

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

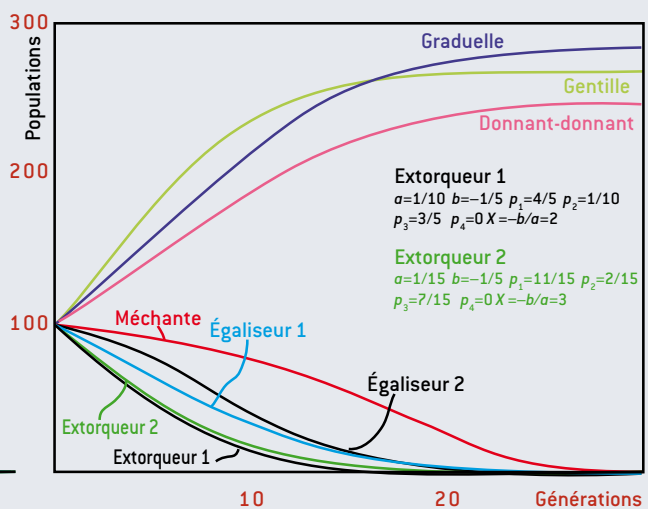
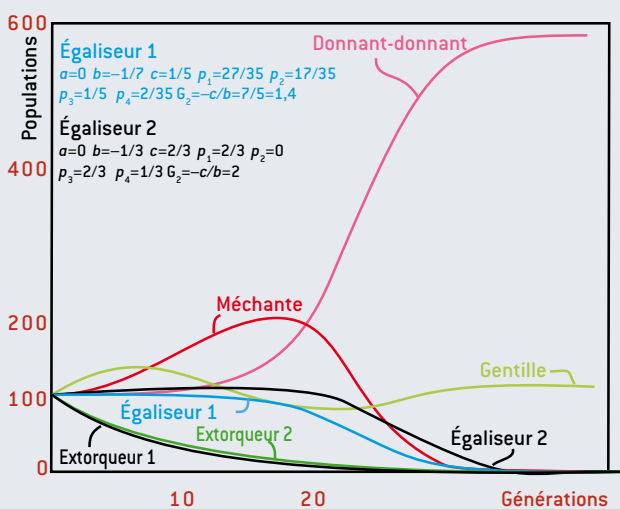
clair, si la seconde veut gagner plus, donc augmenter $(G_2 - P)$, cela entraîne mécaniquement que la stratégie ZD augmente son gain moyen, dont l'écart à P est toujours X fois l'écart à P du gain de la seconde.

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

plus de points en moyenne en acceptant d'être battu. La stratégie *Méchante* (qui joue toujours *t*) bat toute autre stratégie (c'est une évidence) ; par exemple, sur une partie de 100 coups contre *Donnant-donnant*, elle obtient 104 points, alors que *Donnant-donnant* en gagne 99. La stratégie *Gentille* (qui joue toujours *c*) ne bat pas *Donnant-donnant* et gagne 300 points en 100 coups contre *Donnant-donnant*, laquelle gagne aussi 300 points. Contre *Donnant-donnant*, *Méchante* gagne peut-être, mais elle a tort de gagner car en faisant comme *Gentille*, elle obtiendrait un bien meilleur score.

Mieux vaut s'entendre avec son adversaire

La plupart des ZD-stratégies sont un peu dans la même situation : elles ne battent leur adversaire qu'en renonçant elles-mêmes à de bons scores. Les *Extorqueurs* gagnent sur les stratégies auxquelles on les oppose, mais aux dépens du total des points gagnés. D'ailleurs, une stratégie *X-Extorqueur* (avec $X > 1$) qui joue contre elle-même n'obtient (d'après la théorie qu'on vérifie par simulation) qu'un seul point en moyenne par coup, ce qui est très médiocre. Il n'est pas vrai qu'être une bonne stratégie, c'est battre son adversaire. Mieux vaut ne pas le battre, bien s'entendre avec lui et marquer beaucoup de points.

Le cas de *Donnant-donnant* est remarquable et tel qu'on a l'impression d'un paradoxe quand on énonce ses propriétés : *Donnant-donnant* ne bat aucune stratégie individuellement et se fait battre par de nombreuses stratégies et pourtant, c'est une bonne stratégie qui gagne de nombreux tournois et compétitions évolutives. Elle gagne non pas parce qu'elle force les autres à gagner moins qu'elle, comme le fait une stratégie *Extorqueur*, mais parce qu'elle punit les stratégies qui ne veulent pas coopérer. Elle force la coopération. Face à elle, ou bien vous gagnerez peu de points, ou bien vous coopérerez, ce qui sera bon pour elle et pour vous.

C'est donc un contresens que croire que les ZD-stratégies sont bonnes. Dans des rencontres un contre un, elles dominent,

comme le fait *Méchante*, mais globalement elles jouent piteusement ! Outre l'erreur de croire que battre son adversaire, c'est gagner des points, un autre oubli a conduit à croire que les ZD-stratégies étaient dominantes : pour s'imposer, il faut jouer correctement contre soi-même. C'est important dans les tournois, mais plus encore dans les compétitions évolutives. En effet, si vous l'emportez lors des premières générations, l'arène se peuple de stratégies identiques à vous, et vous allez donc très fréquemment les rencontrer. Si vous jouez mal face à vous-même, cela se retourne contre vous.

Rien n'est faux dans les résultats mathématiques de W. Press et F. Dyson, mais en n'abordant que le problème « qui gagne dans un combat un contre un ? » et en oubliant les questions « Combien de points sont gagnés ? » et « Face à toi-même, te fais-tu du tort ? », les théorèmes démontrés ne permettent pas de conclure que les ZD-stratégies sont de bonnes stratégies.

Le résultat de W. Press et F. Dyson sur l'inutilité pour une stratégie d'avoir de la mémoire est juste dans le sens précis de ce qu'ils démontrent : si *A* est face à une stratégie *B* ayant une mémoire de *k* coups, il existe une stratégie *A'* qui obtient le même score moyen face à *B* et qui n'a qu'une mémoire de *k* coups. Mais cela ne signifie pas que face à deux stratégies *B* et *C* ayant une mémoire de *k* coups, il existe une stratégie *A'* à mémoire de *k* coups qui fasse le même score face à *B* et face à *C*. En effet, celle, *A'*, qui peut remplacer *A* face à *B* n'est pas nécessairement la même que celle, *A''*, qui peut remplacer *A* face à *C*. Le résultat de W. Press et F. Dyson sur l'inutilité de la mémoire est valable uniquement dans les combats un contre un ; il cesse d'être vrai dès qu'on envisage des tournois ou des évolutions.

Il est très utile que des chercheurs viennent explorer un domaine qui n'est pas le leur, et si W. Press et F. Dyson n'avaient pas été aussi confiants en leurs résultats, ils n'auraient pas proposé un titre provocateur à leur article – titre qui, après analyse, est faux – et ils n'auraient pas attiré l'attention sur ce qui, malgré tout, présente un véritable intérêt. *Se non è vero è bene trovato !* ■

L'AUTEUR



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

BIBLIOGRAPHIE

C. Adami et A. Hintze, Evolutionary instability of zero determinant strategies demonstrates winning is not everything, *Nature Comm.*, vol. 4, 2193, 2013 (en accès libre).

C. Hilbe et al., Evolution of extortion on iterated prisoner's dilemma games, *PNAS*, vol. 110, pp. 6913-6918, 2013.

• W. Press et F. Dyson, Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent, *PNAS*, vol. 109, pp. 10409-10413, 2012.

K. Sigmund, *The Calculus of Selfishness*, Princeton University Press, 2010.

B. Beauflis, Modèles et simulations informatiques des problèmes de coopération entre agents, Thèse de doctorat, Université de Lille, 2000.

M. Boerlijst et al., Equal pay for all prisoners, *American Mathematical Monthly*, vol. 104, pp. 303-307, 1997.

W. Poundstone, *Le dilemme du prisonnier*, Cassini, 1993.

B. Beauflis, J.-P. Delahaye et P. Mathieu, Our meeting with Gradual: A good strategy for the iterated prisoner's dilemma, *Proceedings of Artificial Life V*, MIT Press/Bradford Books, pp. 202-209, 1996.