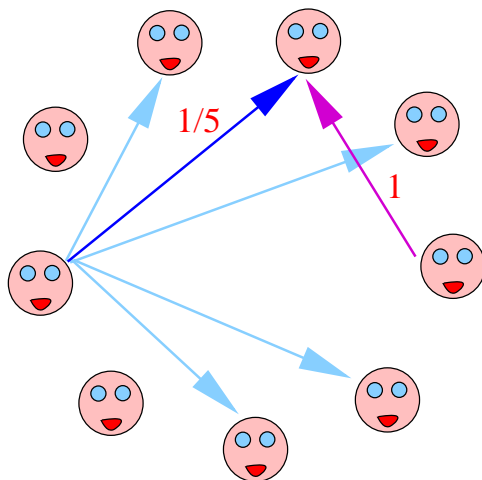
SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes !*



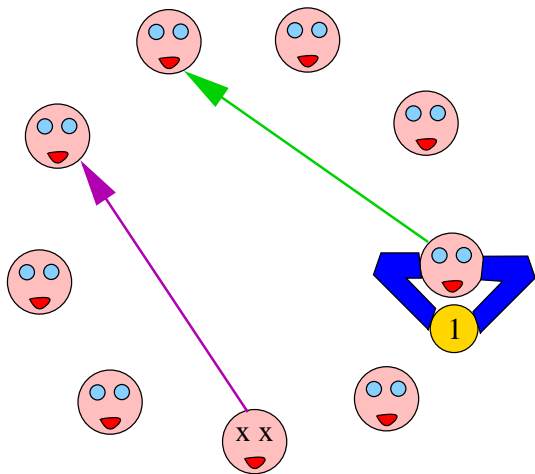
SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes !*

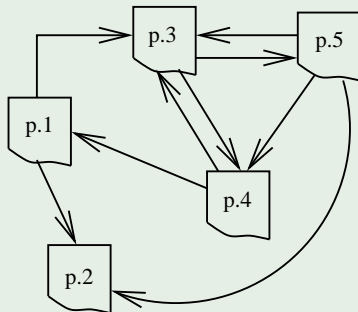


SECONDE OBSERVATION

- *ne pas accorder la même importance à tous les votes (2) !*



EXEMPLE



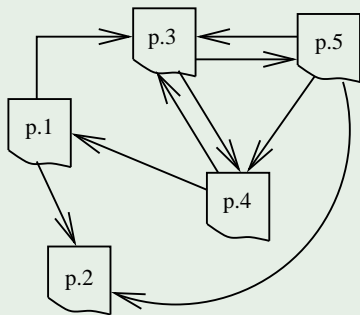
Attribuer un **score** (PageRank) à chacune des pages

Modèle de S. Brin et L. Page :

- R1.** On accorde plus d'importance, i.e., un score de "PageRank" plus élevé, aux pages référencées par des pages qui font elles-mêmes autorité dans le domaine, c'est-à-dire qui ont un PageRank élevé;
- R2.** On accorde d'autant moins de crédit à une référence, si elle provient d'une page qui dispose de nombreux liens.



SUITE DE L'EXEMPLE



$$\left\{ \begin{array}{l} s_1 = s_4/2 \\ s_2 = s_1/2 + s_5/3 \\ s_3 = s_1/2 + s_4/2 + s_5/3 \\ s_4 = s_3/2 + s_5/3 \\ s_5 = s_3/2 \end{array} \right.$$

Inconnues s_1, s_2, s_3, s_4, s_5

On est ramené à résoudre un **système d'équations linéaires** ...
Mais ici... des millions d'équations et d'inconnues !

DEUX QUESTIONS PRIMORDIALES

- ▶ **Existence d'une solution (non nulle) ?**
- ▶ **Unicité de la solution ?**

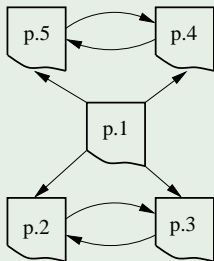
CONTINUONS L'EXEMPLE...

$$\left\{ \begin{array}{l} s_1 = s_4/2 \\ s_2 = s_1/2 + s_5/3 \\ s_3 = s_1/2 + s_4/2 + s_5/3 \\ s_4 = s_3/2 + s_5/3 \\ s_5 = s_3/2 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} s_1 = s_4/2 \\ s_5 = s_3/2 \\ s_2 = s_4/4 + s_3/6 \\ s_3 = s_4/4 + s_4/2 + s_3/6 \\ s_4 = s_3/2 + s_3/6 \end{array} \right.$$

$$\left\{ \begin{array}{l} s_4 = 10s_3/9 \\ s_4 = 6s_3/9 \end{array} \right.$$

La seule solution est $s_1 = s_2 = s_3 = s_4 = s_5 = 0$.

UN SECOND EXEMPLE. . .



$$\left\{ \begin{array}{l} s_2 = s_1/4 + s_3 \\ s_3 = s_1/4 + s_2 \\ s_4 = s_1/4 + s_5 \\ s_5 = s_1/4 + s_4 \end{array} \right.$$

on trouve par exemple,

- ▶ $s_1 = s_2 = s_3 = 0$, $s_4 = s_5 = 1/2$ ou bien,
- ▶ $s_1 = s_4 = s_5 = 0$, $s_2 = s_3 = 1/2$.

CONCLUSION

On ne peut *a priori*

- ▶ ni garantir l'existence d'une solution $\neq 0$:-)
- ▶ ni garantir l'unicité de la solution :-)

SOLUTION

→ Perturber légèrement le modèle initial
pour obtenir un système “proche” mais avec de “belles”
propriétés.

CONCLUSION

On ne peut *a priori*

- ▶ ni garantir l'existence d'une solution $\neq 0$:-)
- ▶ ni garantir l'unicité de la solution :-)

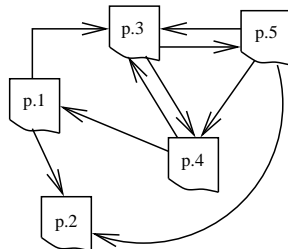
SOLUTION

→ **Perturber légèrement le modèle initial**
pour obtenir un système “proche” mais avec de “belles”
propriétés.

LES MATRICES SONT DE LA PARTIE

Reprenons l'exemple initial...

$$\begin{cases} s_1 = s_4/2 \\ s_2 = s_1/2 + s_5/3 \\ s_3 = s_1/2 + s_4/2 + s_5/3 \\ s_4 = s_3/2 + s_5/3 \\ s_5 = s_3/2 \end{cases}$$



$$\underbrace{\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix}}_{\mathbf{s}} = \underbrace{\begin{pmatrix} 0 & 0 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 0 & 1/3 \\ 1/2 & 0 & 0 & 1/2 & 1/3 \\ 0 & 0 & 1/2 & 0 & 1/3 \\ 0 & 0 & 1/2 & 0 & 0 \end{pmatrix}}_{\mathbf{B}} \underbrace{\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix}}_{\mathbf{s}} .$$

REMARQUE

Si $\mathbf{s} = \mathbf{B}\mathbf{s}$, alors $2\mathbf{s} = \mathbf{B}2\mathbf{s}$, $3\mathbf{s} = \mathbf{B}3\mathbf{s}$, \dots , $t\mathbf{s} = \mathbf{B}t\mathbf{s}$.

Tout multiple d'une solution est encore solution.

CONSÉQUENCE

On peut donc imposer $s_1 + s_2 + s_3 + s_4 + s_5 = 1$.

EXEMPLE

Si $s_1 = 1$, $s_2 = 2$, $s_3 = 3$, $s_4 = s_5 = 0$, alors

$$s_1 + s_2 + s_3 + s_4 + s_5 = 6$$

et

$$(s'_1, s'_2, s'_3, s'_4, s'_5) = \left(\frac{s_1}{6}, \frac{s_2}{6}, \frac{s_3}{6}, \frac{s_4}{6}, \frac{s_5}{6}\right)$$

est encore solution

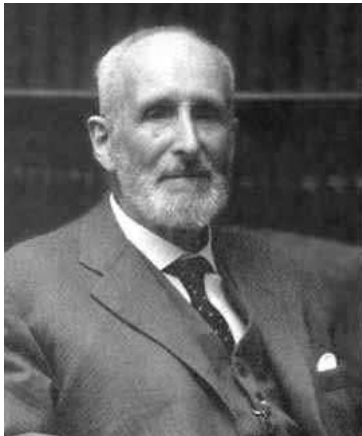
PERTURBATION

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 0 & 1/3 \\ 1/2 & 0 & 0 & 1/2 & 1/3 \\ 0 & 0 & 1/2 & 0 & 1/3 \\ 0 & 0 & 1/2 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1/5 & 0 & 1/2 & 0 \\ 1/2 & 1/5 & 0 & 0 & 1/3 \\ 1/2 & 1/5 & 0 & 1/2 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 0 \end{pmatrix}.$$

$$0,85 \cdot \begin{pmatrix} 0 & 1/5 & 0 & 1/2 & 0 \\ 1/2 & 1/5 & 0 & 0 & 1/3 \\ 1/2 & 1/5 & 0 & 1/2 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 0 \end{pmatrix} + 0,15 \cdot \begin{pmatrix} 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix}$$

$$= \begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} = \mathbf{G}.$$

La matrice \mathbf{G} est *primitive*,
on peut appliquer
la **théorie de Perron**



Oskar Perron (1880–1975)

THÉORÈME DE PERRON 1907

Soit $A \geq 0$ une matrice carrée primitive de dimension n .

- ▶ La matrice A possède un **vecteur propre** $v_A \in \mathbb{R}^n$ (resp. $w_A \in \mathbb{R}^n$) dont les composantes sont toutes strictement positives et correspondant à une valeur propre $\lambda_A > 0$,

$$A v_A = \lambda_A v_A \text{ (resp. } \widetilde{w}_A A = \lambda_A \widetilde{w}_A).$$

- ▶ Cette valeur propre λ_A possède une **multiplicité** algébrique (et géométrique) **simple**.
- ▶ Tout vecteur propre de A dont les composantes sont strictement positives est un multiple de v_A .
- ▶ Toute autre valeur propre $\mu \in \mathbb{C}$ de A est telle que $|\mu| < \lambda_A$.
- ▶ $A^k = \lambda_A^k v_A \widetilde{w}_A + o(\lambda_A^k)$

Introduction to Dynamical Systems



Michael Brin
and Garrett Stuck

CAMBRIDGE

INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$

p.3

p.5

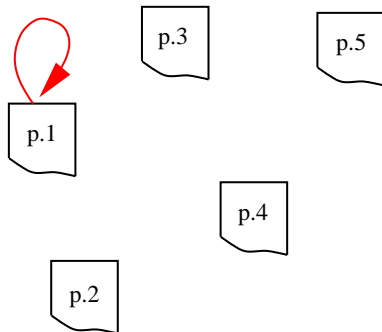
p.1

p.4

p.2

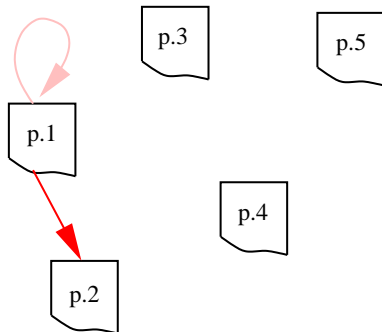
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



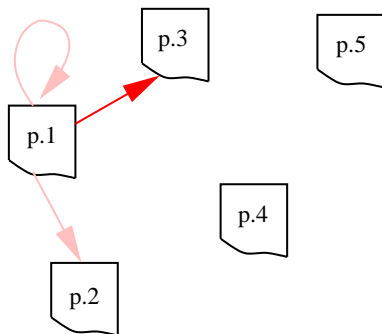
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



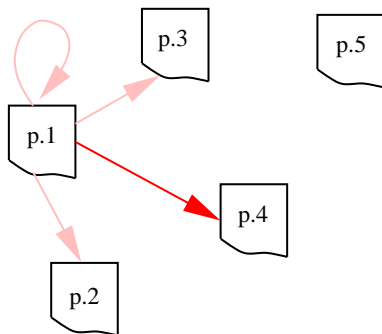
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



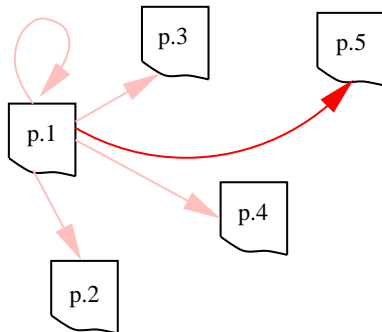
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



INTERPRÉTATION PROBABILISTE

$$0,85 \cdot \begin{pmatrix} 0 & 1/5 & 0 & 1/2 & 0 \\ 1/2 & 1/5 & 0 & 0 & 1/3 \\ 1/2 & 1/5 & 0 & 1/2 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 1/3 \\ 0 & 1/5 & 1/2 & 0 & 0 \end{pmatrix} + 0,15 \cdot \begin{pmatrix} 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix}$$

- ▶ On lance une pièce (**Pile** 85%, **Face** 15%)
 - ▶ **Pile** : on choisit au hasard un lien depuis la page courante
 - ▶ **Face** : on se téléporte aléatoirement

INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$

p.3

p.5

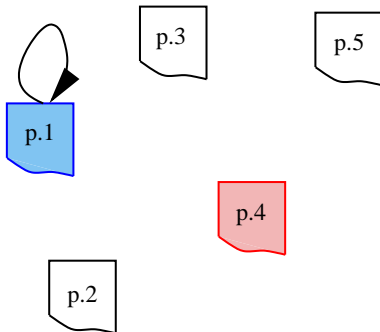
p.1

p.4

p.2

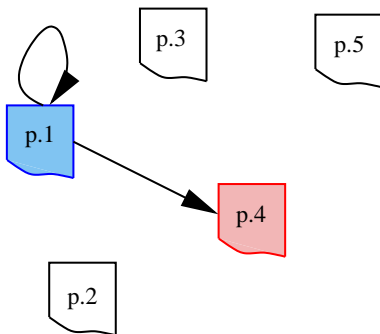
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



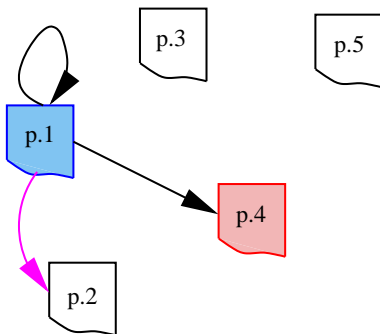
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



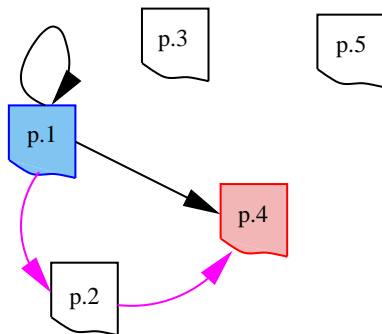
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



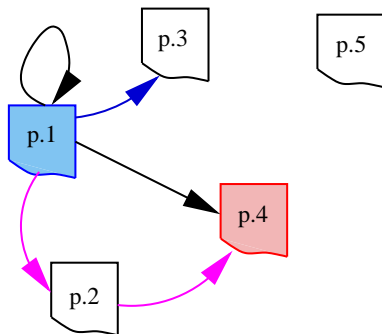
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



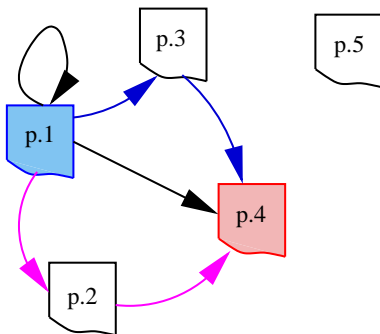
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



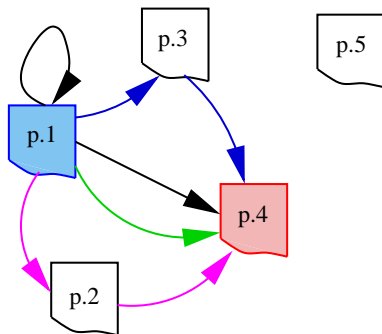
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



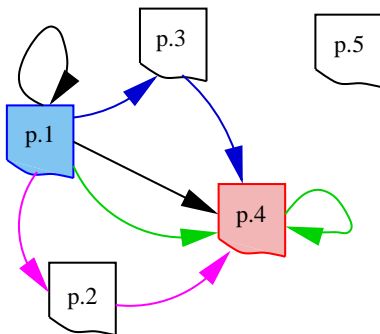
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



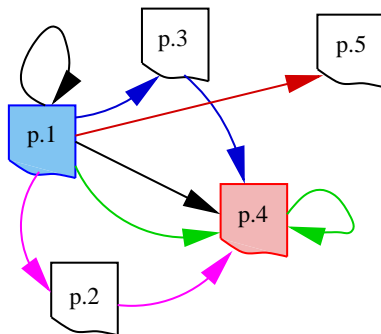
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & \mathbf{3/100} & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



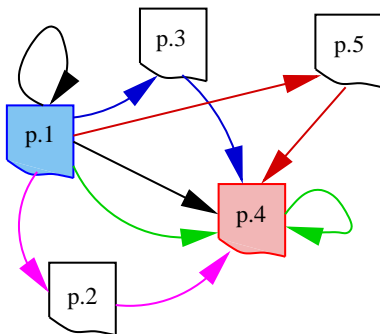
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & \mathbf{3/100} & 47/150 \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



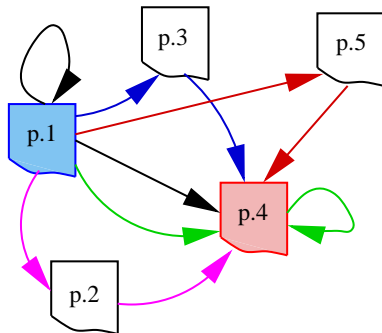
INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & \mathbf{3/100} & \mathbf{47/150} \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



INTERPRÉTATION PROBABILISTE

$$\begin{pmatrix} 3/100 & 1/5 & 3/100 & 91/200 & 3/100 \\ 91/200 & 1/5 & 3/100 & 3/100 & 47/150 \\ 91/200 & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & \mathbf{1/5} & \mathbf{91/200} & \mathbf{3/100} & \mathbf{47/150} \\ 3/100 & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix} \begin{pmatrix} \mathbf{3/100} & 1/5 & 3/100 & 91/200 & 3/100 \\ \mathbf{91/200} & 1/5 & 3/100 & 3/100 & 47/150 \\ \mathbf{91/200} & 1/5 & 3/100 & 91/200 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 47/150 \\ \mathbf{3/100} & 1/5 & 91/200 & 3/100 & 3/100 \end{pmatrix}$$



$$\frac{3}{100} \frac{3}{100} + \frac{1}{5} \frac{91}{200} + \frac{91}{200} \frac{91}{200} + \frac{3}{100} \frac{3}{100} + \frac{47}{150} \frac{3}{100}$$

LA QUESTION...

Que représente \mathbf{G}^2 ?

$$\mathbf{G}^2 = \mathbf{G} \cdot \mathbf{G} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$
$$= \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

LA QUESTION...

Que représente \mathbf{G}^2 ?

$$\mathbf{G}^2 = \mathbf{G} \cdot \mathbf{G} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$
$$= \begin{pmatrix} * & * & * & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

LA QUESTION...

Que représente \mathbf{G}^2 ?

$$\mathbf{G}^2 = \mathbf{G} \cdot \mathbf{G} = \begin{pmatrix} * & * & * & * & * \\ \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \begin{pmatrix} * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \end{pmatrix}$$
$$= \begin{pmatrix} * & * & * & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

LA QUESTION (2)...

Que représente \mathbf{G}^3 ?

$$\mathbf{G}^3 = \mathbf{G}^2 \cdot \mathbf{G} = \begin{array}{c} \overbrace{\begin{pmatrix} * & * & * & * & * \\ \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}}^{\mathbf{G}^2} \cdot \overbrace{\begin{pmatrix} * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & \mathbf{X} & * & * \end{pmatrix}}^{\mathbf{G}} \\ \\ = \begin{pmatrix} * & * & * & * & * \\ * & * & \mathbf{X} & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \end{array}$$

$$\mathbf{G}^5 = \begin{pmatrix} 0.14721 & 0.13301 & 0.12861 & 0.13367 & 0.13989 \\ 0.18196 & 0.18931 & 0.19412 & 0.18584 & 0.18545 \\ 0.26597 & 0.26067 & 0.26295 & 0.25602 & 0.26253 \\ 0.16641 & 0.17481 & 0.17349 & 0.17906 & 0.17107 \\ 0.23844 & 0.2422 & 0.24084 & 0.2454 & 0.24107 \end{pmatrix}$$

$$\mathbf{G}^{10} = \begin{pmatrix} 0.13568 & 0.13553 & 0.13545 & 0.13559 & 0.13561 \\ 0.18801 & 0.18804 & 0.1881 & 0.18799 & 0.18802 \\ 0.26173 & 0.26161 & 0.26159 & 0.26159 & 0.26166 \\ 0.17304 & 0.17319 & 0.17322 & 0.17319 & 0.17311 \\ 0.24155 & 0.24163 & 0.24164 & 0.24165 & 0.24159 \end{pmatrix}$$

$$\mathbf{G}^{20} = \begin{pmatrix} 0.13556 & 0.13556 & 0.13556 & 0.13556 & 0.13556 \\ 0.18804 & 0.18804 & 0.18804 & 0.18804 & 0.18804 \\ 0.26163 & 0.26163 & 0.26163 & 0.26163 & 0.26163 \\ 0.17316 & 0.17316 & 0.17316 & 0.17316 & 0.17316 \\ 0.24162 & 0.24162 & 0.24162 & 0.24162 & 0.24162 \end{pmatrix}$$

$$\mathbf{G}^{200} = \begin{pmatrix} 0.13556 & 0.13556 & 0.13556 & 0.13556 & 0.13556 \\ 0.18804 & 0.18804 & 0.18804 & 0.18804 & 0.18804 \\ 0.26163 & 0.26163 & 0.26163 & 0.26163 & 0.26163 \\ 0.17316 & 0.17316 & 0.17316 & 0.17316 & 0.17316 \\ 0.24162 & 0.24162 & 0.24162 & 0.24162 & 0.24162 \end{pmatrix}$$

CONSÉQUENCE DU THM. DE PERRON

Tous les éléments d'une ligne tendent vers une même valeur !

On a trouvé les scores !

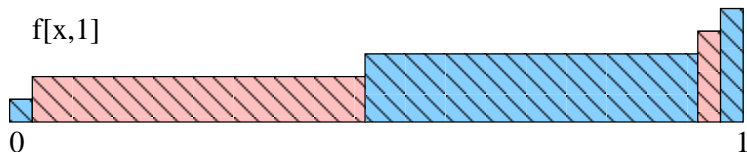
INTERPRÉTATION

Après un *long* "surf", ces valeurs correspondent à la proba. de se trouver sur une page donnée.

Plus une page est importante, plus la probabilité de s'y trouver est grande.

PROGRAMMATION

```
f[x_,i_]:=
1+If[x<g[[1,i]],0,1]
+If[x<g[[1,i]]+g[[2,i]],0,1]
+If[x<g[[1,i]]+g[[2,i]]+g[[3,i]],0,1]
+If[x<g[[1,i]]+g[[2,i]]+g[[3,i]]+g[[4,i]],0,1];
```



PREMIÈRE COLONNE DE LA MATRICE G

$3/100, 91/200, 91/200, 3/100, 3/100$

PROGRAMMATION

```
(* 1000 choix aléatoires depuis la page 1 *)
```

```
Table[f[Random[],1],{i,1,1000}]
```

```
{3,2,2,1,3,3,2,2,2,2,3,3,3,4,2,1,2,2,2,2,3,3,  
2,2,3,3,2,2,2,3,2,3,3,3,2,1,2,3,3,5,4,3,3,3,2,3,  
2,3,3,2,2,3,3,2,2,3,3,3,3,3,3,2,5,3,3,3,2,3,2,4,  
2,2,3,3,4,3,3,3,3,4,2,5,1,2,2,3,2,3,3,2,3,...}
```

```
Table[Length[Position[%,i]],{i,1,5}]
```

```
{40,430,465,31,34}
```

PREMIÈRE COLONNE DE LA MATRICE G

3/100, 91/200, 91/200, 3/100, 3/100

PROGRAMMATION

(* un "surf" aléatoire depuis la page 1 *)

```
NestList[f[Random[],#]&,1,100]
```

```
{1,1,2,4,1,2,5,5,5,5,2,1,2,5,2,3,4,1,3,2,3,4,  
3,4,3,4,1,3,4,1,2,4,3,4,1,2,2,3,4,5,5,2,4,3,4,3,  
5,5,3,4,3,4,3,4,1,3,4,3,5,2,1,2,4,1,2,1,1,2,5,2,  
1,2,3,3,4,1,3,5,2,4,1,2,2,4,1,3,5,4,3,5,4,3,...}
```

```
Table[Length[Position[%,i]],{i,1,5}]
```

```
{17,23,22,22,17}
```

```
NestList[f[Random[],#]&,1,100]
```

```
{1,2,2,1,2,2,5,3,4,1,2,4,1,3,5,2,2,2,5,5,5,5,  
3,4,1,3,4,5,5,2,3,5,5,5,2,5,3,4,3,5,2,1,3,5,...}
```

```
Table[Length[Position[%,i]],{i,1,5}]
```

```
{15,23,24,13,26}
```

LA PREMIÈRE COLONNE DE G^{200}

0.13556 0.18804 0.26163 0.17316 0.24162

```
NestList[f[Random[],#]&,1,10000];  
Table[Length[Position[%,i]],{i,1,5}]  
{1333,1912,2620,1724,2412}  
NestList[f[Random[],#]&,2,10000];  
Table[Length[Position[%,i]],{i,1,5}]  
{1318,1886,2619,1688,2490}
```

LA PREMIÈRE COLONNE DE G^{200}

0.13556 0.18804 0.26163 0.17316 0.24162

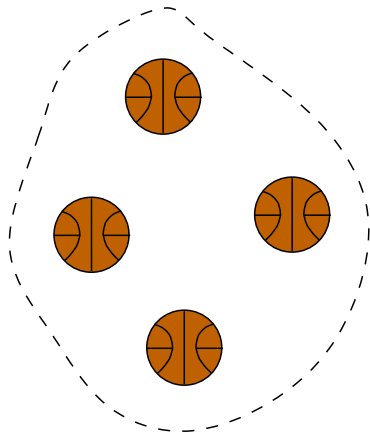
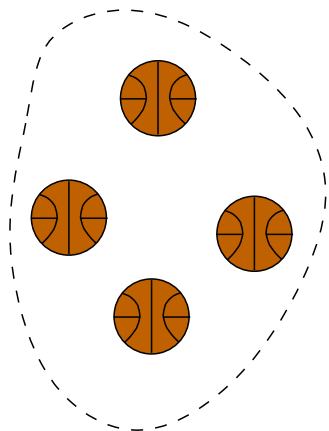
Donc, on calcule \mathbf{G}^{200} avec une matrice $10^{10} \times 10^{10}$



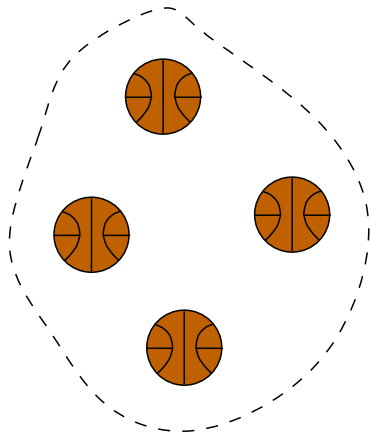
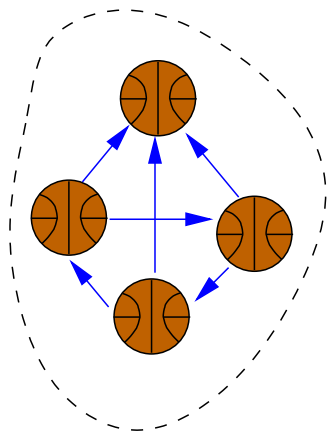
- R1.** On accorde plus d'importance, i.e., un score de "PageRank" plus élevé, aux pages référencées par des pages qui font elles-mêmes autorité dans le domaine, c'est-à-dire qui ont un PageRank élevé;
- R2.** On accorde d'autant moins de crédit à une référence, si elle provient d'une page qui dispose de nombreux liens.
- ⊕ Perturber *légèrement* le modèle initial pour obtenir une matrice "proche" mais pour laquelle on peut appliquer le théorème de Perron.

Algèbre, calcul matriciel, probabilités, analyse !

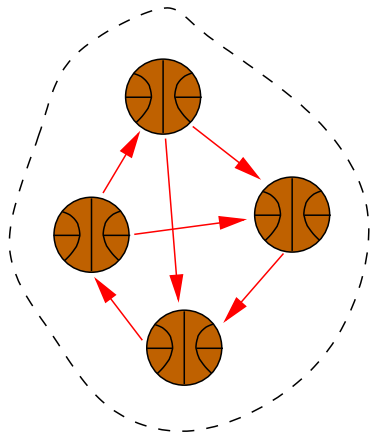
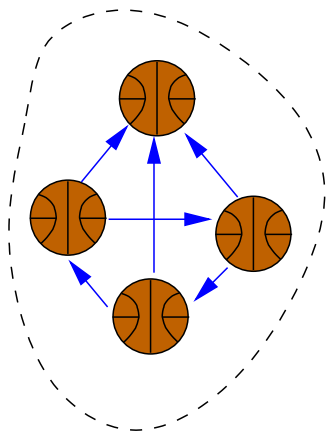
APPLICATION À LA NBA



APPLICATION À LA NBA



APPLICATION À LA NBA



APPLICATION À LA NBA

