

> individus emportent 6 points en jouant [c, c], alors qu'ils en gagnent 5 en jouant [c, t] et 2 en jouant [t, t]. L'intérêt collectif est de jouer [c, c], mais une analyse rationnelle égoïste conduit à [t, t], qui est collectivement le pire cas! Dans ce jeu à «somme non nulle», ce que certains gagnent n'est pas nécessairement pris à d'autres et donc bien s'entendre est payant.

Le dilemme est «itéré» quand la situation se présente de façon répétée aux deux individus. Jouer consiste alors à choisir une stratégie déterminant, chaque joueur étant informé du passé, comment jouer le coup suivant. Six stratégies sont définies dans l'encadré 2.

Nous décrirons cinq méthodes équitables et stables de classification d'un ensemble de stratégies. La durée (nombre de coups) des rencontres pour nos calculs est toujours 100, mais dès que le nombre de coups dépasse une dizaine, les résultats changent très peu.

Nous verrons que les méthodes aboutissent parfois à des classements différents. Il est difficile d'analyser les divergences entre méthodes, qui sont pourtant reproductibles, car si l'on refait les évolutions probabilistes, chaque méthode produit en pratique toujours le même classement.

Précisons qu'aucune de ces méthodes ne permet de donner un classement absolu entre toutes les stratégies envisageables: quand on change l'ensemble des compétiteurs, une stratégie bonne peut devenir mauvaise et inversement. Le problème posé n'est pas de construire un ordre absolu entre toutes les stratégies imaginables, mais de déterminer si, un ensemble de compétiteurs étant fixé, les divers systèmes de classifications entre ces compétiteurs divergent sensiblement, et si oui pourquoi.

Dans la suite, on se donne un ensemble de stratégies $S = \{s_1, s_2, \dots, s_n\}$ qu'on tente de classer par ordre de réussite quand elles jouent les unes avec les autres au dilemme du prisonnier.

Ce problème de la classification d'un ensemble de stratégies a été posé dès les premiers articles de Robert Axelrod et William Hamilton sur le sujet, en 1981. Dans son livre (*voir la bibliographie*), Robert Axelrod a lui-même largement pris la peine de comparer et analyser les différences entre les méthodes d'évaluation des gains que nous nommons M0 (nombres de parties gagnées), M1 (nombre de points gagnés), M2 (évolution déterministe).

Dans les recherches publiées sur le sujet, plusieurs autres méthodes ont été envisagées pour mener les classifications, mais le plus souvent les auteurs choisissent M1. Les méthodes M3 et M4 que nous décrivons sont des variantes probabilistes de M1 et M2; alors que nous pensions qu'elles devaient conduire à des classements identiques ou proches de ce que donnent M1 ou M2, nous avons découvert que les résultats sont plus compliqués et parfois étonnants.

M0: LE NOMBRE DE VICTOIRES

Notons M0 le tournoi des victoires. Son idée est simple: chaque paire s_i, s_j de stratégies est considérée, elles jouent une partie l'une contre l'autre et celle qui emporte le plus de points gagne. Le score de chaque stratégie est compté de la manière suivante: +1 pour chaque partie gagnée, -1 pour chaque partie perdue et 0 en cas d'égalité. Avec les six stratégies données en exemple à l'encadré 2, on obtient, en s'aidant du tableau de points gagnés: *Méchante* 4, *Méfiant* 3, *Donnant-donnant* -1, *Sondeur* -1, *Graduelle* -2, *Gentille* -3. On notera donc:

M0: Méc > Méf > Don = Son > Gra > Gen.

Le résultat n'est pas surprenant, car *Méchante* ne perd aucune partie: à chaque fois qu'elle joue, elle obtient au moins autant de points que son adversaire, quoi qu'il joue. Ce nombre de victoires est cependant une piètre mesure de la qualité des stratégies, car pour

3

CLASSEMENTS RESSEMBLANTS, MAIS DIFFÉRENTS

Nous avons recherché un cas amenant au désaccord le plus extrême.

Aux stratégies *Méfiant* et *Sondeur*, nous avons associé *Pttc*, la stratégie qui joue périodiquement t, t, c, et *Pcct*, la stratégie qui joue périodiquement c, c, t. Les quatre stratégies ont toute la propriété de prendre l'initiative de trahir et c'est sans doute ce qui explique le désordre résultant. Les classements obtenus divergent sensiblement, même si l'on ne prend pas en compte M0. Aucun classement n'est meilleur que les autres.

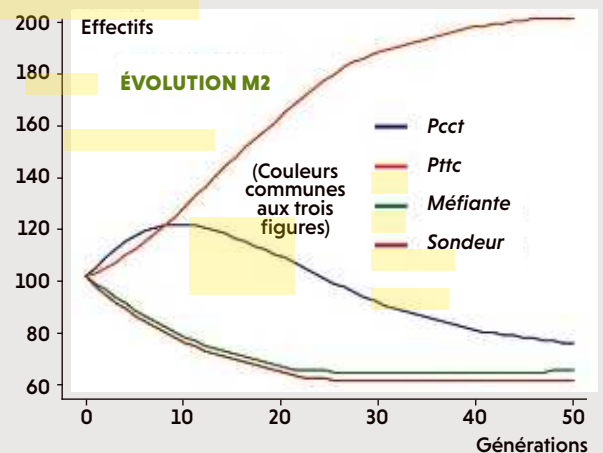
M0 : Pttc = Méc > Son > Pcct

M1 : Pcct > Pttc > Méc > Son

M2 : Pttc > Pcct > Méc > Son

M3 : Méc > Son > Pcct > Pttc

M4 : Méc > Son > Pttc > Pcct



gagner souvent, il faut accepter d'emporter peu de points. Cela apparaît avec M1.

M1: LE TOURNOI DES POINTS GAGNÉS

Si l'on évalue la valeur d'une stratégie dans l'ensemble S des six stratégies au nombre de points qu'elle emporte quand elle rencontre chaque stratégie de S (y compris elle-même), on obtient un nouveau classement (voir le tableau des points gagnés dans l'encadré 2) :

M1: Son > Gra > Don > Méf > Gen > Méc.

Méchante est passée de première à dernière!

La divergence avec le classement M0 fondé sur les victoires est étonnante. Cela s'explique assez bien. Battre son adversaire en étant intransigeant est facile, mais a pour conséquence qu'on gagne peu de points si l'adversaire ne se laisse pas faire. *Méchante* bat tout le monde en n'acceptant aucune coopération et cela la condamne à récolter peu de points. *Sondeur*, plus subtile, ne cherche à exploiter que celles qui se laissent faire comme *Gentille*, ce qui la propulse en tête.

Les classements M0 et M1 sont bien sûr équitables: la méthode de classification ne favorise *a priori* aucune stratégie. Ces classements sont stables car déterministes: ils ne dépendent d'aucun paramètre aléatoire. Notons que si l'on considère que l'important est de réussir à coopérer avec les individus rencontrés et non de les écraser, le classement par victoires (M0) n'est pas pertinent, alors que celui par nombre de points gagnés (M1) est bien plus sensé. Cependant, M1 n'est pas l'idée ultime, car cette méthode ne prend pas en compte l'idée d'évolution par sélection naturelle.

M2: COMPÉTITION ÉVOLUTIONNISTE DÉTERMINISTE

Les modèles évolutionnistes modélisent les compétitions entre stratégies en s'inspirant de ce qui se passe dans la nature quand des espèces

sont en compétition: celles qui réussissent à gagner beaucoup de points voient leurs effectifs augmenter, aux dépens de celles qui en gagnent peu, dont les effectifs diminuent.

La troisième méthode, M2, a été proposée par Robert Axelrod. L'évolution des effectifs d'une population de stratégies à la génération $n+1$ dépend de sa réussite en nombre de points gagnés dans la «soupe» de la génération n . Au départ, une soupe d'individus comprend P_1 individus utilisant la stratégie s_1 , P_2 individus utilisant la stratégie s_2 , etc. Cette soupe est la génération 1. Chaque individu rencontre tous les autres individus de la soupe. L'effectif des individus jouant s_1 dans la composition de la génération 2 est proportionnel aux points que gagnent les s_1 dans la soupe de la génération 1. De même pour s_2 , et ainsi de suite. Les bonnes stratégies de la génération 1 ont de plus grandes descendance. On s'arrange (par une règle de trois) pour que le nombre total d'individus des générations successives reste constant. La génération 3 est calculée à partir de la génération 2 de la même façon, et ainsi de suite.

Ainsi, les stratégies qui réussissent à coopérer et gagnent donc beaucoup de points prospèrent, et celles qui jouent mal, par exemple parce que trop agressives, déclinent. Cependant, si une stratégie s réussit bien parce qu'elle exploite des stratégies s' qui jouent mal et qui disparaissent progressivement, alors au bout de quelques générations, s , n'ayant plus personne à exploiter, va jouer mal dans les nouvelles soupes et son effectif va alors diminuer. Être bon à ce jeu, c'est savoir exploiter les mauvaises stratégies, mais surtout bien jouer avec les bonnes stratégies. Le modèle est naturel, équitable et il est stable car déterministe. C'est le plus simple des modèles évolutionnistes et le plus souvent utilisé pour classer les stratégies au dilemme du prisonnier. ➤

