

Robert AXELROD
DONNANT DONNANT
THEORIE DU COMPORTEMENT COOPERATIF
Traduit de l'américain par Michèle Garène
Odile JACOB, Paris, 1992 (1984)

Les jeux à somme nulle désignent des jeux dans lequel il y a un perdant et un gagnant pour chaque partie, soit au total $+1-1=0$. Les jeux à somme non nulle font que le résultat global est gagnant ou perdant pour les deux. Dans les systèmes humains, au niveau écologique, nous sommes dans des jeux à somme non nulle. D'où l'intérêt de réfléchir aux intérêts communs. Le changement climatique en est un exemple caricatural puisque les intérêts égoïstes et le bien commun vont dans des directions souvent divergentes...

Le dilemme du prisonnier¹, modélisation d'un jeu à somme non nulle, est une expérience de pensée qui m'a longtemps laissé dubitatif : deux suspects, interrogés séparément, sont placés devant les choix suivants : soit aucun de reconnaît sa culpabilité, et ils sont alors condamnés à 6 mois de prison chacun, soit les deux se dénoncent mutuellement et chacun fera 5 ans de prison, soit si l'un dénonce l'autre qui ne le dénonce pas, le dénonciateur est relaxé et le dénoncé est condamné à 10 ans de prison. Le calcul égoïste pousserait à la dénonciation, le calcul coopératif à ne rien avouer... la définition même du problème éliminant la qualité de la relation entre les deux partenaires du dilemme, autant jouer la réponse à pile ou face puisqu'on n'a droit qu'à une réponse...

J'avais par contre retenu de l'ouvrage de Theodor Caplow, *deux contre un*², que la meilleure stratégie était, dans une relation durable, de jouer un premier coup coopératif puis de s'aligner sur la réponse obtenue, qu'elle soit coopérative ou hostile. J'aurais dû aller à sa source, au livre de Robert Axelrod, ce que je ne fais que maintenant, c'est-à-dire bien tard... Le problème est posé très différemment dans la simulation d'un dilemme du prisonnier itératif, dans lequel on apprend en jouant et rejouant de nombreuses fois. Les conséquences n'étant plus que de gagner quelques points, et non plus des années de prison, les joueurs, en réalité des ordinateurs, s'inscrivent dans un processus à long terme et peuvent essayer diverses procédures pour tenter de gagner ensemble un maximum de points. Des programmes informatiques différents se confrontent, chacun se voulant plus malin que les autres.

Si les conditions initiales paraissent très simplifiées par rapport à la complexité des relations humaines, elles aboutissent à des leçons tout à fait intéressantes. Et il est possible de faire varier ces conditions initiales, de les enrichir de règles plus ou moins compliquées, ce qui ouvrent à des conclusions inattendues.

Retenons-en quelques-unes : si l'échange ne s'inscrit pas dans la durée, il y a tout intérêt à ne pas coopérer. Par contre, si la relation s'inscrit dans un temps long, avec le souvenir des coups précédents et l'anticipation des coups à venir, la coopération devient le meilleur choix, en se nuancant de qualités morales, humaines, ce qui n'est pas le moindre paradoxe puisque les programmes informatiques n'ont, *a priori*, aucun jugement moral. Mais, à nos yeux, ils en manifestent pourtant malgré eux. Ainsi on considéra qu'il y a de la « *bienveillance* » si un joueur n'est jamais le premier à faire défection (à trahir). Mais si ce joueur souhaite ne pas être dupé, il faudra aussi qu'il manifeste de la « *susceptibilité* », c'est-à-dire qu'il ne réponde pas par la coopération face à une défection, mais il devrait aussi faire preuve d'« *indulgence* » et ne pas s'enfermer dans un cercle infini de rétorsion réciproque pour revenir dans la coopération. De plus, contrairement à l'idée de jouer au plus malin, il vaut mieux faire preuve de « *transparence* », c'est-à-dire avoir une stratégie clairement lisible pour ses partenaires, amis ou ennemis. Court terme et long terme semblent les dimensions les plus influentes sur les choix des partenaires. De nombreux exemples, et de nombreuses variantes sont exposées et viennent enrichir nos réflexions stratégiques.

¹ Proposé par Albert W. Tucker en 1950.

² Caplow T. *Deux contre un, les coalitions dans les triades*. ESF, Paris, 1984