

1. Surprise pour le temps moyen d'attente

Quand on lance deux fois une pièce de monnaie, la probabilité d'obtenir FF est la même que la probabilité d'obtenir PF et vaut $1/4$. En lançant indéfiniment la pièce, la fréquence d'apparition de FF sera $1/4$, de même que pour PF. On est tenté d'en déduire que, quand on lance une pièce de monnaie indéfiniment, on attend en moyenne autant de temps pour obtenir FF que PF. C'est faux : le temps moyen d'attente de FF est de six lancers alors que le temps moyen d'attente de PF est de quatre lancers. Cela choque l'intuition et paraît même impossible !



Pour se persuader de l'exactitude de nos affirmations, on peut mener une vérification expérimentale, ou raisonner.

Exemple 1. Le temps d'attente moyen de la séquence P (une fois pile) est 2. En effet :

– Une fois sur deux, on obtient P dès le premier tirage (temps d'attente égal à 1). Dans la moyenne, on a donc 1 avec un poids $1/2$.

– Une fois sur quatre, on a le début FP, et le temps d'attente a donc été 2. Dans la moyenne, on a 2 avec un poids $1/4$.

– Une fois sur huit, on a le début FFP. Dans la moyenne, on a donc 3 avec un poids $1/8$; etc.

Le temps d'attente moyen de P est la somme de la série :

$1/2 + 2 \times 1/4 + 3 \times 1/8 + 4 \times 1/16 + \dots$ Cette série se calcule à partir de la série géométrique classique $1/2 + 1/4 + 1/8 + \dots = 1$.

$1/2 + 2 \times 1/4 + 3 \times 1/8 + 4 \times 1/16 + \dots =$

$1/2 + 1/4 + 1/8 + 1/16 + \dots$

$+ 1/4 + 1/8 + 1/16 + \dots$

$+ 1/8 + 1/16 + \dots$

$+ 1/16 + \dots$

La première ligne vaut 1, la deuxième $1/2$ (c'est la première divisée par 2), la troisième $1/4$ (c'est la première divisée par 4), etc. Le tout vaut ainsi $1 + 1/2 + 1/4 + 1/8 + \dots = 2$.

Exemple 2. Calculons maintenant le temps moyen d'attente T de la séquence FF. En envisageant les deux possibilités pour le premier tirage, on a :

$T = 1/2 \times (\text{temps moyen d'attente de FF si le premier tirage a donné F}) + 1/2 \times (\text{temps moyen d'attente de FF si le premier tirage a donné P})$.

Le temps moyen d'attente de FF quand le premier tirage a donné P est bien sûr une unité de plus que le temps moyen d'attente de FF, car quand on a obtenu P en premier, on a perdu un coup : le P obtenu ne sert à rien et ne servira pas. Le second terme de la somme est donc $1/2 \times (1 + T)$.

Pour le premier terme, on écrit que le temps moyen d'attente de FF si le premier tirage a donné F est égal à $1/2 \times (\text{temps moyen d'attente de FF si le premier tirage a donné F et le second a donné F}) + 1/2 \times (\text{temps moyen d'attente de FF si le premier tirage a donné F et le second P})$.

La première grande parenthèse vaut évidemment 2. Quant à la seconde, en raisonnant comme précédemment, on voit qu'elle vaut $(2 + T)$. En regroupant ce que nous venons d'obtenir, on a : $T = 1/2 \times [1/2 \times 2 + 1/2 \times (2 + T)] + 1/2 \times (1 + T)$, ce qui, après résolution, donne $T = 6$.

Exemple 3. Par la même méthode, en utilisant le résultat de l'exemple 1, on obtient que le temps d'attente moyen T' de PF vérifie :

$T' = 1/2 \times (\text{temps moyen d'attente de PF si le premier tirage a donné P}) + 1/2 \times (\text{temps moyen d'attente de PF si le premier tirage a donné F}) = 1/2 \times (1 + \text{temps moyen d'attente de F}) + 1/2 \times (1 + T') = 1/2 \times (1 + 2) + 1/2 \times T'$. On en déduit que $T' = 4$.

on oppose PPP et FFPP, car la seconde gagne avec une probabilité $7/12$.

(f) Des séquences peuvent avoir des chances égales de gagner dans une compétition à quatre (25 % de chance chacune) et pourtant, dans des compétitions deux à deux, être toujours inégales.

Ainsi, les séquences PPF, PFF, FPP et FFP gagnent chacune dans 25 % des cas quand on les fait concourir toutes les quatre ensemble, mais ne font jamais jeu égal quand on les oppose deux à deux.

(g) Il se peut que dans des compétitions deux à deux, on obtienne des cycles du type $[S_1 \text{ bat } S_2, S_2 \text{ bat } S_3, \dots, S_{n-1} \text{ bat } S_n, \text{ mais } S_n \text{ bat } S_1]$. Un tel cycle est obtenu avec les quatre séquences PPF, PFF, FFP et FPP (voir l'encadré 3) :

PPF bat PFF avec une probabilité $2/3$,
PFF bat FFP avec une probabilité $3/4$,
FFP bat FPP avec une probabilité $2/3$,
FPP bat PPF avec une probabilité $3/4$.

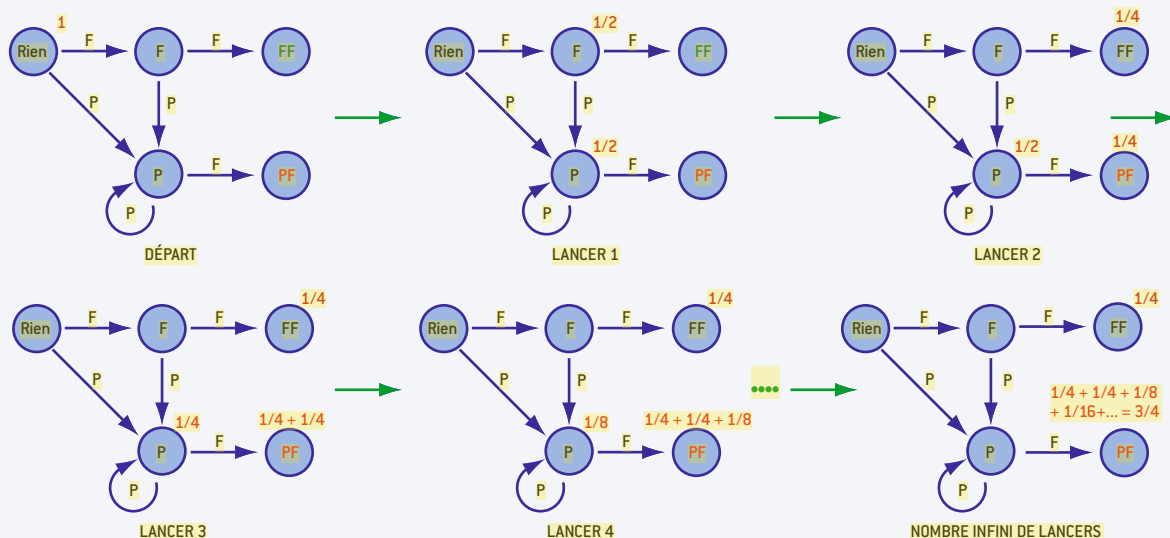
Corriger les fausses simplicités

Nous avons donc affaire à un paradoxe de non-transitivité : l'intuition nous souffle que chaque séquence possède une force propre qui lui permet de se présenter en tête devant une autre séquence et que, donc, si S_1 est plus forte que S_2 et S_2 plus forte que S_3 , alors nécessairement S_1 est plus forte que S_3 . C'est une illusion, car une séquence peut être « forte » devant une deuxième et « faible » devant une troisième. En sport

et dans les problèmes d'élection, de telles situations se produisent aussi.

Les encadrés expliquent comment mener les calculs qui vous permettront de vérifier ce qui vient d'être affirmé. L'examen attentif de ces méthodes affine notre compréhension et corrige nos jugements spontanés, dont l'inadéquation est flagrante. En relisant le raisonnement B présenté plus haut et en le comprenant, on finit par accepter que PF gagne contre FF et que, finalement... cela n'aurait pas dû nous étonner. De même, la méthode des graphes d'états (voir l'encadré 2) conduit à saisir clairement pourquoi PPF bat par exemple PFF. Plus généralement, elle permet de visualiser le rapport de force entre séquences et de comprendre (et calculer) que ce rapport se

2. La méthode des graphes d'états



Deux séquences de P et de F étant données, par exemple PF et FF, on lance une pièce de monnaie jusqu'à ce que l'une se présente. Ce sera la gagnante. Contrairement à ce que le bon sens suggère, la séquence PF gagnera plus souvent (trois fois sur quatre en moyenne) que la séquence FF.

Pour le prouver, un raisonnement direct est possible (voir le texte), mais la méthode des graphes d'états présentée ici se généralise à tout couple de séquences et même à des compétitions entre k séquences. C'est elle qui permet de déterminer (sans simulation) les probabilités de gains de ce type de compétitions.

On dessine un graphe dont les nœuds sont tous les débuts possibles de l'une des séquences en compétition. On relie le nœud N au nœud N' si N est le début de N'. Grâce à ce graphe, on calcule les probabilités des différentes situations, lancer après lancer, en opérant un suivi de la probabilité que chaque nouveau tirage fait avancer sur le graphe.

- Au départ, il n'y a aucun P ou F connu : avec une probabilité égale à 1, nous nous trouvons dans l'état « Rien ».

- Après un lancer, le 1 initial s'est séparé en deux fois 1/2, qui se trouvent sur les états P et F. Cela signifie qu'après un lancer, on a une chance sur deux d'être dans l'état P et autant pour l'état F.

- Après le second lancer, la probabilité 1/2 de l'état F s'est scindée en deux fois 1/4. Le premier 1/4 est allé se placer sur FF qui, main-

tenant, est marqué par 1/4. L'état P a donné la moitié de sa valeur à l'état PF (marqué maintenant par 1/4) et l'autre moitié de sa valeur à lui-même, ce qui, en additionnant avec le 1/4 qui vient de l'état F, donne une probabilité 1/2 pour PF. Ces marques correspondent aux probabilités qu'on a, après deux lancers, de se trouver dans les états notés sur le graphe.

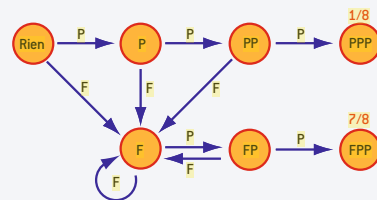
La circulation des probabilités marquées sur les nœuds du graphe se poursuit selon les mêmes règles : chaque valeur est coupée en deux et va vers les nœuds voisins en suivant les flèches. Cela donne ainsi, lancer après lancer, les probabilités (en rouge) de se trouver dans un état ou un autre. Bien sûr, quand une probabilité arrive sur l'une des séquences en compétition, elle n'en bouge plus. Bien sûr aussi, lors de ces calculs et quelle que soit l'étape, la somme des probabilités placées sur les nœuds marqués vaut 1.

Il est clair ici que le 1/4 de FF ne changera plus, et que les 3/4 restants vont peu à peu arriver sur PF qui, à l'infini, sera donc marqué par 3/4. L'état du graphe à l'infini indique les probabilités respectives de gain des séquences PF et FF : 1/4 et 3/4.

Appliquée au graphe ci-contre, la méthode permet de savoir très rapidement que PPP ne gagne qu'avec une probabilité de 1/8 contre FPP (seul 1/8 de la probabilité arrive sur PPP, et le reste finit donc sur FPP, car on continue à jouer tant que l'on n'a pas une des deux séquences). Dans le cas de graphes très compliqués, l'étude de la circulation des probabilités

d'un nœud à l'autre devient difficile à suivre. Cependant, elle se ramène à un problème classique d'algèbre linéaire qu'on sait parfaitement traiter, et donc la méthode des graphes d'états est générale pour connaître les probabilités respectives de gains quel que soit le nombre de séquences en compétition.

Il est intéressant de noter que la circulation des probabilités dans le graphe d'états fonctionne selon un principe identique à celui utilisé par le moteur de recherche Google pour attribuer des notes PageRank aux pages du Web, notes qui déterminent les rangs d'affichage des pages quand vous soumettez une requête. Au départ, une note égale est attribuée à chaque page ; à chaque étape de redistribution, cette note est fractionnée et passe aux pages citées (ce qui permet aux pages souvent citées d'avoir de bonnes notes). Quelques itérations – et non pas une infinité – de ce procédé de redistribution de PageRank conduisent à une valeur satisfaisante de la valeur de notoriété d'une page. Ce procédé itératif peut se paralléliser facilement et fonctionne donc même avec des milliards de pages à noter.





Jean-Paul DELAHAYE est professeur à l'Université de Lille et chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL).

construit à partir de la structure des deux séquences, et qu'il n'y a pas de force propre à une séquence indépendante des compétitions à laquelle on la soumet.

Connaître la fréquence de survenue d'une séquence S_1 et son temps moyen d'attente ne suffit pas pour prédire sa force quand elle sera confrontée une autre séquence S_2 , car cette force se décide face à S_2 et est spécifique de la compétition particulière S_1 contre S_2 , contrairement à ce que notre jugement premier nous faisait croire. L'explication de l'inadéquation de nos appréciations spontanées est sans doute à rechercher dans l'expérience que nous avons des nombreuses situations où la force d'un concurrent dans

une compétition provient d'une aptitude qu'il possède en propre (par exemple, en course à pied, son aptitude à aller vite) et qui n'est pas liée aux adversaires qu'il affronte.

Mener le calcul pour savoir qui gagne est un travail délicat par la méthode des graphes d'états. Peut-on arriver au résultat plus rapidement ? Un algorithme miraculeux de John Conway détermine le temps moyen d'attente d'une séquence et la probabilité de gagner d'une séquence par rapport à une autre (voir l'encadré 4).

On ne comprend pas pourquoi l'algorithme de Conway, d'une simplicité déroutante, donne la bonne réponse ... comme bien d'autres résultats de J. Conway. Ce mathématicien n'a jamais publié la preuve que son algorithme calcule convenablement (cela ne signifie pas qu'il n'en connaissait pas). Des démonstrations ont été publiées, dont aucune n'est simple. Nous disposons, grâce au génie de J. Conway, d'un procédé assez rapide pour savoir qui gagne dans une compétition à deux et les temps moyens d'attente. Toutefois, à ma connaissance, il n'existe pas d'algorithme équivalent pour savoir qui gagne dans une compétition entre trois séquences ou plus.

3. La non-transitivité de la relation de force

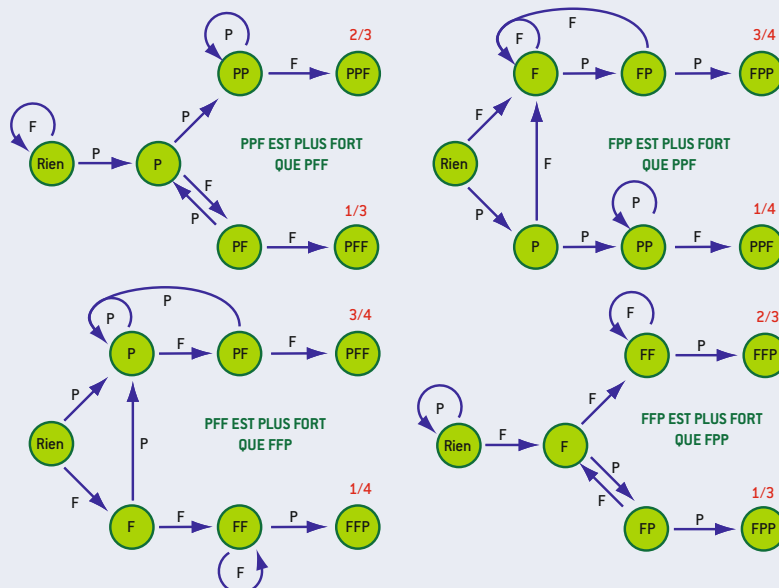
L'examen des quatre graphes d'états correspondant aux compétitions PPF-PFF, PFF-FFP, FFP-FPP et FPP-PFF indique qu'à chaque confrontation, la première séquence bat la seconde : nous avons donc un cycle paradoxal.

Le 2/3 en faveur de PPF contre PFF se justifie de la manière suivante en utilisant la méthode du graphe d'états. Au départ, on place 1 sur le nœud Rien. On suit ce que devient ce 1, lancer après lancer. Malgré la flèche du nœud Rien vers lui-même, toute la probabilité 1 se retrouve à la limite sur le nœud P. Celui-ci distribue 1/2 à PP qui finit par l'envoyer sur PPF. Le nœud P distribue aussi 1/2 à PF, qui envoie la moitié (c'est-

à-dire 1/4) à PFF. Le 1/4 que PF renvoie à P sera renvoyé pour moitié (1/8) à PP et pour moitié à PF, qui en enverra la moitié à PFF (1/16), etc. Le nœud PFF au total recevra donc $1/4 + 1/16 + 1/64 + \dots = 1/3$ (somme des termes d'une suite géométrique). Bien sûr tout le reste, c'est-à-dire 2/3, arrive sur PPF.

Le résultat 3/4 en faveur de PPF quand il est opposé à FFP est plus facile encore à justifier, puisque le nœud FF ne reçoit que 1/4 en tout.

Les deux autres graphes d'états sont identiques aux deux premiers à des changements de noms près, et donc se ramènent aux deux premiers. Le cycle est entièrement justifié.



Choisir à tour de rôle

Avec les suites de P ou F, on peut jouer ou parier. La méthode consistant à proposer la séquence perdante (par exemple FF) alors que vous vous réservez la bonne (PF) risque de susciter la méfiance. Mieux vaut laisser choisir une séquence, puis choisir ensuite une autre séquence de même longueur qui vous fera gagner.

Comment, face à une séquence S donnée, déterminer une séquence qui batte S ? Un examen systématique de toutes les séquences et la comparaison de chaque possibilité avec l'algorithme de Conway donnerait le résultat. De tête, le calcul serait compliqué, voire impossible, si le premier joueur a choisi une séquence de plus de dix piles ou faces. Apprendre par cœur le tableau des probabilités de gains (calculé à l'avance) n'est pas plus envisageable, même en se limitant aux séquences de taille inférieure à cinq. Comment s'y prendre ?