

2. Les ZD-stratégies

Une stratégie à mémoire d'un coup est définie par les probabilités p_1, p_2, p_3, p_4 de jouer c lorsque le dernier coup a été respectivement $[c, c], [c, t], [t, c]$ ou $[t, t]$. Notons $Strat(p_1, p_2, p_3, p_4)$ cette stratégie. W. Press et F. Dyson considèrent une classe de stratégies $Strat(p_1, p_2, p_3, p_4)$ dépendant de trois paramètres a, b et c , nommées ZD-stratégies et notées $ZD(a, b, c)$. Les équations générales reliant les paramètres p_1, p_2, p_3, p_4 pour les ZD dans le cas $\{R = 3, S = 0, T = 5, P = 1\}$ sont :

$$\begin{aligned} p_1 - 1 &= 3a + 3b + c; \\ p_2 - 1 &= 5b + c; \\ p_3 &= 5a + c; \\ p_4 &= a + b + c. \end{aligned}$$

W. Press et F. Dyson démontrent que

lorsqu'on oppose $ZD(a, b, c)$ à une stratégie $Strat$, et que l'on note G_1 le gain moyen par coup de la première et G_2 le gain moyen par coup de la seconde, alors ces gains moyens vérifient : $aG_1 + bG_2 + c = 0$. Le gain moyen G_2 ne dépend plus de ce que $Strat$ joue ! La stratégie $Strat$ est « contrôlée » par la ZD-stratégie.

Lorsque $a = 0$ et $b \neq 0$, on a : $p_1 = 3b + c + 1, p_2 = 5b + c + 1, p_3 = c, p_4 = b + c$ et $G_2 = -c/b$. Autrement dit, $Strat$ a un gain moyen indépendant des probabilités qui la définissent, gain qui ne dépend que de la ZD-stratégie qui lui fait face, $ZD(a, b, c)$. Une telle ZD-stratégie est nommée *Égaliseur*.

Prenons un exemple d'*Égaliseur* : $a = 0; b = -1/3; c = 2/3$;

$p_1 = 2/3; p_2 = 0; p_3 = 2/3; p_4 = 1/3$; $G_2 = -cb = 2$. Voici les résultats des rencontres entre cet *Égaliseur* et diverses stratégies classiques :

Égaliseur = 2,5 $\leftrightarrow$ *Donnant-donnant* = 2

Égaliseur = 3 $\leftrightarrow$ *Graduelle* = 2

Égaliseur = 3 $\leftrightarrow$ *Gentille* = 2

Égaliseur = 1 $\leftrightarrow$ *Méchante* = 2

Égaliseur = 2 $\leftrightarrow$ *Périodique-cct* = 2

Égaliseur = 1 $\leftrightarrow$ *Rancunière* = 2

Égaliseur = 2 $\leftrightarrow$ *Égaliseur* = 2

Égaliseur force le gain moyen de l'adversaire, mais cela se fait parfois à ses dépens : contre *Rancunière*, elle n'obtient qu'un point en moyenne par coup et comme *Égaliseur* force les stratégies rencontrées à avoir un faible score, elle en est victime lorsqu'elle joue contre elle-même !



William Press



Freeman Dyson

à un adversaire qui coopère. Pire encore, l'article présentait comme les meilleures possibles des stratégies simples (n'utilisant, pour se déterminer, que le dernier coup joué), ce qui est en contradiction avec l'idée d'un perfectionnement possible illimité des stratégies de plus en plus complexes.

L'étonnement était d'autant plus grand que les résultats avancés étaient des preuves mathématiques dans un domaine où il est difficile de démontrer des résultats, la raison étant que l'espace associé au problème est infini et discret. Il s'agit de l'espace de tous les algorithmes pouvant servir de stratégie, espace qui n'est muni d'aucune structure topologique en lien direct avec les performances des stratégies. Les méthodes habituelles fondées sur le continu ne donnent rien et la richesse énorme de l'espace des stratégies envisageables ne se laisse pas réduire à quelques équations ou considérations logiques. Les résultats de W. Press et F. Dyson remettaient en cause les innombrables conclusions admises et utilisées en théorie de l'évolution, en économie et en sciences sociales.

Le travail de W. Press et F. Dyson s'organise autour de deux arguments tirés de

théorèmes. Le premier théorème concerne le dilemme itéré dans une version limitée aux stratégies probabilistes à mémoire d'un coup : les stratégies *Lunatique*, *Donnant-donnant* et *Pavlov* appartiennent à cette catégorie, mais pas la stratégie *Graduelle* qui consulte tous les coups déjà joués.

Stratégies probabilistes à mémoire d'un coup

Une stratégie à mémoire d'un coup est définie par quatre paramètres p_1, p_2, p_3, p_4 indiquant la probabilité de jouer c lorsque le dernier coup a été $[c, c], [c, t], [t, c]$ ou $[t, t]$. Notons $Strat(p_1, p_2, p_3, p_4)$ cette stratégie générale (on supposera que la stratégie coopère au premier coup).

La stratégie *Donnant-donnant* est $Strat(1, 0, 1, 0)$: elle coopère avec une probabilité de 100 % si le dernier coup est $[c, c]$ ou $[t, c]$, et trahit sinon. De même, on vérifie sans peine que la stratégie *Lunatique* est $Strat(1/2, 1/2, 1/2, 1/2)$ et que *Pavlov* est $Strat(1, 0, 0, 1)$.

W. Press et F. Dyson considèrent une classe particulière des stratégies $Strat(p_1, p_2, p_3, p_4)$ dépendant de trois

paramètres a, b et c et dénommées ZD-stratégies. Nous les noterons $ZD(a, b, c)$ (voir la figure 2). Le ZD provient d'un déterminant nul (zero determinant) qui apparaît dans les raisonnements conduisant aux théorèmes.

Le plus important est que les deux chercheurs démontrent qu'elles ont la propriété d'avoir un gain moyen lié linéairement à celui de leur adversaire quand celui-ci est pris parmi les stratégies à mémoire d'un coup. Ainsi, quand on oppose $ZD(a, b, c)$ à une stratégie $Strat$ et qu'on note G_1 le gain moyen par coup de la première et G_2 le gain moyen par coup de la seconde, ces gains moyens vérifient $aG_1 + bG_2 + c = 0$.

Si $a = 0$ et $b \neq 0$, on a $G_2 = -c/b$: $Strat$ a un gain moyen indépendant des probabilités qui la définissent, gain qui ne dépend que de la ZD-stratégie qui lui fait face, $ZD(0, b, c)$! Une telle ZD-stratégie est nommée *Égaliseur*. Contre elle, toutes les stratégies à mémoire d'un coup obtiennent le même gain moyen connu d'avance : $-c/b$. Inutile de se débattre face à une stratégie *Égaliseur*, vous gagnerez $-c/b$ et pas plus ! Les valeurs possibles pour $-c/b$ sont les valeurs situées entre P et R (c'est-à-dire entre 1 et 3 dans le cas classique).

3. Les extorqueurs

Parmi les ZD-stratégies découvertes par W. Press et F. Dyson certaines opèrent une forme d'extorsion. En effet, si $c = -(a + b)P$, on démontre que le gain moyen G_1 de la stratégie ZD contre une autre (laquelle obtient le gain moyen de G_2) vérifie : $G_1 - P = X(G_2 - P)$ avec $X = -b/a$. En

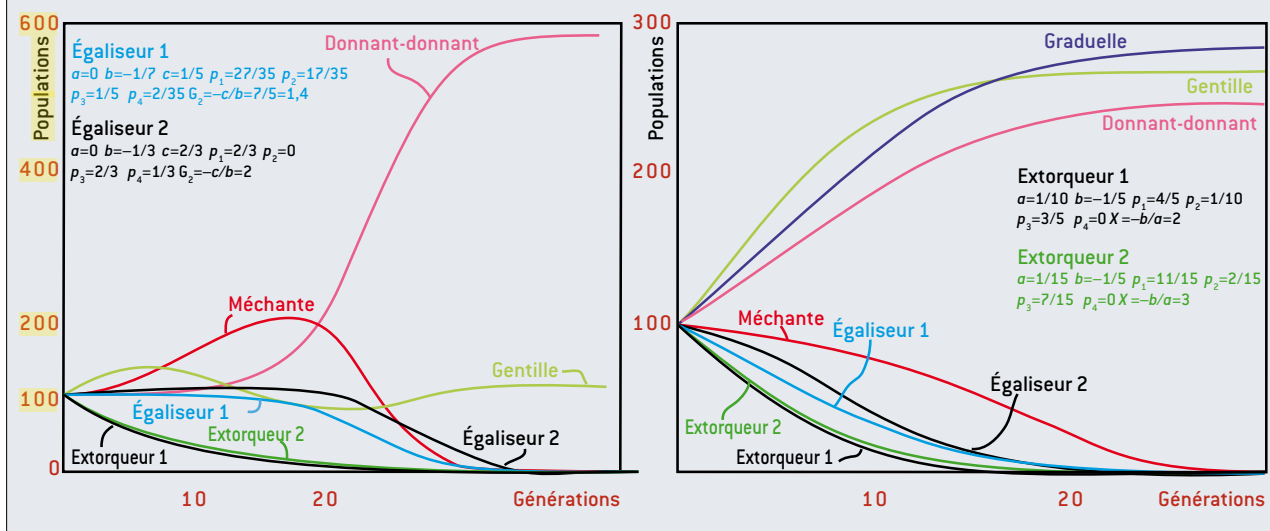
clair, si la seconde veut gagner plus, donc augmenter $(G_2 - P)$, cela entraîne mécaniquement que la stratégie ZD augmente son gain moyen, dont l'écart à P est toujours X fois l'écart à P du gain de la seconde.

Les quatre paramètres définissant ces stratégies *Extorqueurs* sont donnés

par les équations : $p_1 = 2a + 2b + 1$; $p_2 = 4b - a + 1$; $p_3 = 4a - b$; $p_4 = 0$.

Prenons deux stratégies *Extorqueurs* de cette catégorie, deux stratégies *Égaliseurs* dont les paramètres sont donnés ci-dessous et regardons leurs évolutions avec les stratégies classiques, dans un cas sans *Graduelle*

et dans l'autre cas avec. Un des gros défauts des stratégies *Extorqueurs* est qu'elles jouent mal contre elles-mêmes et contre n'importe quelle autre stratégie *Extorqueur*. Dans la simulation évolutive, les *Égaliseurs* et les *Extorqueurs* ne subsistent pas et *Graduelle* prend la première place.



La stratégie ZD[0, -1/3, 2/3] (voir la figure 2) qui est Strat[2/3, 0, 2/3, 1/3] force le gain moyen de son adversaire, qui sera de 2 points en moyenne par coup, quelle que soit la stratégie à mémoire d'un coup prise par l'adversaire. Il y a mieux encore : les ZD-stratégies dénommées *X-Extorqueur*.

Elles ne vous permettent d'avoir un bon résultat que si vous leur en accordez un proportionnellement meilleur. Ce sont celles où $c = -(a + b)P$. En jouant contre elle, la relation entre votre gain moyen G_2 et le sien G_1 est $G_1 - P = X(G_2 - P)$, où $X = -b/a$.

Cela signifie que ce qu'elles gagnent en moyenne en plus de P (pris égal à 1 dans le cas usuel) est proportionnel à ce que leur adversaire gagne en moyenne en plus de P . Si vous jouez contre un *X-Extorqueur*, vous ne pouvez améliorer votre score moyen qu'en améliorant le sien : il vous taxe en proportion X de vos revenus !

Considérons ZD[1/10, -1/5, 1/10] (*Extorqueur 1* de la figure 3) qui est la stratégie Strat[4/5, 1/10, 3/5, 0]. C'est

une stratégie 2-*Extorqueur* : elle double pour elle l'écart à P . Par exemple, quand elle joue contre *Pavlov*, ce dernier gagne en moyenne 1,62 [0,62 point en plus de 1] alors que l'*Extorqueur* obtient en moyenne 2,24 [1,24 point en plus de 1].

Pas plus de mémoire

Le second théorème important de l'article de W. Press et F. Dyson indique que si, dans une partie supposée infinie, la stratégie A est face à une stratégie B ayant une mémoire de k coups, il existe alors une stratégie A' qui obtient le même score moyen face à B et qui n'a qu'une mémoire de k coups. Pour obtenir un certain résultat face à une stratégie donnée, il n'est jamais nécessaire d'avoir plus de mémoire qu'elle.

La combinaison des deux résultats mathématiques de W. Press et F. Dyson affirme que face à une stratégie *Égaliseur* ou *Extorqueur*, ce ne sont pas seulement toutes les stratégies à mémoire d'un coup

qui se trouvent contraintes, mais toutes les stratégies à mémoire finie. On est tenté d'en conclure que : « les stratégies mémorisant plus que le dernier coup sont inutiles et que l'on dispose avec les ZD-stratégies de stratégies dominantes dans l'absolu pour le dilemme itéré du prisonnier ».

Certains ont d'ailleurs interprété ainsi les théorèmes démontrés, et le titre retenu pour leur article suggère que c'est le cas de W. Press et F. Dyson (bien que le contenu de leur article soit plus prudent). Pourtant, la double affirmation sur l'inutilité des stratégies à mémoire étendue et sur la dominance absolue des ZD-stratégies est totalement fautive, comme l'ont prouvé plusieurs articles parus en 2013 (voir la bibliographie). Voici pourquoi.

Considérons d'abord l'affirmation de dominance des ZD-stratégies. On sait depuis longtemps qu'au dilemme itéré du prisonnier, battre son adversaire [obtenir plus de points que lui] peut se faire aux dépens de celui qui gagne et qui aurait pu obtenir