Jacques Cellier

Jeu de mots

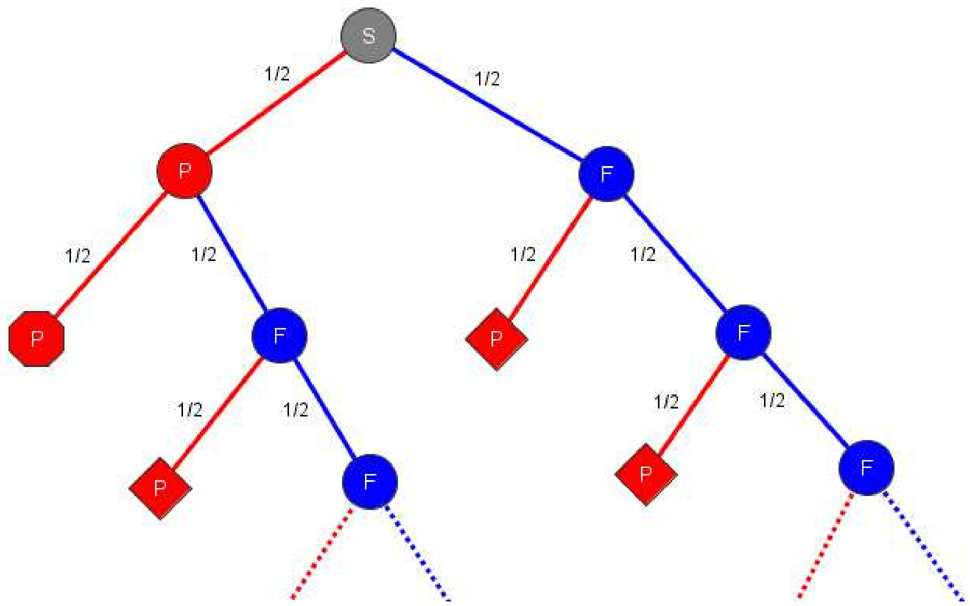
Introduction

Le jeu de pile ou face de Walter Penney

Ce jeu oppose 2 joueurs avec comme seul matériel requis une pièce de monnaie supposée équilibrée. Chacun choisit une séquence de résultats pile (P) ou face (F) de même longueurs[1]. Ils procèdent alors à des jets successifs de la pièce jusqu'à l'apparition d'une séquence choisie par un des joueurs lequel est déclaré vainqueur. Compte tenu de l'équiprobabilité entre pile et face, l'intuition première suggère que les deux joueurs ont la même chance de gain. Il n'en n'est rien : tous les mots ne sont pas égaux ! Avant de plonger dans les arcanes mathématiques du problème, un petit exemple permet d'entrevoir la raison de ce « paradoxe ».

Alice choisit le mot PP, Bob choisit FP.

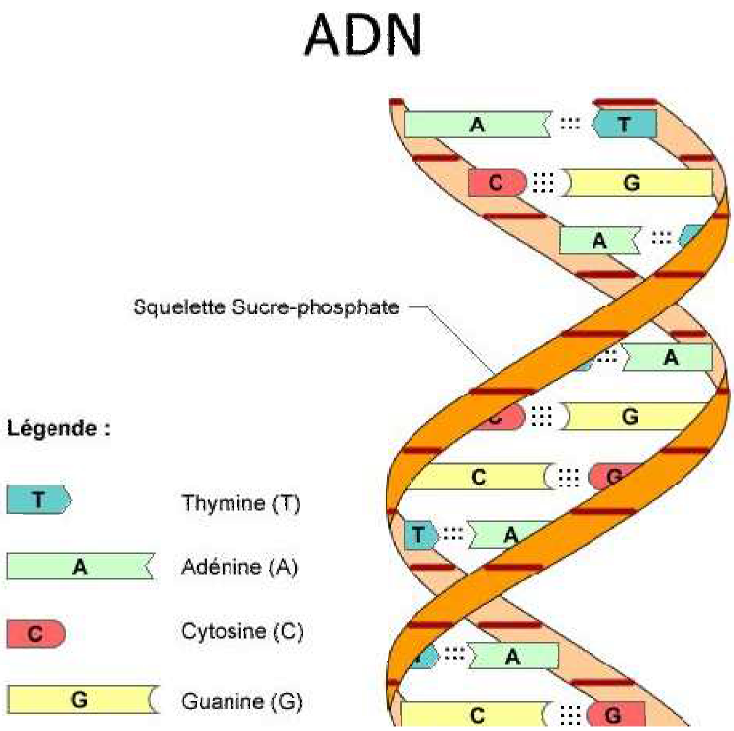
Au deuxième jet Alice a 1 chance sur 4 de gagner et Bob de même. Mais si Alice n'a pas gagné à ce deuxième jet, elle n'a plus aucune chance de gagner ensuite. Car, dans ce cas soit Bob a gagné, soit le dernier un jet a donné F et alors soit le suivant donne P assurant la victoire Bob soit il donne F et le problème se repose dans les mêmes termes. Alice n'a qu'une chance sur 4 de gagner contre 3 sur 4 pour Bob !



[1] Par la suite nous parlerons de mots sur l'alphabet $\Omega = \{P, F\}$.

Le jeu de la vie

Le premier exemple cité, même s'il conduit à des développements mathématiques intéressants, peut sembler futile. Mais on peut évoquer un exemple sur un autre alphabet $\Omega = \{A, C, G, T\}$ qui l'est beaucoup moins[2]…



[2] Voir [9].

Formalisation

Un dispositif émet de manière aléatoire et équiprobable et indépendante des lettres extraites d'un alphabet Ω de k lettres ($k \geq 2$). La probabilité d'apparition de chacune des lettres est donc égale à $\frac{1}{k}$. On va se pencher sur la question de l'apparition d'un mot donné dans la suite de lettres émises. Lorsque $k=2$ et $\Omega = \{P, F\}$ on se retrouve devant un jeu de pile ou face de Walter Penney.

Notations

Un mot est une liste de lettres appartenant à l'alphabet Ω de k lettres et pour un mot non vide A :

- ✓ $|A|$ désigne la longueur du mot.
- ✓ Si $A = a_1 a_2 \cdots a_n$ et si $1 \leq i \leq n$,
  - ${}_iA = a_1 \cdots a_i$ est le préfixe de longueur i du mot A,
  - $A_i = a_{n-i+1} \cdots a_n$ est le suffixe de longueur i du mot A.
- ✓ $P[A] = \frac{1}{k^n}$ la probabilité d'apparition de A en $|A| = n$ émissions si équiprobabilité..
- ✓ τ_A désigne la variable aléatoire égale au temps d'attente (nombre d'émissions) de la première apparition du mot A.
- ✓ $p_t(A) = P[\tau_A = t]$: probabilité que le mot A apparaisse la première fois à l émission t.
- ✓ $q_t(A) = P[\tau_A > t]$: probabilité que le mot A n'apparaisse pas dans les t premières émissions.
- ✓ Si Z est une proposition logique $1_Z = \begin{cases} 1 \text{ si } Z \text{ est vraie} \\ 0 \text{ sinon} \end{cases}$

Remarques :

Si $|A| = n$, alors :

- ✓ $p_0(A) = p_1(A) = \cdots = p_{n-1}(A) = 0$
- ✓ $q_0(A) = q_1(A) = \cdots = q_{n-1}(A) = 1$

Probabilité d'apparition d'un mot

1. Proposition

Soi A un mot de longueur $n > 1$ sur l'alphabet Ω. Son apparition dans une séquence infinie d'émissions est presque sûre.

Autrement dit, donnez à un singe un clavier et l'éternité, il finira par taper « *La recherche du temps perdu* »

DEMONSTRATION :

Pour $i = 1 \cdots n-1$, appelons intervalle d'ordre i un intervalle du type $t \cdots t+n-1$ où $t = i \bmod n$. Ainsi, les intervalles d'ordre 1 sont : $1 \cdots n, \quad n+1 \cdots 2n, \quad 2n+1 \cdots 3n, \quad$ etc …

La probabilité que le mot A n'apparaisse pas à un de ces intervalles est : $\left(1-\frac{1}{k^n}\right)$. Par indépendance, la probabilité que mot A n'apparaisse à aucun intervalle d'ordre i est donc égale à

$$\underset{s \to +\infty}{Lim}\left(1-\frac{1}{k^n}\right)^s = 0$$

L'évènement [le mot A n'apparaît jamais] est la réunion finie des évènements [le mot A n'apparaît jamais dans un intervalle d'ordre i], tous de probabilité nulle est donc lui-même de probabilité nulle. Donc l'évènement contraire [le mot A n'apparaît au moins une fois] est de probabilité 1.

Préliminaires

Relations entre $p_t(A)$ et $q_t(A)$

2. Proposition

Pour $t \geq 1$ *:*

1. $q_{t-1}(A) = q_t(A) + p_t(A)$
2. $q_t(A) = \sum_{i=t+1}^{+\infty} p_t(A)$
3. $E(\tau_A) = \sum_{i=1}^{+\infty} q_i(A)$

DEMONSTRATION :

1. $[\tau_A > t-1] = [\tau_A \geq t] = [\tau_A > t] \cup [\tau_A = t]$

2. $[\tau_A > t] = \bigcup_{i=t+1}^{+\infty} [\tau_A = i]$

3. Résulte de l'application du lemme suivant.

3. Lemme

Soit X *une variable aléatoire à valeur dans* $\mathbb{N}$, *alors :*

$$E(X) = \sum_{j=0}^{+\infty} P[X > j]$$

DEMONSTRATION :

$$E(X) = \sum_{i=1}^{+\infty} i\, P[X=i] = \sum_{i=1}^{+\infty}\sum_{j=1}^{i} P[X=i] = \sum_{j=1}^{+\infty}\sum_{i=j}^{+\infty} P[X=i] = \sum_{j=1}^{+\infty} P[X > j-1] = \sum_{j=0}^{+\infty} P[X > j]$$

4. Lemme

Pour un mot A *de longueur* n *:*

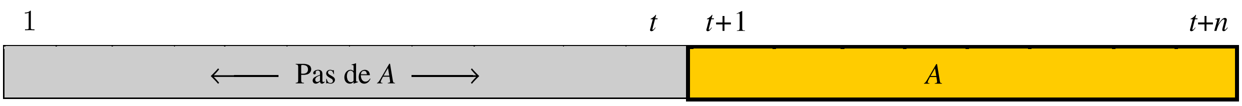
$$q_t(A) = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times p_{t+i}(A) \times \frac{P[A_{n-i}]}{P[A]}$$

Et, sous l'hypothèse d'équiprobabilité :

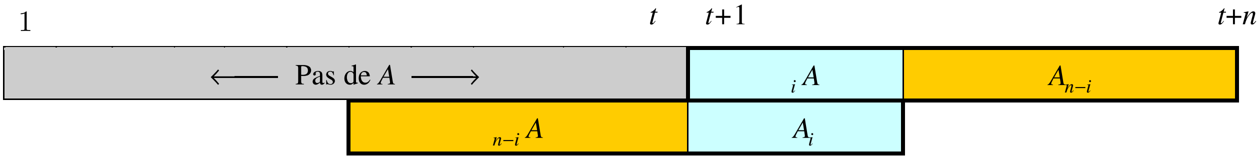
$$q_t(A) = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times p_{t+i}(A) \times k^i$$

DEMONSTRATION :

Soit M_t l'événement : A n'apparaît pas dans les t premières émissions et apparaît aux émissions $t+1 \cdots t+n$:



Par indépendance, on a : $P[M_t] = q_t(A) \times P[A]$. Cependant, le mot A peut apparaître, avant les émissions $t+1 \cdots t+n$. Plus précisément, pour $i = 1 \cdots n$, le mot A peut apparaître à l'émission : $t+i$ sous une condition de chevauchement : $A_i = {}_iA$. Voir schéma



L'évènement M_t est donc la réunion des événements disjoints $M_t^i =$ « le mot A apparaît la première fois à l'émission $t+i$ ». Si $A_i = {}_iA$, $P\left[M_t^i\right] = p_{t+i}(A) \times P\left[A_{n-i}\right]$ et finalement :

$$P\left[M_t\right] = q_t(A) \times P\left[A\right] = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times p_{t+i}(A) \times P\left[A_{n-i}\right]$$

Sous l'hypothèse d'équiprobabilité $P\left[A\right] = \dfrac{1}{k^n}$ et $P\left[A_{n-i}\right] = \dfrac{1}{k^{n-i}}$

$$q_t(A) = \sum_{i=1}^{n} 1_{A_i = {}_iA_i} \times p_{t+i}(A) \times k^i$$

Polynôme de corrélation de 2 mots

Ce polynôme reflète les possibilités de « chevauchement » d'un mot sur l'autre.

5. Définition

Soit A, B deux mots, le polynôme de corrélation de A sur B est définie par :

$$K_{A,B}\left(X\right) = \sum_{i=1}^{|A|} 1_{A_i = {}_iB} X^{i-1}$$

Exemple :

$A =$ PFPF, $B =$ FPFF

PFPF		
FPFF	0	
FPFF	1	$K_{A,B}\left(X\right) = X^2 + 1$
FPFF	0	
FPFF	1	

Remarques :

1. Si $|A| = n$, il s'agit donc d'un polynôme de degré inférieur ou égal à $n-1$ dont les coefficients valent 1 ou 0 selon que le suffixe de longueur *i* de *A* est égal ou non au préfixe de longueur *i* de *B*,.
2. En règle générale, $K_{A,B}\left(X\right) \neq K_{B,A}\left(X\right)$. Ainsi, dans l'exemple $K_{B,A}\left(X\right) = 0$.
3. Lorsque $A = B$, on parle de polynôme d'auto-corrélation. Comme $A_n = {}_nA$, le coefficient de degré $n-1$ de $K_{A,A}\left(X\right)$ est égal à 1.
4. Si $K_{A,B}\left(X\right) = \alpha_n X^{n-1} + \cdots + \alpha_2 X + \alpha_1$, $\alpha_n \cdots \alpha_2 \alpha_1$ est la représentation en base $k \geq 2$ du nombre $K_{A,B}\left(k\right) = \alpha_n k^{n-1} + \cdots + \alpha_2 k + \alpha_1$.
5. Si $A \neq B$, $K_{A,A}(k) > K_{A,B}(k)$ car le coefficient de degré $n-1$ de $K_{A,A}\left(X\right)$ est égal à 1 tandis que celui de $K_{A,B}\left(X\right)$ est égal à 0.

Temps moyen d'attente d'un mot

L'apparition d'un mot est presque sûre. C'est aussi le cas pour le retour à l'origine de l'ivrogne traité au chapitre précédent. Mais pour ce dernier le temps moyen de retour à l'origine est infini. Quant est-il pour le temps moyen d'apparition d'un mot ?

6. Proposition

Pour un mot A sur l'alphabet de k lettres Ω*, le temps moyen d'apparition dans une suite d'émissions, sous l'hypothèse d'équiprobabilité, est :*

$$E(\tau_A) = k \times K_{A,A}(k)$$

DEMONSTRATION :

Selon le corollaire 3 : $E(\tau_A) = \sum_{i=1}^{+\infty} q_i(A)$

Selon le lemme 4, pour tout $t \geq 0$: $q_t(A) = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times p_{t+i}(A) \times k^i$, on a, en tenant compte de $p_1(A) = \cdots = p_{n-1}(A) = 0$ et en notant $\alpha_i = 1_{A_i = {}_iA} \times k^i$:

$$\begin{array}{lllll}
q_0(A) = & \cdots\cdots\cdots & \cdots\cdots\cdots & & \alpha_n \times p_n(A) \\
q_1(A) = & \cdots\cdots\cdots & \cdots\cdots & \alpha_{n-1} p_n(A) + & \alpha_n p_{n+1}(A) \\
\vdots & & & \vdots & \\
q_{n-1}(A) = & \alpha_1 p_n(A) + & \cdots\cdots\cdots & + & \alpha_n p_{2n-1}(A) \\
\vdots & & & \vdots & \\
q_t(A) = & \alpha_1 p_{t+1}(A) + & \cdots\cdots\cdots & + & \alpha_n p_{t+n}(A) \\
\vdots & & & \vdots &
\end{array}$$

Par sommation :

$$\sum_{t=0}^{+\infty} q_t(A) = (\alpha_1 + \cdots + \alpha_n) \sum_{t=n}^{+\infty} p_t(A)$$

Comme A étant presque sûre et que $p_1(A) = \cdots = p_{n-1}(A) = 0$, $\sum_{t=n}^{+\infty} p_t(A) = 1$.

Par définition de α_i

$$\alpha_1 + \cdots + \alpha_n = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times k^i = k \sum_{i=1}^{n} 1_{A_i = {}_iA} \times k^{i-1} = k\, K_{A,A}(k)$$

Exemples :

Alphabet	Mot	$E(\tau_A)$
{P,F}	PP	6
:	FP	4
:	PFPF	20
:	FPFF	18
:	FPFFP	36
:	PPPPPPPPPP	2 046
{A,C,G,T}	ACGT	256
:	GGCCAATT	65 536

Alice et Bob jouent aux mots

En introduction, on a découvert, sur un exemple simple, l'inégalité des mots dans le jeu de Penney. Si Alice et Bob ayant chacun choisi un mot jouent, on sait s'après la proposition 1 que l'apparition de l'un ou l'autre des mots choisis est presque sûre. La question dont la réponse va décider du vainqueur est : quel mot va sortir en premier ? La proposition suivante, élargie aux alphabets de k lettres, donne la côte (rapport des probabilités) d'un mot A contre un mot B.

7. Proposition (formule de Conway)

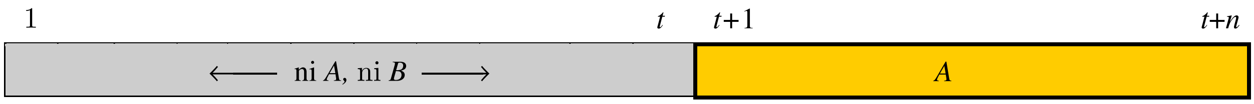
Si Alice choisit le mot A et Bob un autre mot B, de même longueur, dans un alphabet Ω de k lettres, le rapport de la probabilité de gain d'Alice sur celle de Bob (la cote) est, sous l'hypothèse d'équiprobabilité :

$$\frac{P[\text{Victoire d'Alice}]}{P[\text{Victoire de Bob}]}=\frac{K_{B,B}(k)-K_{B,A}(k)}{K_{A,A}(k)-K_{A,B}(k)}$$

DEMONSTRATION :

Cette formule a été énoncée par John Conway qui n'a pas publié de démonstration. Celle qui suit s'inspire de celle de Pavel A. Pevzner et reprend la démarche du lemme 4 et de la proposition 6.

Soit M_t l'événement : ni A, ni B n'apparaissent pas dans les t premières émissions et apparaît à l' émission $t+n$:



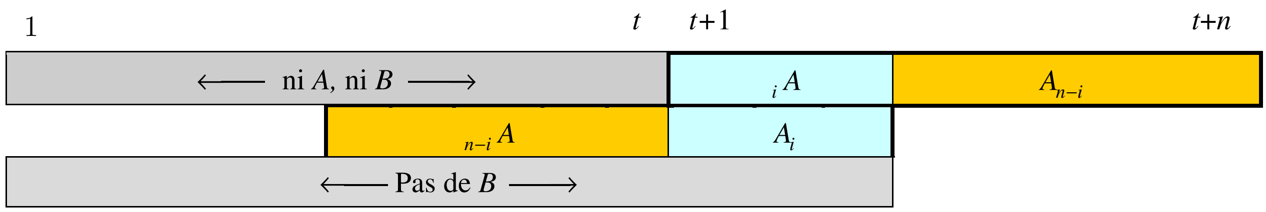
En notant $q_t(AB)$ la probabilité que ni A, ni B n'apparaissent pas dans les t premières émissions, il vient par indépendance : $P[M_t]=q_t(AB)\times P[A]$.

Par ailleurs, M_t peut se conclure par la victoire d'Alice avant l'émission $t+n$ sous condition de chevauchement $A_i={}_iA$. Et ceci se produit avec une probabilité égale à :

$$\sum_{i=1}^{n} 1_{A_i={}_iA}\times p'_{t+i}(A)\times P[A_{n-i}]$$

où $p'_{t+i}(A)$ est la probabilité que A apparaisse pour le première fois à l'émission $t+i$ avant B, autrement dit la probabilité que A l'emporte à l'émission $t+i$.

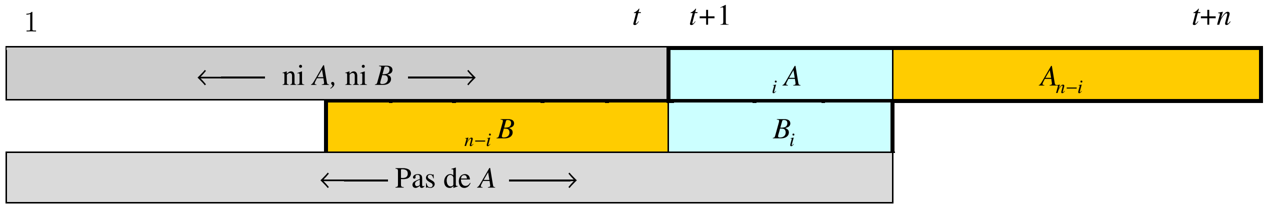
Selon le schéma :



Mais Bob peut l'emporter si le mot B se glisse avant A à l'émission $t+i$ pour $i=1\cdots n-1$ sous condition de chevauchement $B_i={}_iA$ et ceci avec une probabilité égale à :

$$\sum_{i=1}^{n-1} 1_{B_i = {}_iA} \times p'_{t+i}(B) \times P\left[A_{n-i}\right]$$

Selon le schéma :



Comme $A \neq B$, $1_{B_n = {}_nA} = 0$, ce qui permet d'étendre la sommation précédente au rang n. On obtient :

$$P\left[M_t\right] = q_t(AB) \times P\left[A\right] = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times p'_{t+i}(A) \times P\left[A_{n-i}\right] + \sum_{i=1}^{n-1} 1_{B_i = {}_iA} \times p'_{t+i}(B) \times P\left[A_{n-i}\right]$$

Puis par l'hypothèse d'équiprobabilité :

$$q_t(AB) = \sum_{i=1}^{n} 1_{A_i = {}_iA} \times k^i \times p'_{t+i}(A) + \sum_{i=1}^{n-1} 1_{B_i = {}_iA} \times k^i \times p'_{t+i}(B)$$

Il reste, comme pour la proposition 22, à procéder à une sommation qui va donner :

$$\sum_{t=0}^{+\infty} q_t(AB) = k\,K_{A,A}(k) \sum_{t=n}^{+\infty} p'_t(A) + k\,K_{B,A}(k) \sum_{t=n}^{+\infty} p'_t(B)$$

Sachant que $p'_t(A)$ (resp. $p'_t(B)$) est la probabilité de la victoire de A (resp. B) à l'émission t,

$$\sum_{t=n}^{+\infty} p'_t(A) = P\left[\text{Victoire d'Alice}\right]$$

$$\sum_{t=n}^{+\infty} p'_t(B) = P\left[\text{Victoire de Bob}\right]$$

Il s'en suit que :

$$\sum_{t=0}^{+\infty} q_t(AB) = k\,K_{A,A}(k) P\left[\text{Victoire d'Alice}\right] + k\,K_{B,A}(k) P\left[\text{Victoire de Bob}\right]$$

Mais rien n'empêche d'inverser A et B pour obtenir :

$$\sum_{t=0}^{+\infty} q_t(AB) = k\,K_{B,B}(k) P\left[\text{Victoire de Bob}\right] + k\,K_{A,B}(k) P\left[\text{Victoire d'Alice}\right]$$

De ces deux égalités, résulte :

$$K_{A,A}(k) P\left[\text{Victoire d'Alice}\right] + K_{B,A}(k) P\left[\text{Victoire de Bob}\right] =$$
$$K_{B,B}(k) P\left[\text{Victoire de Bob}\right] + K_{A,B}(k) P\left[\text{Victoire d'Alice}\right]$$

$$K_{A,A}(k) \frac{P\left[\text{Victoire d'Alice}\right]}{P\left[\text{Victoire de Bob}\right]} + K_{B,A}(k) = K_{B,B}(k) + K_{A,B}(k) \frac{P\left[\text{Victoire d'Alice}\right]}{P\left[\text{Victoire de Bob}\right]}$$

$$\frac{P\left[\text{Victoire d'Alice}\right]}{P\left[\text{Victoire de Bob}\right]} = \frac{K_{B,B}(k) - K_{B,A}(k)}{K_{A,A}(k) - K_{A,B}(k)}$$

Remarque :

Compte tenu de la relation qui joint la cote (c) d'un évènement à sa probabilité (p) :

$$c = \frac{p}{1-p} \Leftrightarrow p = \frac{c}{1+c}$$

il est aisé de déduire $P[\text{Victoire d'Alice}]$ ou $P[\text{Victoire de Bob}]$ de la formule de Conway.

Exemples :

Avec l'alphabet `{P,F}` :

✓ A=`PP`, B=`FP`

$$K_{B,B}(2) = 2\ ,\ K_{B,A}(2) = 1\ ,\ K_{A,A}(2) = 3\ ,\ K_{A,B}(2) = 0 \Rightarrow \frac{P[\text{Victoire d'Alice}]}{P[\text{Victoire de Bob}]} = \frac{1}{3}$$

On retrouve bien ce que nous avions vu en introduction 1 chance sur 4 pour Alice, contre 3 sur 4 pour Bob.

✓ A=`PPP`, B=`FPP`

$$K_{B,B}(2) = 4\ ,\ K_{B,A}(2) = 3\ ,\ K_{A,A}(2) = 7\ ,\ K_{A,B}(2) = 0 \Rightarrow \frac{P[\text{Victoire d'Alice}]}{P[\text{Victoire de Bob}]} = \frac{1}{7}$$

✓ A=`PFPF`, B=`FPFF`

$$K_{B,B}(2) = 9\ ,\ K_{B,A}(2) = 0\ ,\ K_{A,A}(2) = 10\ ,\ K_{A,B}(2) = 5 \Rightarrow \frac{P[\text{Victoire d'Alice}]}{P[\text{Victoire de Bob}]} = \frac{9}{5}$$

Mais, surgit un nouveau paradoxe : le résultat est en faveur d'Alice or le temps moyen d'attente du mot d'Alice est de 20 contre 18 pour celui de Bob ! On ne peut donc se fonder sur les temps moyens d'attente pour savoir quel est le vainqueur le plus probable.

Avec l'alphabet `{A,C,G,T}`

✓ A= `AGCTA`, B = `AAGCT`

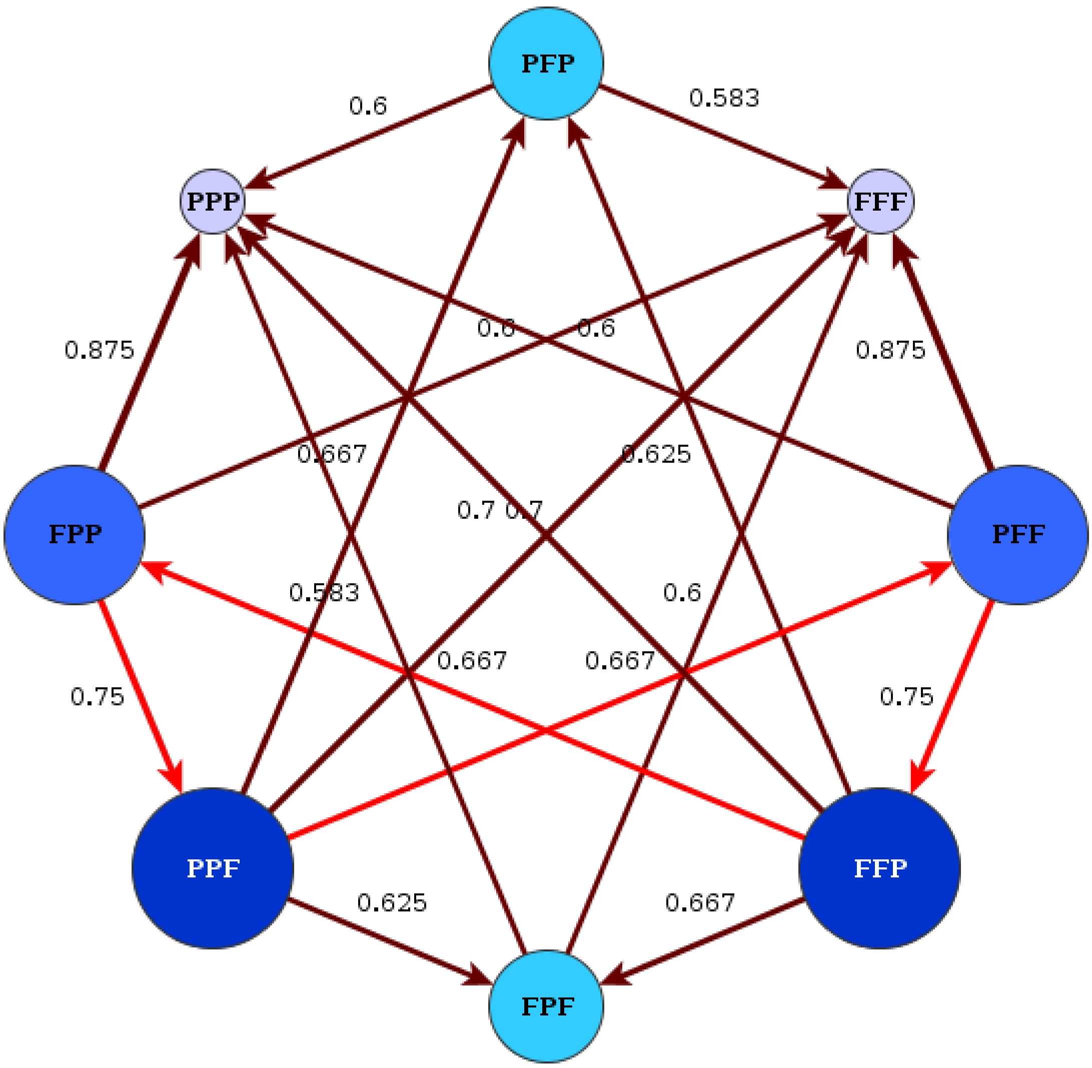
$$K_{B,B}(4) = 256\ ,\ K_{B,A}(4) = 64\ ,\ K_{A,A}(4) = 257\ ,\ K_{A,B}(4) = 1 \Rightarrow \frac{P[\text{Victoire d'Alice}]}{P[\text{Victoire de Bob}]} = \frac{192}{256} = \frac{3}{4}$$

Autour de la relation de domination d'un mot sur l'autre

Sur l'ensemble Ω^n des mots de n lettres, on peut définir une relation binaire $\succ$, dite de domination, par :

$$A \succ B \text{ si et seulement si } P[\text{Victoire d'Alice}] > P[\text{Victoire de Bob}]$$ [3]

Le graphe de cette relation est représenté pour $\{\mathtt{P},\mathtt{F}\}^3$ sur la figure ci-dessous. On peut y observer :



Représentation du graphe de $\prec$ sur $\{\mathtt{P},\mathtt{F}\}^3$.

L'aire des disques est proportionnelle au nombre de « victoires ».

Les nombres sur les flèches indiquent la probabilité de la victoire.

L'absence de flèches entre deux mots signifie «match nul »

- ✓ Cette relation n'est pas totale. Si $P[\text{Victoire d'Alice}] = P[\text{Victoire de Bob}]$ alors les 2 mots font jeu égal. Et l'on a ni $A \succ B$ ni $B \succ A$. Exemple : `PPP` et `FFF`.

---

[3] En supposant, bien sûr, qu'Alice choisisse le mot A et Bob le mot B.

- ✓ Il existe des jamais gagnants : `FFF` et `PPP`. Ils font match nul entre eux et curieusement aussi avec des « très forts »
- ✓ Mais aucun mot n'échappe à la défaite. Même les mots « les plus forts » comme `PPF` et `FFP` sont dominés (par `FPP` pour le premier, par `PFF` pour le second).
- ✓ La relation n'est pas transitive : `FPP` $\succ$ `PPF` et `PPF` $\succ$ `FPF`, mais `FPP et FPF` font jeu égal.
- ✓ Il existe des cycles comme celui représenté en rouge sur le graphe.

Comment Bob peut battre Alice ?

Supposant qu'Alice choisisse son mot A au vu de Bob. Ce dernier peut toujours choisir un mot B tel que $A \prec B$ et qui lui assure une victoire statistique. L'observation du graphe précédent montre que, pour les mots de trois lettres sur l'alphabet `{P,F}` que tout mot a un « dominant ». Mais comment, en général, en déterminer un ?

8. Proposition

Si Alice choisit le mot $A = a_1 a_2 \cdots a_n$ $(n \geq 3)$ *et Bob choisit* $B = b a_1 a_2 \cdots a_{n-1}$, *alors, pour un choix « judicieux » de b :*

$$P[\text{Victoire de Bob}] > P[\text{Victoire d'Alice}]$$

DEMONSTRATION :

Remarques préliminaires :

1. Pour le moment, le choix judicieux sera : $b \neq a_2$.
2. $A \neq B$. En supposons le contraire, on aurait : $b = a_1$ et $a_1 = a_2$ en contradiction avec le choix de $b \neq a_2$.
3. Soit 2 polynômes de corrélation $K_1(X)$ et $K_2(X)$. Si le degré de $K_1(X)$ est strictement supérieur au degré de $K_2(X)$, alors $K_1(k) > K_2(k)$. Cela résulte de la remarque 4 de la définition 5 : le nombre $K_1(k)$ ayant une représentation en base k « plus longue » que $K_2(k)$.

Minoration et majoration des termes de la formule de Conway :

Nous allons nous pencher sur les termes de plus haut degré des polynômes $K_{\bullet,\bullet}(X)$.

1. Comme $K_{A,A}(X)$ est un polynôme de degré $n-1$:

$$K_{A,A}(k) \geq k^{n-1} \quad (1)$$

2. Pour $K_{B,B}(X)$, on a :

   $1_{B_n = {}_nB} = 1$

   $1_{B_{n-1} = {}_{n-1}B} = 0$. En effet, $B_{n-1} = a_1 a_2 \cdots a_{n-1}$ et ${}_{n-1}B = b a_1 a_2 \cdots a_{n-2}$. Si ils étaient égaux, on aurait : $b = a_1$ et $a_1 = a_2$ en contradiction avec le choix de $b \neq a_2$.

   $1_{B_{n-2} = {}_{n-2}B} = 0$ En effet, $B_{n-2} = a_2 \cdots a_{n-1}$ et ${}_{n-2}B = b a_1 a_2 \cdots a_{n-3}$ et $b \neq a_2$.

   Donc $K_{B,B}(X) = X^{n-1} + 0X^{n-2} + 0X^{n-3} + \cdots$ et finalement :

$$K_{B,B}(k) < k^{n-1} + k^{n-3} \quad (2)$$

3. Pour $K_{A,B}(X)$, on a :

   $1_{A_n = {}_nB} = 0$ puisque $A \neq B$.

   $1_{A_{n-1} = {}_{n-1}B} = 0$. En effet : $A_{n-1} = a_2 \cdots a_n$ et ${}_{n-1}B = ba_1a_2\cdots a_{n-2}$ et $b \neq a_2$.

   Si $\boxed{k \geq 3}$, en complétant le choix « judicieux » de b par $b \neq a_3$, on a :

   $1_{A_{n-2} = {}_{n-2}B} = 0$. En effet : $A_{n-2} = a_3 \cdots a_n$ et ${}_{n-2}B = ba_1a_2\cdots a_{n-3}$

   Donc $K_{A,B}(X) = 0X^{n-1} + 0X^{n-2} + 0X^{n-3} + \cdots$ et finalement :

   $$K_{A,B}(k) < k^{n-3} \quad (3)$$

4. Pour $K_{B,A}(X)$, on a :

   $1_{B_n = {}_nA} = 0$ puisque $A \neq B$.

   $1_{B_{n-1} = {}_{n-1}A} = 1$. En effet : $B_{n-1} = a_1a_2\cdots a_{n-1}$ et ${}_{n-1}A = a_1a_2\cdots a_{n-1}$.

   Donc $K_{B,A}(X) = 0X^{n-1} + 1X^{n-2} + \cdots$ et finalement :

   $$K_{B,A}(k) \geq k^{n-2} \quad (4)$$

Il reste à rassembler les morceaux :

(1) et (3) $\Rightarrow K_{A,A}(k) - K_{A,B}(k) > k^{n-1} - k^{n-3}$

(2) et (4) $\Rightarrow K_{B,B}(k) - K_{B,A}(k) < k^{n-1} + k^{n-3} - k^{n-2}$

D'où :

$$\frac{P[\text{Victoire de Bob}]}{P[\text{Victoire d'Alice}]} = \frac{K_{A,A}(k) - K_{A,B}(k)}{K_{B,B}(k) - K_{B,A}(k)} > \frac{k^{n-1} - k^{n-3}}{k^{n-1} + k^{n-3} - k^{n-2}} = \frac{k^2 - 1}{k^2 - k + 1} \geq 1 \text{ pour } k \geq 2$$

Mais il reste le cas $k = 2$ et qui est précisément celui du jeu de pile ou face de Penney. Si $A_{n-2} \neq {}_{n-2}B$, l'argumentation précédente reste valide, mais quid du cas : $A_{n-2} = {}_{n-2}B$?

Dans ce cas on a : $a_3a_4a_5\cdots a_n = ba_1a_2\cdots a_{n-3}$ avec $b \neq a_2$. Il en résulte une périodicité de A :

$$\left|\begin{aligned} & a_1 = a_4 = a_7 = \cdots a_{1+3i} \\ & a_2 = a_5 = a_8 = \cdots a_{2+3i} \\ b = {}& a_3 = a_6 = a_9 = \cdots a_{3+3i} \end{aligned}\right.$$

Suivant la valeur de $n \bmod 3$, le mot A est donc d'une des 3 formes :

0. $A = \underbrace{a_1a_2a_3\cdots a_1a_2a_3}_{i \text{ répétitions}}$, et alors $B = \underbrace{a_3a_1a_2\cdots a_3a_1a_2}_{i \text{ répétitions}}$
1. $A = \underbrace{a_1a_2a_3\cdots a_1a_2a_3}_{i \text{ répétitions}}\, a_1$ et alors $B = \underbrace{a_3a_1a_2\cdots a_3a_1a_2}_{i \text{ répétitions}}\, a_3$
2. $A = \underbrace{a_1a_2a_3\cdots a_1a_2a_3}_{i \text{ répétitions}}\, a_1a_2$ et alors $B = \underbrace{a_3a_1a_2\cdots a_3a_1a_2}_{i \text{ répétitions}}\, a_3a_1$

Quant au triplet $a_1a_2a_3$, sachant que $a_3 = b$ est différent de a_2 il doit être d'une des 4 forme suivante : `FPF`, `PPF`, `PFP`, `FFP`. Par raison de symétrie, on peut se limiter à l'étude des deux premiers triplets.

On va se pencher sur l'écriture binaire des $K_{\bullet,\bullet}(2)$ dont on rappelle que les bits sont les coefficients des polynômes $K_{\bullet,\bullet}(X)$.

Dans le cas de la forme 0 pour A :

- ✓ Pour $K_{A,A}(2)$, on a le chevauchement trivial, $\begin{array}{l} a_1a_2a_3\cdots a_1a_2a_3 \\ a_1a_2a_3\cdots a_1a_2a_3 \end{array}$ puis compte tenu de la périodicité tous les 3 décalages, jusqu'à $\begin{array}{ll} a_1a_2a_3\cdots a_1a_2a_3 & \\ & a_1a_2a_3\cdots a_1a_2a_3 \end{array}$. Sous forme binaire : $K_{A,A}(2)=\underbrace{100\,100\cdots100}_{i-1\text{ répétitions}}t_{A,A}$ où $t_{A,A}=1\varepsilon\varepsilon$ avec $\varepsilon=0$ ou 1.
- ✓ Pour $K_{B,B}(2)$, c'est la même chose : $K_{B,B}(2)=\underbrace{100\,100\cdots100}_{i-1\text{ répétitions}}t_{B,B}$.
- ✓ Pour $K_{A,B}(2)$, le premier chevauchement sera obtenu avec un décalage de 2 et $\begin{array}{ll} a_1a_2a_3\cdots a_1a_2a_3 & \\ & a_3a_1a_2\cdots a_3a_1a_2 \end{array}$ et se répétera tous les trois décalages jusqu'à : $\begin{array}{ll} a_1a_2a_3\cdots a_1a_2a_3 & \\ & a_3a_1a_2\cdots a_3a_1a_2 \end{array}$ où sachant que $a_3\neq a_2$, il n'y aura pas de chevauchement. Sous forme binaire : $K_{A,B}(2)=\underbrace{001\,001\cdots001}_{i-1\text{ répétitions}}t_{A,B}$, où $t_{A,B}=0\varepsilon\varepsilon$ avec $\varepsilon=0$ ou 1.
- ✓ Pour $K_{B,A}(2)$, le premier chevauchement sera obtenu avec un décalage de 1 $\begin{array}{ll} a_3a_1a_2\cdots a_3a_1a_2 & \\ & a_1a_2a_3\cdots a_1a_2a_3 \end{array}$ et se répétera tous les trois décalages jusqu'à : $\begin{array}{ll} a_3a_1a_2\cdots a_3a_1a_2 & \\ & a_1a_2a_3\cdots a_1a_2a_3 \end{array}$ où la encore le chevauchement n'aura pas lieu. Sous forme binaire : $K_{B,A}(2)=\underbrace{010\,010\cdots010}_{i-1\text{ répétitions}}t_{B,A}$ où $t_{A,B}=0\varepsilon\varepsilon$ avec $\varepsilon=0$ ou 1.

Pour la formes 1 (resp. 2) de A, on obtient les mêmes résultats à ceci près que les bits terminaux comportent 3 ε (resp. 4 ε).

Les valeurs des bits terminaux dépendent de $n \bmod 3$ et du triplet répété `FPF` ou `PPF`. Le tableau suivant fournit ces valeurs pour les différent cas :

	A	B	*n mod 3*	$t_{A,A}$	$t_{A,B}$	$t_{B,B}$	$t_{B,A}$	$(t_{A,A}-t_{A,B})-(t_{B,B}-t_{B,A})$
Binaire	PPF	FPP	0	100	001	100	011	
	PPFP	FPPF	1	1001	0010	1001	0100	
	PPFPP	FPPFP	2	10011	00100	10010	01001	
	FPF	FFP	0	101	001	100	010	
	FPFF	FFPF	1	1001	0011	1001	0101	
	FPFFP	FFPFF	2	10010	00100	10011	01001	
Décimal	PPF	FPP	0	4	1	4	3	2
	PPFP	FPPF	1	9	2	9	4	2
	PPFPP	FPPFP	2	19	4	18	9	6
	FPF	FFP	0	5	1	4	2	2
	FPFF	FFPF	1	9	3	9	5	2
	FPFFP	FFPFF	2	18	4	19	9	4

Plutôt que d'examiner les cas un par un, ce qui ne présente pas de difficultés mais serait quelque peu fastidieux, on va se livrer à une approche synthétique.

Posons $x_i = \underbrace{001\cdots001}_{i-1 \text{ répétitions}}$ avec la convention $x_1 = 1$. Sachant qu'une multiplication par 2 décale les bits d'un cran vers la gauche, on peut écrire :

- ✓ $K_{A,A}(2) = x_i 2^{s+2} + t_{A,A}$
- ✓ $K_{B,B}(2) = x_i 2^{s+2} + t_{B,B}$
- ✓ $K_{A,B}(2) = x_i 2^{s} + t_{B,B}$
- ✓ $K_{B,A}(2) = x_i 2^{s+1} + t_{B,A}$

où $s = 3, 4, 5$ selon que les $t_{\bullet,\bullet}$ comportent 3, 4 ou 5 bits.

$$\begin{aligned}\left(K_{A,A}(2) - K_{A,B}(2)\right) - \left(K_{B,B}(2) - K_{B,A}(2)\right) &= x_i 2^{s+1} - x_i 2^{s} + \left(t_{A,A} - t_{A,B}\right) - \left(t_{B,B} - t_{B,A}\right) \\ &= x_i 2^{s} + \left(t_{A,A} - t_{A,B}\right) - \left(t_{B,B} - t_{B,A}\right)\end{aligned}$$

Or selon le tableau : $\left(t_{A,A} - t_{A,B}\right) - \left(t_{B,B} - t_{B,A}\right) \geq 2$ et donc :

$$\left(K_{A,A}(2) - K_{A,B}(2)\right) - \left(K_{B,B}(2) - K_{B,A}(2)\right) \geq x_i 2^{s} + 2$$

Ce qui assure la victoire statistique de Bob.

Remarques :

- ✓ On peut montrer que, pour $i > 1$, $x_i = 1 + 2^3 + \cdots 2^{3(i-1)} = \dfrac{8^i - 1}{7}$
- ✓ La stratégie de Bob consistant à choisir $b \neq a_2$ et éventuellement $b \neq a_3$ lui donne une domination sur Alice, mais rien dans ce que nous avons démontré ne garantit qu'elle est optimale. L. J. GUIBAS and A. M. ODLYZKO dans [6] et Daniel FELIX DANS [4] abordent la question de la stratégie optimale, mais c'est nettement plus compliqué.

Une autre approche via les chaînes de Markov

Introduction : graphe de transitions

Construction du graphe $G=(S,F)$ associé à un mot $A=a_1a_2\cdots a_n$

L'idée sous-jacente est que le mot A va s'écrire, au cours des émissions aléatoires, lettre par lettre. On définit alors l'ensemble S dit espace des « états » comme étant l'ensemble des préfixes de A (Ø et A compris) :

$$S=\{\varnothing, a_1, a_1a_2, \cdots, a_1a_2\cdots a_n\}.$$

Ces $n+1$ états sont les nœuds du graphe.

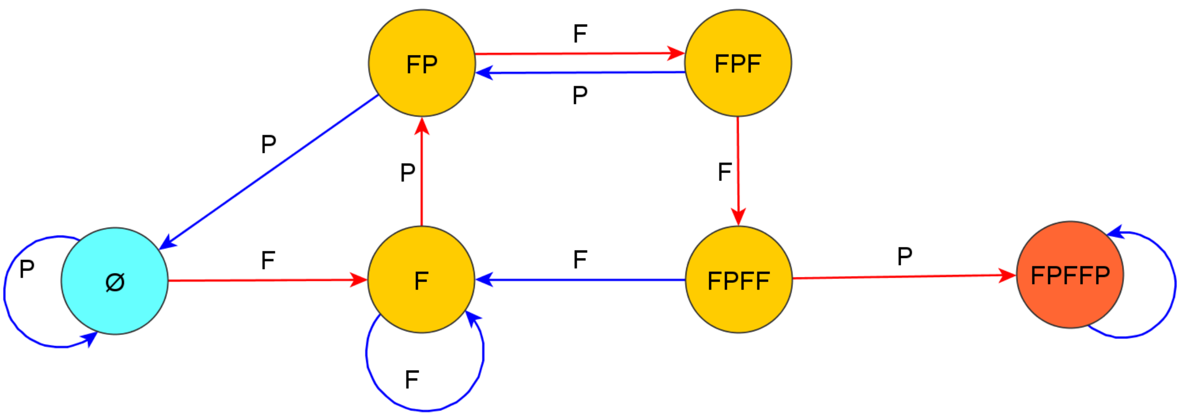
Soit : ${}_rA\in S$ un préfixe de longueur $0\leq r<n$.

- ✓ On met une flèche étiquetée par a_{r+1} du nœud ${}_rA$ vers le nœud ${}_{r+1}A$. Explication : si on se trouve dans l'état ${}_rA$ et que la lettre émise à l'étape $r+1$ est précisément a_{r+1}, on avance dans la construction de A et on passe à l'état ${}_{r+1}A$.
- ✓ Soit $b\in\Omega, b\neq a_{r+1}$. On met une flèche étiquetée par b vers l'état de longueur m maximale tel que $(a_1a_2\cdots a_rb)_m\in S$. Explication : si la lettre émise à l'étape $r+1$ est différente de b, tout n'est pas perdu ! Il se peut qu'un suffixe de $a_1a_2\cdots a_rb$ soit un préfixe de A, auquel cas on retourne à cet état (qui peut être éventuellement $\varnothing$).

Enfin, on met une boucle sur A. Explication : arrivé à l'état A, le jeu est gagné et on reste dans le même état.

Le graphe $G=(S,F)$ ainsi obtenue est un graphe orienté avec boucles. Partant d'un état quelconque, il existe toujours un chemin vers l'état final A. Le degré extérieur de chaque état $\neq A$ est égal à $k=|\Omega|$.

Exemple 1 : A = FPFFP



- ✓ Partant de l'état $\varnothing$, tant qu'aucun F n'est émis, on reste dans cet état.
- ✓ Partant de l' état FPF, si un P est émis, alors le suffixe FP de FPFP est le plus long appartenant à S d'où la flèche vers FP.
- ✓ En revanche partant de l'état FP, si P est émis, alors aucun suffixe de FPP n'appartient à S d'où la flèche vers $\varnothing$.

Construction du graphe $G=(S,F)$ associé à 2 mots

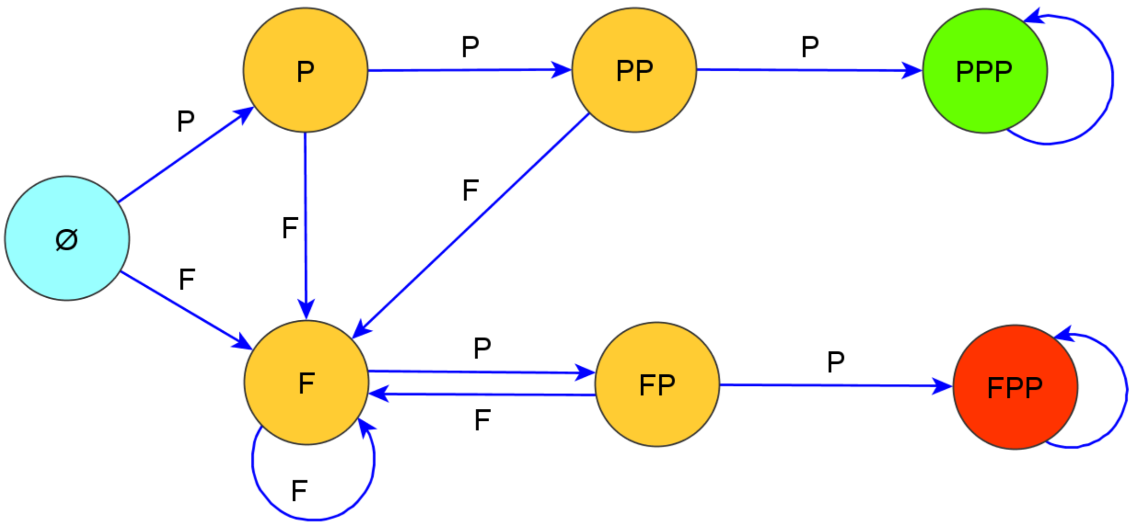
$$A=a_1a_2\cdots a_n\ ,\ B=b_1b_2\cdots b_n$$

La construction est similaire au cas d'un seul mot. La différence essentielle réside dans la définition de l'espace S des états qui est alors la réunion des préfixes de A et des préfixes de B.

$$S=\{\varnothing, a_1, b_1, a_1a_2, b_1b_2, \cdots, A, B\}$$

Pour les flèches, le principe est le même avec une boucle sur chaque état final A et B.

Exemple 2 : A = PPP, B = FPP



On peut remarquer qu'un seul chemin mène de l'état $\varnothing$ vers l'état PPP, alors que plusieurs mènent de l'état $\varnothing$ vers l'état FPP. C'est une illustration de l'avantage de B sur A (pour mémoire la côte de B contre A est de 7).

Des graphes aux matrices stochastiques

Les graphes précédents sont entièrement décrits par des tableaux du type :

	Ø	F	FP	FPF	FPFF	FPFFP
Ø	P	0	P	0	0	0
F	F	F	0	0	F	0
FP	0	P	0	P	0	0
FPF	0	0	F	0	0	0
FPFF	0	0	0	F	0	0
FPFFP	0	0	0	0	P	1

	Ø	P	F	PP	FP	PPP	FPP
Ø	0	0	0	0	0	0	0
P	P	0	0	0	0	0	0
F	F	F	F	F	F	0	0
PP	0	P	0	0	0	0	0
FP	0	0	P	0	0	0	0
PPP	0	0	0	P	0	1	0
FPP	0	0	0	0	P	0	1

En rouge, les états finaux.

Si on remplace dans les tableaux précédents les étiquettes des flèches par les probabilités, correspondantes, par exemple, pour le jeu de pile ou face équiprobable, P et F par $\frac{1}{2}$, le graphe décrit les probabilité de transition d'un état vers un autre. À chaque graphe de ce type on peut associer une matrice $M=\left(M_{i,j}\right)$ avec $M_{i,j}=P[i/j]$ la probabilité de passer à l'état i à partir de l'état j. On a donc : $M_{i,j}=\begin{cases}>0 \text{ si } j\to i\in F\\ 0 \text{ sinon}\end{cases}$. Par exemple, pour les deux exemples précédents :

$$M_1=\begin{pmatrix}1/2&0&1/2&0&0&0\\1/2&1/2&0&0&1/2&0\\0&1/2&0&1/2&0&0\\0&0&1/2&0&0&0\\0&0&0&1/2&0&0\\0&0&0&0&1/2&1\end{pmatrix}\qquad M_2=\begin{pmatrix}0&0&0&0&0&0&0\\1/2&0&0&0&0&0&0\\1/2&1/2&1/2&1/2&1/2&0&0\\0&1/2&0&0&0&0&0\\0&0&1/2&0&0&0&0\\0&0&0&1/2&0&1&0\\0&0&0&0&1/2&0&1\end{pmatrix}$$

De telles matrices à termes positifs ou nuls dont la somme de chaque colonne est égale à 1 sont dites stochastiques colonnes[4]. Elles décrivent complètement les probabilités de transitions. Leur étude permet de retrouver les résultats précédents, mais requièrent des notions concernant l'algèbre linéaire et les chaînes de Markov à états finis. On pourra se nourrir sur ces sujets avec le [2] pour l'algèbre linéaire en général et avec le [10] pour les chaînes de Markov en particulier.

[4] La convention usuelle en probabilité est d'utiliser des matrices stochastiques ligne. Mais, par habitude de l'algèbre linéaire nous préférons la convention inverse.

Dans la section suivante, on va s'efforcer de donner un aperçu des résultats utiles pour notre problème de jeu de mots.

Chaines de Markov

Brève introduction aux chaines de Markov

Elles servent à modéliser l'évolution d'un système évoluant aléatoirement d'un état à l'autre au fil du temps (discret). Soit $S=\{e_1,\cdots,e_s\}$ un ensemble de ces s états que l'on désignera plus simplement par leur indice. Pour $\tau\in\mathbb{N}$, on notera X_τ la variable aléatoire représentant l'état du système à l'instant τ et $P_\tau=\begin{pmatrix} P[X_\tau=1] \\ \vdots \\ P[X_\tau=s] \end{pmatrix}$ le vecteur des probabilités à l'instant τ. Sachant que le système doit, à chaque instant se trouver dans un état de *S*, on a bien sûr : $\sum_{i=1}^{s} P[X_\tau=i]=1$.

Les chaines de Markov sont à mémoire « courte » et « homogène » ce qui se traduit par la formule suivante :

$$\forall\tau\in\mathbb{N}: \quad P[X_{\tau+1}=i/X_\tau=j]=P[X_1=i/X_0=j]$$

En notant $M_{i,j}=P[X_1=i/X_0=j]$ on définit une matrice dite de transition qui décrit complètement le comportement de la chaîne. Cette matrice dite stochastique colonne a les propriétés suivantes :

1. Tous ses termes appartiennent à $[0\cdots1]$.
2. La somme de chacune des colonnes est égale à 1.
3. Son rayon spectral $\rho(M)$ est égal à 1 et elle admet 1 comme valeur propre.
4. $\forall\tau\in\mathbb{N}: \quad P_{\tau+1}=MP_\tau$.
5. $\forall\tau\in\mathbb{N}: \quad P_\tau=M^\tau P_0$.
6. $\forall\tau\in\mathbb{N}: \quad M^\tau$ est une matrice stochastique colonne.
7. $(M^\tau)_{i,j}$ représente la probabilité d'atteindre l'état i à partir de l'état j en τ étapes. Autrement dit : $(M^\tau)_{i,j}=P[X_\tau=i/X_0=j]$.

ATTENTION : $(M^\tau)_{i,j}$ désigne le terme ligne i colonne j de la matrice M^τ (et non le terme $M_{i,j}$ de la matrice M élevé à la puissance τ).

En vue de nous confronter à notre problème de jeu de mots, on va se pencher sur un type particulier de chaînes de Markov : les chaînes absorbantes.

Chaînes absorbantes

9. Définitions

1. *Un état i est dit accessible à partir d'un état j si et seulement si il existe un entier* τ *tel que :* $(M^{\tau})_{i,j} > 0$
2. *Un état i est dit absorbant si et seulement si* $M_{i,i} = 1$. *On notera* $\mathcal{A}$*, l'ensemble des états absorbants.*
3. *Une chaîne de Markov est dite absorbante, si et seulement si :*
   - ✓ *il existe au moins un état absorbant,*
   - ✓ *pour tout état j, il existe un état absorbant i accessible à partir de j.*

Commentaires :

En terme de graphe :

1. signifie qu'il existe un chemin de j vers i.
2. signifie qu'un état absorbant n'émet aucune flèche sauf la boucle sur lui-même.
3. signifie qu'à partir d'un état quelconque, il existe toujours un chemin vers un état absorbant.

En observant nos deux graphes exemple, on peut d'ores et déjà constater qu'ils possèdent des états absorbants (A pour le premier, A et B pour le second) et qu'à partir de tout état, il existe un chemin vers un état absorbant.

En plaçant en dernier les états absorbant (donc avec un indice dans l'intervalle $[s-a+1 \cdots s]$), la matrice M de taille $s \times s$ d'une chaîne absorbante ayant a états absorbants peut s'écrire sous forme dite canonique : $M = \left(\begin{array}{c|c} Q & 0 \\ \hline R & I \end{array}\right)$ où :

- ✓ Q est une matrice carrée de taille : $s-a$
- ✓ I est la matrice identité de taille a.
- ✓ 0 est la matrice nulle de taille $(s-a) \times a$
- ✓ R est une matrice de taille $a \times (s-a)$

Pour toutes les propositions qui vont suivre, on suppose désormais que <u> P est la matrice d'une chaîne absorbante mise sous forme canonique.</u>

10. Proposition

Pour $\tau \geq 1$, $M^{\tau} = \left(\begin{array}{c|c} Q^{\tau} & 0 \\ \hline R\left(\sum_{i=0}^{\tau-1} Q^i\right) & I \end{array}\right)$

DEMONSTRATION :

Une simple récurrence suffit…

11. Proposition

$\underset{\tau \to +\infty}{Lim}\, Q^{\tau} = 0$

DEMONSTRATION :

Soit $j \notin \mathcal{A}$ un état non absorbant. Par définition d'une chaîne absorbante, il existe $\tau \geq 1$ et $v \in \mathcal{A}$ tels que : $(M^{\tau})_{v,j} > 0$. On peut observer de plus que : $(M^{\tau})_{v,j} > 0 \Rightarrow (M^{\tau+1})_{v,j} > 0$. En effet :

$$(M^{\tau+1})_{v,j} = \sum_{c=1}^{s} M_{v,c} (M^{\tau})_{c,j} \geq \underbrace{M_{v,v}}_{=1} (M^{\tau})_{v,j} = (M^{\tau})_{v,j}$$

Définissons alors T comme la première date pour laquelle $\forall j \notin \mathcal{A}, \exists v \in \mathcal{A} : (M^{\tau})_{v,j} > 0$.

Comme la somme de la colonne j de M^T est égale à 1, on a :

$$\forall j \notin \mathcal{A} : \sum_{i \notin \mathcal{A}} (M^T)_{i,j} = \sum_{i \notin \mathcal{A}} (Q^T)_{i,j} < 1$$

Or $\left\| Q^T \right\|_1 = \max_{j \notin \mathcal{A}} \left(\sum_{i \notin \mathcal{A}} (Q^T)_{i,j} \right)$, et[5] $\rho\left(Q^T\right) \leq \left\| Q^T \right\|_1$. Donc : $\rho\left(Q^T\right) < 1$ et $\rho(Q) < 1$. Cette dernière inégalité assurant que : $\underset{r \to +\infty}{Lim} (Q^{\kappa}) = 0$.

12. Proposition et définition

$I - Q$ est inversible et $N = (I - Q)^{-1} = \sum_{\tau=0}^{+\infty} Q^{\tau}$. La matrice N est appelée <u>matrice fondamentale</u> de la chaîne absorbante de matrice M.

DEMONSTRATION :

$$\begin{aligned}(I - Q)\left(I + Q + \cdots + Q^{\tau}\right) &= I + Q + \cdots + Q^{\tau} \\ &\quad - Q - \cdots - Q^{\tau} - Q^{\tau+1} = I - Q^{\tau+1}\end{aligned}$$

Et d'après la proposition précédente : $(I - Q) \underset{\tau \to +\infty}{Lim} \left(I + Q + \cdots + Q^{\tau}\right) = I$. Donc $(I - Q)$ est inversible et $(I - Q)^{-1} = \sum_{\tau=0}^{+\infty} Q^{\tau}$.

Cette matrice fondamentale va nous livrer beaucoup d'informations sur le comportement de la chaîne :

13. Théorème

Soit N la matrice fondamentale d'une chaîne absorbante mise sous forme canonique $M = \left(\begin{array}{c|c} Q & 0 \\ \hline R & I \end{array}\right)$, alors :

1. $\underset{\tau \to +\infty}{Lim} M^{\tau} = \left(\begin{array}{c|c} 0 & 0 \\ \hline RN & I \end{array}\right)$
2. *Quelque soit l'état initial X_0, l'absorption est presque sûre. Autrement dit, la probabilité qu'il existe un temps τ tel que $X_{\tau} \in \mathcal{A}$ est égale à 1.*

[5] $\rho(M)$, le rayon spectral, est le plus grand module des valeurs propres de M (cf ;[2].VII.11 et 12).

3. $\sum_{i \notin \mathcal{A}} N_{i,j}$ *est le temps d'absorption, autrement dit, le temps moyen d'attente pour arriver à un état absorbant, sachant que* $X_0 = j$ *.*
4. *Pour* $j \notin \mathcal{A}, v \in \mathcal{A}$ *,* $(RN)_{i,j}$ *est la probabilité d'accéder au* $i^{\text{ème}}$ *état absorbant* $v = s - a + i$ *, sachant que* $X_0 = j$ *.*
5. *Pour* $i, j \notin \mathcal{A}$ *,* $N_{i,j}$ *est l'espérance du nombre d'apparitions de l'état i, sachant que* $X_0 = j$ *.*

DEMONSTRATION :

1.

Des propositions 11 et 13 il résulte :

$$\underset{\tau \to +\tau}{Lim}\, M^{\tau} = \underset{\tau \to +\tau}{Lim} \left(\begin{array}{c|c} Q^{\tau} & 0 \\ \hline R\left(\sum_{i=0}^{\tau-1} Q^{i}\right) & I \end{array} \right) = \left(\begin{array}{c|c} 0 & 0 \\ RN & I \end{array} \right).$$

2.

Soit la variable aléatoire : $Y = \min_{\tau \in \mathbb{N}} \left(\tau / X_{\tau} \in \mathcal{A} \right)$ représentant le premier passage dans un état absorbant. L'évènement $\left[Y > \tau \right]$ est la réunion des événements $\left[X_0 \notin \mathcal{A} \right], \left[X_1 \notin \mathcal{A} \right], \cdots, \left[X_{\tau} \notin \mathcal{A} \right]$.

La magie des états absorbants c'est que quand on y tombe, on n'en sort plus ! On a donc, pour $\tau' \leq \tau : X_{\tau} \notin \mathcal{A} \Rightarrow X_{\tau'} \notin \mathcal{A}$. Tous ces événements sont donc en fait inclus dans $\left[X_{\tau} \notin \mathcal{A} \right]$.

La probabilité que, partant d'un état non absorbant $j \notin \mathcal{A}$, la chaîne reste dans un état non absorbant jusqu'au temps τ est :

$$P\left[Y > \tau / X_0 = j \right] = P\left[X_{\tau} \notin \mathcal{A} / X_0 = j \right] = \sum_{i \notin \mathcal{A}} P\left[X_{\tau} = i / X_0 = j \right] = \sum_{i \notin \mathcal{A}} (M^{\tau})_{i,j} = \sum_{i \notin \mathcal{A}} (Q^{\tau})_{i,j}$$

La probabilité que la chaîne reste toujours dans un état non absorbant est, en tenant compte de la proposition 11 :

$$\underset{\tau \to +\infty}{Lim}\, P\left[X_{\tau} \notin \mathcal{A} / X_0 = j \right] = \sum_{i \notin \mathcal{A}} \underset{\tau \to +\infty}{Lim} (Q^{\tau})_{i,j} = 0$$

L'événement contraire $\left[\exists \tau \in \mathbb{N} : X_{\tau} \in \mathcal{A} \right]$ est de probabilité 1, autrement dit presque sûr.

3.

Revenons à la variable *Y* déjà définie. Son espérance, est le temps moyen d'attente d'un état absorbant. En supposant, comme précédemment que $X_0 = j$, selon le lemme 3 :

$$E(Y / X_0 = j) = \sum_{\tau=0}^{+\infty} P\left[Y > \tau / X_0 = j \right].$$

Donc :

$$E(Y / X_0 = j) = \sum_{\tau=0}^{+\infty} P\left[X_{\tau} \notin \mathcal{A} / X_0 = j \right] = \sum_{\tau=0}^{+\infty} \sum_{i \notin \mathcal{A}} P\left[X_{\tau} = i / X_0 = j \right]$$

$$= \sum_{\tau=0}^{+\infty} \sum_{i \notin \mathcal{A}} (Q^{\tau})_{i,j} = \sum_{i \notin \mathcal{A}} \sum_{\tau=0}^{+\infty} (Q^{\tau})_{i,j} = \sum_{i \notin \mathcal{A}} N_{i,j}$$

4.

Soit B la matrice de taille $a\times(s-a)$ telle que $B_{i,j}$ désigne la probabilité que partant de l'état non absorbant $X_0=j$, il existe un temps $\tau\in\mathbb{N}$ tel que X_τ soit le $i^{\text{ème}}$ état absorbant (donc d'indice $v=s-a+i$).

Partant de $X_0=j$, pour atterrir au bout d'un certain temps dans l'état absorbant v, il faut :

- ✓ Soit que $X_1=v$ avec une probabilité de $P\left[X_1=v/X_0=j\right]=M_{v,j}=R_{i,j}$.
- ✓ Soit que $X_1=j'$ où $j'\notin\mathcal{A}$ avec une probabilité de $Q_{j',j}$ et que partant de j' on atterrisse sur v au bout d'un certain temps avec une probabilité de $B_{i,j'}$. Par indépendance on a donc :

$$B_{i,j}=R_{i,j}+\sum_{j'=1}^{s-a}B_{i,j'}\,Q_{j',j}=R_{i,j}+(BQ)_{i,j}$$

Matriciellement on a donc :

$$B=R+BQ\Leftrightarrow B(I-Q)=R\Leftrightarrow B=R(I-Q)^{-1}\Leftrightarrow B=RN$$

5.

Soit $i,j\notin\mathcal{A}$, $\mathrm{T}\in\mathbb{N}$. La variable aléatoire $I_\mathrm{T}=\sum_{\tau=0}^{\tau=\mathrm{T}}1_{X_\tau=i}$ représente le nombre de passage par l'état i dans l'intervalle de temps $\left[0\cdots\mathrm{T}\right]$. Son espérance sachant que $X_0=j$ est :

$$E\left(I_\mathrm{T}/X_0=j\right)=E\left(\sum_{\tau=0}^{\tau=\mathrm{T}}1_{X_\tau=i}/X_0=j\right)=\sum_{\tau=0}^{\tau=\mathrm{T}}E\left(1_{X_\tau=i}/X_0=j\right)$$

La variable aléatoire $1_{X_\tau=i}=\begin{cases}1 \text{ si } X_\tau=i\\ 0 \text{ sinon}\end{cases}$ a pour espérance sachant que $X_0=j$:

$$E\left(1_{X_\tau=i}/X_0=j\right)=P\left[X_\tau=i/X_0=j\right]=(M^\tau)_{i,j}=(Q^\tau)_{i,j}$$

Donc :

$$E\left(I_\mathrm{T}/X_0=j\right)=\sum_{\tau=0}^{\tau=\mathrm{T}}E\left(1_{X_\tau=i}/X_0=j\right)=\sum_{\tau=0}^{\tau=\mathrm{T}}(Q^\tau)_{i,j}$$

Lorsque $\mathrm{T}\to+\infty$, selon la proposition 13, $\sum_{\tau=0}^{\tau=\mathrm{T}}(Q^\tau)_{i,j}\to N$. D'où :

$$\underset{\mathrm{T}\to+\infty}{Lim}\,E\left(I_\mathrm{T}/X_0=j\right)=N_{i,j}$$

Application au jeu de mots

Ce sont les points 2, 3. et 4. du théorème 14 qui sont particulièrement importants pour nos jeux de mot pour lesquels on part toujours de l'état $\varnothing$ (d'indice 1).

2. Confirme que le jeu se termine presque surement par l'apparition du mot d'Alice ou du mot de Bob qui constituent les états absorbants de la chaîne.
3. Dans le cas d'un mot unique, $\sum_{i \notin \mathcal{A}} N_{i,1}$ la somme de la première colonne de la matrice fondamentale, est le temps moyen d'attente du mot dont on sait qu'il est égal à $k \times K_{A,A}(k)$.
4. Dans le cas de 2 mots A, B $(RN)_{1,1}$ est la probabilité d'arriver au premier état absorbant, c'est-à-dire au mot A (victoire d'Alice) et $(RN)_{2,1}$, probabilité d'arriver au deuxième état absorbant, c'est-à-dire au mot B (victoire de Bob). Probabilités données par la formule de Conway.

En retournant vers des exemples déjà traités par les polynômes de corrélation, on doit retrouver les mêmes résultats.

Sur les exemples déjà cités

Exemple 1, A = FPFFP (voir tableau page 9) :

Il n'y a qu'un état absorbant A.

La matrice de transition sous forme canonique

$$M_1 = \left(\begin{array}{ccccc|c} 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1/2 & 1 \end{array}\right)$$

La matrice fondamentale

$$N_1 = \begin{pmatrix} 10 & 8 & 8 & 6 & 4 \\ 12 & 12 & 10 & 8 & 6 \\ 8 & 8 & 8 & 6 & 4 \\ 4 & 4 & 4 & 4 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix}$$

Partant de l'état $\varnothing$, le temps moyen pour atteindre l'unique état absorbant A = FPFFP est la somme de la première colonne de N_1, soit : 36, ce qui est précisément la valeur donnée par la formule $2K_{A,A}(2)$.

Exemple 2 : A = PPP, B = FPP (voir page 12)

Il y a 2 états absorbant A et B.

La matrice de transition sous forme canonique

$$M_2 = \left(\begin{array}{ccccc|cc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1 \end{array}\right)$$

La matrice fondamentale

$$N_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 & 0 \\ 7/2 & 3 & 4 & 2 & 2 \\ 1/4 & 1/2 & 0 & 1 & 0 \\ 7/4 & 3/2 & 2 & 1 & 2 \end{pmatrix}$$

Partant de l'état $\varnothing$, le temps moyen pour atteindre un état absorbant (soit A ou B) est donc la somme de la première colonne de N_2 soit :7.

De plus :

$$R_2 N_2 = \begin{pmatrix} 1/8 & 1/4 & 0 & 1/2 & 0 \\ 7/8 & 3/4 & 1 & 1/2 & 1 \end{pmatrix}$$

Donc, partant de l'état $\varnothing$, la probabilités d'arriver à l'état PPP (victoire d'Alice) est de 1/8 contre 7/8 la probabilité d'arriver à l'état FPP (victoire de Bob). Ce que confirme la formule de Conway.

En général

Cette approche matricielle est nettement plus exigeante quant aux pré-requis mathématiques qu'elle suppose. Son intérêt réside dans son champ d'application plus étendu :

- ✓ Elle s'applique avec un alphabet Ω de taille quelconque.

C'était déjà le cas pour la méthode fondée sur les polynômes de corrélation $K_{\bullet,\bullet}(X)$. Mais de plus :

- ✓ Elle s'applique à un nombre de mots en compétition >2.
- ✓ Elle s'applique avec des probabilités d'émission différente pour les lettres de l'alphabet[6].

Données

Alphabet : $L=[l_1,\cdots,l_k]$

Probabilités : $\mathrm{P}:=[p_1,\cdots,p_k]$, p_i désigne la probabilité d'émission de la lettre l_i.

Mots en compétition $[A_1, A_2,\cdots, A_m]$ liste de m mots distincts de même longueur n. $A_j = a_j^1 a_j^2 \cdots a_j^n$

Construction de l'espace des états S

S est la réunion des préfixes des A_j. Les éléments de cet ensemble sont indexés de 1 à $nbs=|S|$, suivant la longueur des mots et suivant l'indice du mot sous-jacent pour former la liste des états.

$S=\left[\varnothing, a_1^1,\cdots,a_1^1 a_1^2,\cdots,A_1,\cdots A_m\right]$

L'essentiel, en vue d'obtenir une matrice de transition sous forme canonique, est que les états finaux $A_1, A_2,\cdots, A_m$ soient en queue de liste.

Ces états finaux sont par définition absorbants puisque $P\left[X_{\tau+1}=A_i / X_\tau = A_i\right]=1$.

[6] Dans [11] et [12], on pourra trouver une définition des polynômes de corrélation étendue pour traiter ces deux dernières situations.

Construction du graphe de transition $G=(S,F)$

La liste S des états étant définie, les flèches du graphe sont déterminées par l'algorithme informel suivant :

```
Pour chaque état non absorbant s Faire
   Pour chaque lettre l∈L Faire
   Si s&l∈S Alors
        ajouter la flèche s --l--> s&l à F
   Sinon
        déterminer le plus long suffixe su de s&l∈S
        ajouter la flèche s --l--> su à F
        Fin Si
        Fin faire
Fin Faire
Pour chaque état absorbant s Faire
        ajouter la flèche s --1--> s à F
Fin Faire
```

Remarques :

1. & désigne l'opérateur de concaténation des mots.
2. On peut tout aussi bien, remplacer l'étiquette l de la flèche par la probabilité d'émission de la lettre l. C'est d'ailleurs ce qui sera fait pour construire la matrice de transition.
3. Quel que soit l'état, il existe un chemin vers un des états absorbants. En effet chaque état étant un préfixe d'un des mots A_j de la liste, il existe un chemin de cet état vers A_j.
4. En terme de complexité d'exécution, le plus pénalisant est la boucle qui visite les états non absorbants qui imbrique en fait deux autres boucles : la boucle qui recherche le plus long suffixe dont le nombre d'itérations est borné par n, la boucle qui parcours l'alphabet qui est exécutée k fois. Avec m mots de n lettres, le nombre d'états non absorbant est borné par $m \times n$. Tout ceci mène à une complexité au pire de $m \times n \times k \times n = mkn^2$.

Implémentation sous forme de procédures Maple

La procédure `etats(Mots)` retourne la liste des états (préfixes des mots de la liste) sans répétition. Elle est ordonnée suivant la longueur des préfixes, puis selon l'ordre des mots de la liste. Elle suppose mais ne vérifie pas que tous les mots de la liste ont même longueur.

```
etats:=proc(Mots::list)
local i,j,nbmots,n::integer,pref::string,S::sequence;
S:="";
nbmots:=nops(Mots);
n:=length(Mots[1]);
for i from 1 to n do
    for j from 1 to nbmots do
    pref:=substring(Mots[j],1..i);
    if not member(pref,[S]) then S:=S,pref; fi;
    od;
od;
RETURN([S]);
end:
```

Exemple :

```
etats(["FFF","FFP","PFP"]);
```

["", "F", "P", "FF", "PF", "FFF", "FFP", "PFP"]

Dans la procédure de détermination de la matrice de transition, mots et lettres seront manipulés par leurs indices. Elle nécessite le recours à une procédure auxiliaire `rang(L,element)` pour déterminer le rang d'un élément dans une liste (retourne -1 si absent) :

```
rang:=proc(L::list,element)
local i::integer;
i:=1;
while i<=nops(L) and L[i]<>element do i:=i+1;od;
if i>nops(L) then i:=-1;fi;
RETURN(i);
end:
```

Enfin, la procédure `mattransitions(Mots, Alphabet,P)` retourne la matrice de transition sous forme canonique. Préalablement, elle affiche la liste des états. Cette procédure a recours à une fonction du « package » `StringTools`. En vue des calculs matriciels ultérieurs comme le calcul de la matrice fondamental, il est aussi utile de charger le « package » `linalg`.

```
mattransitions:=proc(Mots,Alphabet,P::list)
local i,r,
k,s,l,v,nbmots,n,nbs::integer,S::list,M::matrix,su::string;
S:=etats(Mots);
print(S);
nbs:=nops(S);
nbmots:=nops(Mots);
k:=nops(Alphabet);
M:=matrix(nbs,nbs,(i,j)->0);
# états non absorbants
for s from 1 to nbs-nbmots do
   for l from 1 to k do
   su:=cat(S[s],Alphabet[l]);
   v:=rang(S,suivant);
   if v>0 then M[v,s]:=P[l];
   else
   r:=1;
   while rang(S,Drop(su,r))<0 do r:=r+1;od;
   v:=rang(S,Drop(su,r));
   M[v,s]:=P[l];
   fi;
   od;
od;
# états absorbants
for i from nbs-nbmots+1 to nbs do M[i,i]:=1;od;
RETURN(evalm(M));
end:
```

```
M1:=mattransitions(["FPFFP"],["P","F"],[1/2,1/2]);
```

["", "F", "FP", "FPF", "FPFF", "FPFFP"]

$$M1 := \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Si on reprend les exemples 1 et 2 en pipant la pièce avec une probabilité d'émission de $\frac{3}{5}$ pour P contre $\frac{2}{5}$ pour F, voici les résultats :

Exemple 1 bis

Pour la matrice de transition :

```
M1:=mattransitions(["FPFFP"],["P","F"],[3/5,2/5]);
```

["", "F", "FP", "FPF", "FPFF", "FPFFP"]

$$M1 := \begin{bmatrix} \frac{3}{5} & 0 & \frac{3}{5} & 0 & 0 & 0 \\ \frac{2}{5} & \frac{2}{5} & 0 & 0 & \frac{2}{5} & 0 \\ 0 & \frac{3}{5} & 0 & \frac{3}{5} & 0 & 0 \\ 0 & 0 & \frac{2}{5} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{5} & 1 \end{bmatrix}$$

Puis pour la matrice fondamentale :

```
N1:=inverse(diag(1,1,1,1,1)-submatrix(M1,1..5,1..5));
```

$$N1 := \begin{bmatrix} \frac{145}{8} & \frac{125}{8} & \frac{125}{8} & \frac{95}{8} & \frac{25}{4} \\ \frac{475}{36} & \frac{475}{36} & \frac{415}{36} & \frac{325}{36} & \frac{95}{18} \\ \frac{125}{12} & \frac{125}{12} & \frac{125}{12} & \frac{95}{12} & \frac{25}{6} \\ \frac{25}{6} & \frac{25}{6} & \frac{25}{6} & \frac{25}{6} & \frac{5}{3} \\ \frac{5}{3} & \frac{5}{3} & \frac{5}{3} & \frac{5}{3} & \frac{5}{3} \end{bmatrix}$$

Le temps moyen d'attente du mot `FPFFP` en partant de l'état $\varnothing$ est la somme des la première colonne de la matrice soit $\dfrac{3425}{72} \approx 47.57$ donc nettement rallongé par rapport à la pièce équitable.

Exemple 2 bis

```
M2:=mattransitions(["PPP","FPP"],["P","F"],[3/5,2/5]);
```

$$M2 := \begin{array}{c} [\text{""}, \text{"P"}, \text{"F"}, \text{"PP"}, \text{"FP"}, \text{"PPP"}, \text{"FPP"}] \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{3}{5} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & \frac{2}{5} & 0 & 0 \\ 0 & \frac{3}{5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{5} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{5} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{5} & 0 & 1 \end{bmatrix} \end{array}$$

```
N2:=inverse(diag(1,1,1,1,1)-submatrix(M2,1..5,1..5));
```

$$N2 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{3}{5} & 1 & 0 & 0 & 0 \\ \frac{98}{45} & \frac{16}{9} & \frac{25}{9} & \frac{10}{9} & \frac{10}{9} \\ \frac{9}{25} & \frac{3}{5} & 0 & 1 & 0 \\ \frac{98}{75} & \frac{16}{15} & \frac{5}{3} & \frac{2}{3} & \frac{5}{3} \end{bmatrix}$$

```
R2:=submatrix(M2,6..7,1..5);
```

$$R2 := \begin{bmatrix} 0 & 0 & 0 & \frac{3}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{5} \end{bmatrix}$$

```
multiply(R2,N2);
```

$$\begin{bmatrix} \frac{27}{125} & \frac{9}{25} & 0 & \frac{3}{5} & 0 \\ \frac{98}{125} & \frac{16}{25} & 1 & \frac{2}{5} & 1 \end{bmatrix}$$

Donc, partant de l'état $\varnothing$, la probabilités d'arriver à l'état `PPP` (victoire d'Alice) est de $\dfrac{27}{125} = 0.216$ contre $\dfrac{98}{125} = 0{,}7847$ la probabilité d'arriver à l'état `FPP` (victoire de Bob). Les chances d'Alice s'améliorent !

Exemple 3

Et pour terminer, un examen de 3 mots sur l'alphabet $[A,C,G,T]$ avec des probabilités inégales. Pour alléger l'écriture, les matrices sont affichées en décimales.

```
M3:=mattransitions(["TCA","CCT","AGC"],["A","C","G","T"],[3/10,2/10,2/10,3/10])
:
```

["", "T", "C", "A", "TC", "CC", "AG", "TCA", "CCT", "AGC"]

```
evalf(%,2);
```

$$\begin{bmatrix} 0.20 & 0.20 & 0.20 & 0. & 0.20 & 0.20 & 0.20 & 0. & 0. & 0. \\ 0.30 & 0.30 & 0.30 & 0.30 & 0.30 & 0. & 0.30 & 0. & 0. & 0. \\ 0.20 & 0. & 0. & 0.20 & 0. & 0. & 0. & 0. & 0. & 0. \\ 0.30 & 0.30 & 0.30 & 0.30 & 0. & 0.30 & 0.30 & 0. & 0. & 0. \\ 0. & 0.20 & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0.20 & 0. & 0.20 & 0.20 & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0.20 & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0.30 & 0. & 0. & 1. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0.30 & 0. & 0. & 1. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 0.20 & 0. & 0. & 1. \end{bmatrix}$$

```
N3:=inverse(diag(1,1,1,1,1,1,1)-submatrix(M3,1..7,1..7)): evalf(%,2);
```

$$\begin{bmatrix} 4.6 & 3.4 & 3.4 & 3.3 & 2.4 & 2.4 & 2.9 \\ 7.2 & 7.8 & 6.8 & 7.0 & 4.7 & 4.4 & 5.9 \\ 2.3 & 2.0 & 3.0 & 2.2 & 1.3 & 1.4 & 1.7 \\ 7.1 & 6.6 & 6.7 & 7.8 & 4.3 & 4.7 & 5.7 \\ 1.4 & 1.6 & 1.4 & 1.4 & 1.9 & 0.88 & 1.2 \\ 0.94 & 0.89 & 1.1 & 0.90 & 0.82 & 1.8 & 0.73 \\ 1.4 & 1.3 & 1.3 & 1.6 & 0.87 & 0.94 & 2.1 \end{bmatrix}$$

```
R3:=submatrix(M3,8..10,1..7): evalf(%,2);
```

$$\begin{bmatrix} 0. & 0. & 0. & 0. & 0.30 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0.30 & 0. \\ 0. & 0. & 0. & 0. & 0. & 0. & 0.20 \end{bmatrix}$$

```
multiply(R3,N3):evalf(%,3);
```

$$\begin{bmatrix} 0.434 & 0.469 & 0.405 & 0.417 & 0.580 & 0.265 & 0.352 \\ 0.283 & 0.267 & 0.327 & 0.270 & 0.246 & 0.547 & 0.218 \\ 0.283 & 0.264 & 0.267 & 0.312 & 0.173 & 0.188 & 0.430 \end{bmatrix}$$

Donc, partant de l'état $\varnothing$: $P[\mathtt{TCA}] \approx 43{,}4\%$, $P[\mathtt{TCA}] \approx 28{,}3\%$, $P[\mathtt{AGC}] \approx 28{,}3\%$.

Annexe 1 : des fonctions VBA pour Excel

Quatre fonctions VBA pour Excel permettent les calculs. Elles sont incluses dans le classeur : `fonctionsVBA_Jeu_de_mots.xls` dont la première feuille expose le mode d'emploi :

Fonctions VBA pour Jeu de mots

keyb(A;B)

Arguments :

A ; B	2 mots sur l'alphabet

Retourne :

Les coefficients (1 ou 0) du polynôme de corrélation $K_{A,B}(X)$ sous la forme d'un mot sur l'alphabet {1,0}

Exemples :

A= PFPF		B= FPFF	
keyb(A;A)	keyb(A;B)	keyb(B;A)	keyb(B;B)
1010	0101	0000	1001
A= ACGT		B= AGCA	
keyb(A;A)	keyb(A;B)	keyb(B;A)	keyb(B;B)
1000	0000	0001	1001

odds(A;B;k)

Arguments :

A ; B	2 mots sur l'alphabet
k	la taille de l'alphabet, optionnel, par défaut 2

Retourne :

P[A gagne]/P[B gagne]

Exemples :

A= PFPF B= FPFF

Odds(A;B)
1,8

A= ACGT B= AGCA

Odds(A;B;4)
1

key(A;B;k)

Arguments :

A ; B	2 mots sur l'alphabet
k	la taille de l'alphabet, optionnel, par défaut 2

Retourne :

$K_{A,B}(k)$

Exemples :

A= PFPF		B= FPFF	
key(A;A)	key(A;B)	key(B;A)	key(B;B)
10	5	0	9
A= ACGT		B= AGCA	
key(A;A;4)	key(A;B;4)	key(B;A;4)	key(B;B;4)
64	0	1	65

proba(A;B;k)

Arguments :

A ; B	2 mots sur l'alphabet
k	la taille de l'alphabet, optionnel, par défaut 2

Retourne :

P[A gagne]

Exemples :

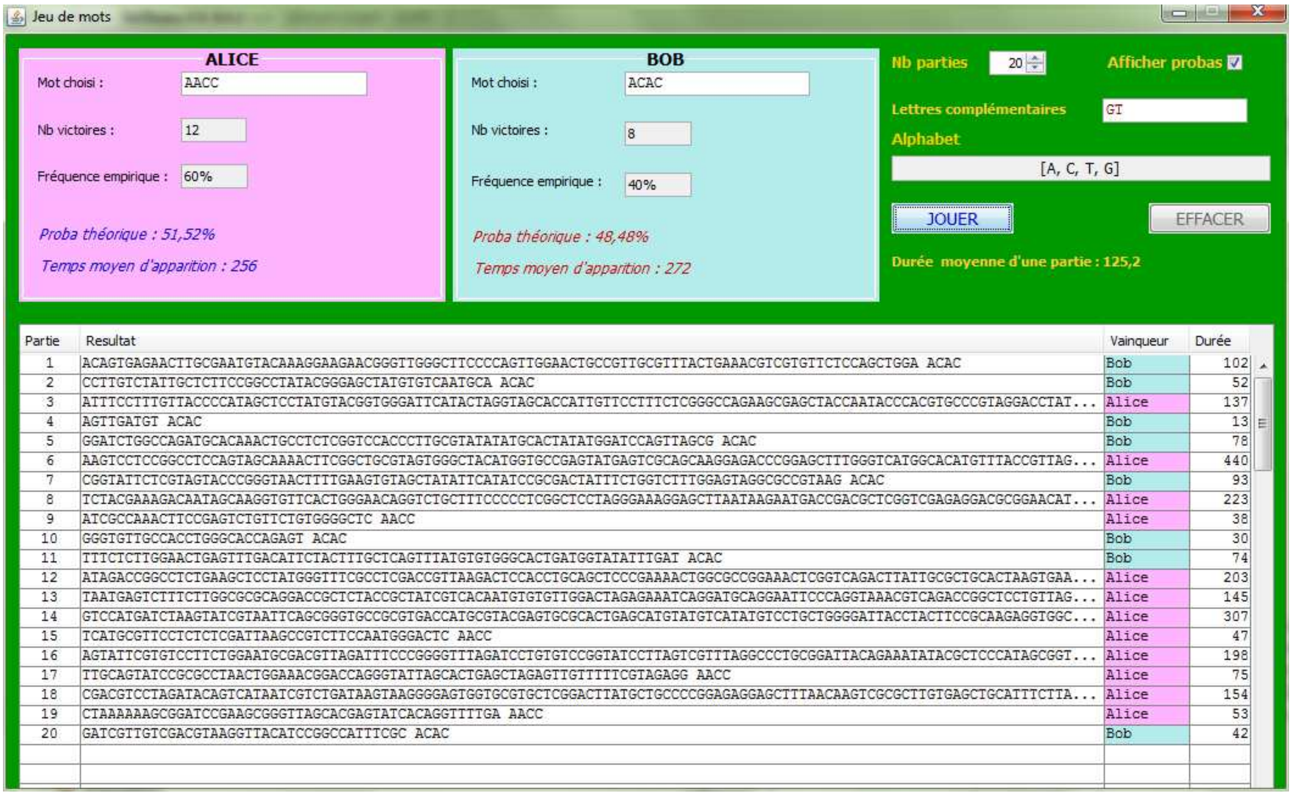
A= PFPF B= FPFF

proba(A;B)
64,29%

A= ACGT B= AGCA

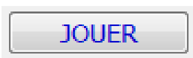
proba(A;B;4)
50,00%

Annexe 2 : une application Java pour simuler les joutes entre Alice et Bob



Fenêtre de l'application `JeuDeMots.jar`

Le mode d'emploi est très simple :

1. Alice écrit son mot dans le champ prévu à cet effet.
2. Bob fait de même.
3. Si l'alphabet inclut d'autres lettres que celles utilisées par Alice et Bob, les écrire à la suite, sans espaces, dans le champ « Lettres complémentaires ».
4. Choisir le nombre de parties entre 1 et 100.
5. Choisir ou non l'affichage des probabilités théoriques en plus des fréquences empiriques à la fin des parties.
6. Lancer la suite de parties par le bouton 

.

Cette application se place dans l'hypothèse d'équiprobabilité des lettres de l'alphabet.

Dans l'affichage du résultat de chaque partie, le mot gagnant est mis en évidence par l'espace qui le précède.

Bibliographie

[1] James BROFOS, A Markov Chain Analysis of a Pattern Matching Coin Game, *arXiv:1406.2212v1 [math.PR]* (2014).

https://arxiv.org/pdf/1406.2212v1

[2] Jacques CELLIER, *Algèbre linéaire, des bases aux applications*, PUR, Rennes, 2008.

[3] Jean-Paul DELAHAYE, Les surprises du jeu de pile ou face, *Pour la Science - n° 409 - Novembre 2011.*

http://www.lifl.fr/~jdelahay/pls/213.pdf

[4] Daniel FELIX, Optimal Penney Ante Strategy via Correlation Polynomial Identities, *Electronic Journal of Combinatorics* 13: 1–15 (2006).

http://www.combinatorics.org/ojs/index.php/eljc/article/download/v13i1r35/pdf

[5] Ronald L. GRAHAM, Donald E. KNUTH, Oren PATASHNIK, *Mathématiques Concrètes*, International Thompson Publishing, Paris, 1998.

[6] L. J. GUIBAS and A. M. ODLYZKO, String overlaps, pattern matching, and nontransitive games, *Journal of Combinatorial Theory* (A) 30 (1981): 183-208.

http://ac.els-cdn.com/0097316581900054/1-s2.0-0097316581900054-main.pdf

[7] R.S. NICKERSON, Penney Ante: Counterintuitive Probabilities in Coin Tossing (2007).

https://www.math.washington.edu/~mathcircle/circle/2015-16/second/PenneyAnte.pdf

[8] A. M. ODLYZKO, Enumeration of strings, *Combinatorial Algorithms on Words (*pp. 205-228), A. Apostolico and Z. Galil (eds.), Springer, 1985.

http://www.dtc.umn.edu/~odlyzko/doc/arch/string.enumerate.pdf

[9] Pavel A. PEVZNER, *Bio-informatique moléculaire*, Springer-Verlag France, Paris 2006.

[10]Yvan. VELENIK, *Probabilités et Statistique,* cours en ligne.

http://www.unige.ch/math/folks/velenik/cours.html

[11] Urszula OSTASZEWSKA, Krzysztof ZAJKOWSKI, On the waiting time till some patterns occur in i.i.d. sequences, *arXiv:1302.4859v1 [math.PR]* (2010).

https://arxiv.org/pdf/1302.4859v1

[12] Krzysztof ZAJKOWSKI, Penney's game between many players, *arXiv:1212.3973v3 [math.PR].* (2014).

https://arxiv.org/pdf/1212.3973v3