

À quoi sert la théorie des jeux ?

Philippe SOLAL

Comment des individus parviennent-ils à résoudre leurs conflits d'intérêt ? Quelles solutions émergent dans un cadre coopératif ? Et dans un contexte non-coopératif ? La science économique s'intéresse à ces questions grâce à l'approche dite de la théorie des jeux, dont l'émergence, entre les années 1920 et les années 1950, est retracée par Robert Leonard dans son dernier ouvrage.

Recensé : Robert Leonard, *Von Neumann, Morgenstern and the Creation of Game Theory. From Chess to Social Science, 1900-1960*. New York, Cambridge University Press, 2010.

La théorie des jeux : origine et développement

En tant que discipline académique, la théorie des jeux a pour objectif de formaliser des situations conflictuelles inhérentes à une communauté d'individus en interaction, de discuter puis de proposer des solutions à ces conflits. La conception des solutions est guidée par des critères d'optimalité individuelle ou collective, de cohérence temporelle, de justice distributive... Schématiquement, la théorie des jeux aborde de deux manières la résolution formelle d'un conflit.

Si les individus ont la possibilité de signer des accords de coopération qui les engagent de manière irrévocable dans un ensemble d'activités, alors une solution se ramène à allouer de la richesse, à imputer des coûts, à répartir des honneurs ou du pouvoir entre les signataires. L'interprétation donnée à une solution est que la collectivité se charge de dédommager les individus ayant coopéré malgré les conflits d'intérêt.

Si les individus évoluent dans un environnement dans lequel aucune institution n'est en mesure de garantir l'exécution d'un accord de coopération, l'accord prend le statut d'une promesse. Les individus ont donc toute licence, dans le cadre de règles préexistantes, pour déployer des stratégies prédatrices, de défense, de contrôle des stratégies des adversaires afin de maximiser la satisfaction de leur intérêt personnel au détriment de l'intérêt collectif. Dans ce contexte, une solution est un ensemble de stratégies, une par individu, qui représente une situation stable, un compromis : chaque individu est satisfait des conséquences de sa stratégie compte tenu des stratégies des autres membres de la communauté ou de l'idée qu'il s'en fait.

Ces deux façons de résoudre un conflit d'intérêt sont à l'origine de deux branches de la théorie des jeux : la branche coopérative et la branche non coopérative. La première de ces deux branches met l'accent sur les accords de coopération réalisables entre des coalitions d'individus, sur les conséquences de ces accords, et sur les procédures de répartition des richesses, des coûts, des honneurs ou des pouvoirs entre les membres d'une coalition. Les critères de justice distributive et d'optimalité collective sont au centre de la conception de ces

solutions. La branche non coopérative de la théorie des jeux étudie des environnements dans lesquels les parties d'un accord de coopération ne sont pas liées par leur engagement ; aucune compensation n'est garantie en cas de défection. En conséquence, l'analyse met l'accent sur la capacité des individus à recueillir de l'information dispersée, à former des anticipations sur le comportement des adversaires, et à ajuster leurs moyens à leurs fins.

En 1944, la publication de l'ouvrage *Theory of Games and Economic Behavior* par le mathématicien John von Neumann (1903-1957) et l'économiste Oskar Morgenstern (1902-1977) déclenche l'institutionnalisation de la théorie des jeux. Cet acte fondateur a structuré la recherche d'une première génération de théoriciens des jeux parmi lesquels on retiendra les noms de Robert Aumann, John Harsanyi, John Nash, Reinhard Selten, Lloyd Shapley, tous lauréats du prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel. La théorie des jeux a révolutionné la micro-économie néo-classique tout en impulsant de nouvelles directions de recherche dans d'autres disciplines scientifiques. Par exemple, en 1973 le théoricien de l'évolution biologique John Maynard Smith et le généticien George Price publient l'article « The Logic of Animal Conflict » dans la revue *Nature*. Ces auteurs considèrent une situation de conflit inter-espèces au sein de laquelle la sélection naturelle opère lentement un tri entre différents comportements. Cette sélection est stimulée par des mécanismes d'interactions bilatérales et de reproduction asexuée. Une interaction bilatérale est modélisée par un jeu non coopératif. Les individus n'ont pas le pouvoir de modifier leur comportement au cours du temps : ils sont programmés pour déployer la stratégie héritée de leur parent. L'utilité obtenue lors d'une interaction bilatérale mesure la capacité de cet individu à s'adapter à l'environnement.

Ainsi, plus une stratégie d'un individu s'avère efficace lors de différentes interactions bilatérales engagées avec chacun de ses opposants, plus ses chances de survie, et donc sa capacité à transmettre cette stratégie à la génération suivante, sont élevées. À partir de ce point de vue, une solution est définie par une collection de stratégies dont la caractéristique principale est de résister aux contraintes sélectives de l'évolution. Précisément, il est exigé, d'une part, que chacune de ces stratégies maximise le succès reproducteur de l'individu qui la déploie compte tenu de la distribution des stratégies dans la population, d'autre part, que cette stratégie résiste à l'introduction de rares éléments mutagènes dans la population. En somme, la stabilité des comportements inter-espèces est l'expression de la résolution d'une situation conflictuelle dans un environnement stratégique. Cet article est à l'origine de la branche évolutionnaire de la théorie des jeux qui a renouvelé la recherche en mathématiques de l'évolution du vivant et a stimulé la recherche en sciences sociales sur les phénomènes de mimétisme, de contagion des comportements, et plus généralement sur les apprentissages collectifs.

Le livre de Robert Leonard reconstitue avec brio en treize chapitres ce processus de formalisation, de pérennisation et d'acceptation de la théorie des jeux au sein de la communauté scientifique. Depuis le début des années 1990, l'auteur accumule et mobilise intelligemment une documentation colossale sur ce thème. La spécificité de sa démarche est d'enchâsser ce processus dans le contexte politique, social, culturel et scientifique des années 1920. La dissolution de l'Autriche-Hongrie en 1918, la montée en puissance du national-socialisme en Allemagne puis dans les pays de l'Europe centrale, l'influence de personnalités scientifiques telles David Hilbert et Karl Menger, la résistance de l'école autrichienne libérale de Friedrich Hayek, Hans Mayer et Ludwig von Mises, l'importance culturelle du jeu d'échecs dans les pays de l'Europe centrale, le rôle de la RAND Corporation dans l'intégration de la théorie des jeux au programme de défense nationale sont autant d'éléments

que l'auteur combine pour fournir une explication saisissante de l'émergence de la théorie des jeux. Face à la richesse du travail présenté par Robert Léonard, nous insisterons sur deux aspects. Le premier concerne l'interaction entre l'école autrichienne et les mathématiques. Le second souligne quelques lignes de force ouvertes par l'ouvrage de John von Neumann et Oskar Morgenstern.

L'école autrichienne et la théorie des jeux

Un pan de l'histoire institutionnelle de la théorie des jeux s'est déroulé à Vienne de la fin des années 1920 jusqu'à la fin des années 1930. Cette période correspond à une phase cruciale du développement de l'école autrichienne jusqu'à son déclin. Ce courant de pensée prend sa source à la fin du XIXe siècle dans les écrits de Carl Menger, Professeur à l'Université de Vienne, et père du mathématicien Karl Menger. Pendant les années 1930, l'école autrichienne a beaucoup influencé le discours sur les fluctuations économiques, la structure institutionnelle de l'économie capitaliste et la philosophie des sciences sociales. Friedrich Hayek, lauréat du prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel pour ses travaux sur le rôle de la monnaie dans les fluctuations économiques, est l'un de ses principaux représentants.

Robert Leonard fait revivre cette période intellectuelle foisonnante au cours de laquelle ont émergé les premiers travaux de John von Neumann en théorie des jeux, et les débats scientifiques entre économistes autrichiens et mathématiciens sur les fondements de l'économie théorique. Néanmoins, le climat politique délétère de cette période incitera nombre de personnalités scientifiques à émigrer en Angleterre, comme ce fut le cas pour Friedrich Hayek, ou aux États-Unis, comme ce fut le cas pour John von Neumann, Oskar Morgenstern, Karl Menger, ou encore Ludwig von Mises.

En 1928, Karl Menger organise le *Mathematisch Kolloquium* auquel participent régulièrement le mathématicien et logicien Kurt Gödel, le mathématicien, économiste et économètre Abraham Wald, Oskar Morgenstern et John von Neumann. Karl Menger fut, aux côtés de John von Neumann, l'un des principaux instigateurs du développement des mathématiques dans les sciences sociales. L'école autrichienne dans son ensemble ayant toujours refusé d'introduire les mathématiques, et plus spécifiquement le formalisme de Hilbert, dans le discours scientifique, c'est au sein de ce colloque que l'application des mathématiques aux problématiques des sciences sociales est discutée. Karl Menger et Oskar Morgenstern, tous deux issus du cercle de l'école autrichienne en économie, partageaient une même vision de l'état de l'économie théorique dans les années 1920 et des problèmes majeurs auxquels elle se devait de fournir une réponse.

Parmi ces problèmes, apparaissaient comme fondamentaux ceux liés à la prise de décision en présence d'incertitude, à la formation des anticipations des agents économiques, à la coordination des décisions individuelles dans un environnement où l'information pertinente, subjective, est dispersée, aux effets des enchaînements des décisions sur la structure du capital. Mais contrairement à l'école autrichienne qui focalisait ses critiques sur l'école néo-classique, Karl Menger, Oskar Morgenstern et John von Neumann ont apporté des réponses à ces problèmes en développant un nouveau programme de recherche pour la science économique. Selon James Buchanan (2001), lauréat du prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel pour ses travaux en théorie des choix publics, la théorie des jeux est l'une des cinq contributions majeures du siècle dernier en sciences économiques. Sensible aux thèses de l'école autrichienne, il note cependant que la

perte d'influence progressive de ce courant de pensée est en grande partie due à son refus d'envisager les mathématiques comme moyen d'investigation des questions économiques.

Une large partie du livre de Robert Léonard est consacrée aux discours de la méthode en sciences économiques, en mathématiques et à ses interactions. Outre les échanges entre les mathématiciens et les économistes autrichiens, Robert Léonard détaille la manière dont les mathématiques pénètrent dans les sciences économiques. Un point important de cette histoire est l'influence du programme de David Hilbert sur la pensée de Karl Menger et John von Neumann. Ce programme met en avant le traitement axiomatique de la théorie. En ce qui concerne les mathématiques, il s'agissait, d'une part, de formaliser le raisonnement propre aux mathématiques, d'autre part, de démontrer l'indépendance et la non contradiction des axiomes des théories mathématiques. L'idée d'appréhender une théorie de manière axiomatique a fait son chemin en théorie des jeux et en économie mathématique grâce, tout d'abord, à John von Neumann et Karl Menger, puis à Lloyd Shapley, John Nash, Kenneth Arrow et Gérard Debreu.

John von Neumann et Oskar Morgenstern

Il est admis que John von Neumann est le principal contributeur des résultats formels contenus dans le livre *Theory of Games and Economic Behavior*. Parmi ces résultats, trois ont marqué la discipline : la caractérisation axiomatique de la règle de décision de l'utilité espérée, le théorème du minmax pour les jeux à deux joueurs à somme nulle, et la définition de l'ensemble stable pour les jeux de coalition à plus de deux joueurs.

Supposons qu'un individu soit en mesure de classer des objets d'un ensemble X à l'aide d'une relation binaire R . Pour x et y de X , $x R y$ signifie que cet individu préfère l'objet x à l'objet y ou que l'objet x domine l'objet y ; et le symbole $\text{non}(x R y)$ signifie que x n'est pas préféré y ou de manière équivalente que x ne domine pas y . On dit qu'une fonction d'utilité u à valeurs dans l'ensemble des nombres réels représente R si pour tout x et y de X on a :

$$x R y \text{ si et seulement si } u(x) \text{ est supérieur ou égal à } u(y).$$

Cela signifie que le classement effectué par R peut être représenté par un classement dans l'ensemble des nombres réels.

Règle de décision de l'utilité espérée

Supposons que X soit composé d'un ensemble fini d'objets et que cet individu doive cette fois-ci classer des distributions de probabilité p sur X , et non pas les éléments de X , tout simplement parce qu'il n'est pas certain d'obtenir les objets de X . Par exemple, considérons deux distributions de probabilité p et q sur $X = \{a, b\}$. On dit que l'individu préfère p à q s'il préfère obtenir l'objet a avec la probabilité $p(a)$, donc obtenir l'objet b avec la probabilité complémentaire $1 - p(a)$, plutôt que d'obtenir ces mêmes objets avec la probabilité $q(a)$ et $1 - q(a)$ respectivement. On notera $P(X)$ l'ensemble des distributions de probabilité sur X , et R cette relation sur $P(X)$.

Le théorème de l'utilité espérée de von Neumann et Morgenstern fournit les conditions nécessaires et suffisantes sous lesquelles il existe une fonction d'utilité u sur X dont l'espérance mathématique représente R : pour tout p et q dans $P(X)$,

$$p R q \text{ si et seulement si } \sum_x p(x) u(x) \text{ est supérieur ou égal à } \sum_x q(x) u(x),$$

où S représente l'opération « somme ». Ces conditions nécessaires et suffisantes portent sur R ; ce sont les axiomes, jugés intuitifs, que R doit satisfaire afin que la règle de l'utilité espérée soit fondée lors d'une prise de décision en environnement incertain.

Théorème du minmax

Supposons qu'il existe deux individus dont les intérêts sont strictement conflictuels dans le sens où si le premier individu choisit l'objet x et le second individu choisit l'objet y , alors l'utilité du premier individu est $u(x, y)$ et l'utilité du second individu est l'opposé de cette quantité, c'est-à-dire $-u(x, y)$, de telle sorte que la somme des utilités est nulle. Une telle situation conflictuelle est donc appelée un jeu à somme nulle. Considérons le tableau suivant représentant les utilités u lorsque l'ensemble des objets disponibles, appelés actions ou stratégies, pour le premier individu est $X = \{a, b, c\}$ et que l'ensemble des stratégies disponibles pour le second individu est $Y = \{A, B, C\}$:

	A	B	C
a	2	0	6
b	6	3	5
c	10	1	0

Ce tableau se lit comme suit : si le premier individu choisit l'action a et le second individu choisit l'action A , alors l'utilité retirée par le premier individu est $u(a, A) = 2$, et l'utilité retirée par le second individu est $-u(a, A)$. Etant donné le choix du second individu, l'intérêt du premier individu est de sélectionner la stratégie, donc une ligne du tableau, qui maximise son utilité. Face à un choix du premier individu, l'intérêt du second individu est de choisir une stratégie, donc une colonne du tableau, qui minimise l'utilité du premier individu. Comment se comporteront ces deux individus ?

Le premier individu peut anticiper que le second individu choisira toujours la stratégie qui minimise son utilité u puisque l'utilité de ce dernier est $-u$. Donc, s'il envisage de sélectionner une quelconque action w dans $\{a, b, c\}$, alors il est convaincu que le second individu choisira une stratégie y dans $\{A, B, C\}$ qui minimise u au point w :

$$u(w, y) = \min u(w, z) \text{ sur les stratégies } z \text{ dans } Y.$$

L'intérêt du premier individu est donc de chercher le maximum parmi les utilités $u(w, y)$ en sélectionnant l'action x dans X de telle sorte que :

$$u(x, y) = \max [\min u(w, z)] = \max u(w, y) \text{ sur les stratégies } w \text{ dans } X.$$

Le second individu appliquera un raisonnement similaire : son intérêt est de minimiser les utilités $u(x, z)$ pour tout z dans Y , où :

$$u(x, z) = \max u(w, z) \text{ sur les actions } w \text{ dans } X.$$

Il sélectionnera donc une action y dans Y telle que :

$$u(x, y) = \min [\max u(w, z)] = \min u(x, z) \text{ sur les actions } z \text{ dans } Y.$$

Une solution pour un jeu à somme nulle est une paire d'actions (x, y) telle que x est dans X , y est dans Y et :

$$u(x, y) = \min [\max u(w, z)] = \max [\min u(w, z)].$$

Dans le tableau ci-dessus, la paire (b, B) procure une utilité de $u(b, B) = 3$ et réalise l'égalité ci-dessus. En effet, si le premier individu déploie la stratégie a , l'intérêt du second individu est de choisir B pour minimiser $u(a, .)$; s'il choisit b , alors le second individu choisira encore B pour minimiser $u(b, .)$; s'il choisit c , le second individu choisira C pour minimiser $u(c, .)$. L'objectif du premier individu est de sélectionner l'action qui maximise son utilité :

$$\max \{u(a, B), u(b, B), u(c, C)\} = \max \{0, 3\}.$$

Il choisira donc b, et ainsi :

$$u(b, B) = \max [\min u(w, z)].$$

Si le second individu choisit A, l'intérêt du premier individu est de choisir c pour maximiser $u(., A)$; s'il choisit B, l'intérêt du premier individu est de choisir b pour maximiser $u(., B)$; s'il choisit C, l'intérêt du premier individu est de choisir a pour maximiser $u(., C)$. Anticipant cela, l'objectif du second individu est de sélectionner la stratégie qui minimise l'utilité de son opposant :

$$\min \{u(c, B), u(b, B), u(a, C)\} = \min \{10, 3, 6\}.$$

Il choisira donc B, et ainsi on aboutit à :

$$u(b, B) = \min[\max u(w, z)].$$

Le théorème du minmax fournit des conditions suffisantes sur les ensembles X, Y et sur la fonction u afin qu'une telle paire de stratégies existe. En effet, l'existence d'une telle paire n'est pas assurée dans chaque jeu à somme nulle. Par exemple, on vérifie aisément que le jeu à somme nulle représenté par le tableau suivant ne possède pas une telle paire de stratégies :

	A	B
a	1	-1
b	-1	1

Ensemble stable

Un ensemble stable de X relativement à une préférence R est un sous-ensemble non vide S de X tel que :

- (a) Pour tout x et tout y de S, non(x R y) ;
- (b) Pour tout x dans X situé hors de S, il existe y dans S tel que y R x.

La condition (a) est une condition de stabilité interne : deux éléments de S ne se dominent pas l'un l'autre, ils sont incomparables du point de l'individu qui a pour relation de préférence R. La condition (b) est une condition de stabilité externe : tout élément situé hors de S est dominé par au moins un élément de S. En particulier, si l'individu préfère un objet situé hors de S à un objet situé dans S, alors ce premier objet ne sera pas choisi par l'individu puisqu'il existe nécessairement un autre objet appartenant à S qui lui sera préféré. Les conditions de stabilité (a) et (b) sont les critères qui permettent de justifier que les éléments de S émergent comme des choix crédibles. Notons qu'il est possible qu'un tel ensemble n'existe pas ou bien que plusieurs ensembles stables coexistent. L'ensemble stable a été appliqué en théorie des jeux coopératifs par von Neumann et Morgenstern, mais il trouve également des applications en théorie des jeux non coopératifs et dans d'autres modèles formels de situations conflictuelles (cf. Greenberg 1990).

Le passage des jeux à deux joueurs aux jeux à un nombre quelconque de joueurs pose la question du traitement de l'action collective. En effet, dès que la possibilité de former librement des accords de coopération est offerte à des coalitions d'individus, la répartition du pouvoir au sein de la population se modifie en fonction de la structure des coalitions au sein de la population. En définissant les jeux de coalition, John von Neumann et Oskar Morgenstern se sont donnés les moyens d'étudier l'action collective et ont interprété un ensemble stable comme un ensemble de recommandations acceptables reflétant un ordre social. Cette conception d'une solution comme un ensemble de recommandations acceptables pour une société a inspiré la théorie des situations sociales à Joseph Greenberg (1990). Ce

dernier présente sa théorie comme plus générale que la théorie des jeux, unificatrice des représentations et solutions pour les formes coopératives et non coopératives de la théorie des jeux.

Robert Léonard met parfaitement en scène le moment de création de l'ouvrage *Theory of Games and Economic Behavior*, les échanges scientifiques entre John von Neumann et Oskar Morgenstern, le rôle pris par chacun d'eux dans le processus de création des concepts et solutions, ainsi que la réception de l'ouvrage dans le milieu académique. Dans le dernier chapitre de son ouvrage, il met l'accent sur le rôle de la Rand Corporation dans le processus d'institutionnalisation de la théorie des jeux. La Rand Corporation, créée en 1948, est une institution à but non lucratif d'aide à la décision dans les secteurs de la sécurité nationale, du droit international, de la santé et de l'éducation. Des ingénieurs, des physiciens, des économistes, des psychologues, des mathématiciens collaborent à travers différents projets pour améliorer le bien-être et la sécurité des États-Unis. La Rand Corporation a immédiatement perçu le potentiel opérationnel de la théorie des jeux pour traiter des situations conflictuelles. Aux côtés de John von Neumann, des mathématiciens et théoriciens des jeux tels Richard Bellman, David Blackwell, John Nash, Herbert Scarf, Lloyd Shapley ont collaboré aux projets de cette institution. De mon point de vue, ce dernier chapitre est le moins réussi de l'ouvrage. Robert Léonard ne met pas suffisamment en lumière l'impact des travaux de John von Neumann et Oskar Morgenstern sur le développement de la théorie des jeux dans les années 1950. Il me semble que les travaux de Lloyd Shapley durant cette période sont un bon exemple de continuité des travaux de von Neumann et Morgenstern. En 1953, Lloyd Shapley a produit deux résultats majeurs. Il a tout d'abord étendu l'analyse des jeux à somme nulle à des situations dynamiques et *de facto* a ouvert la voie à l'étude des situations conflictuelles répétées, un thème majeur de la théorie des jeux. Il a ensuite fourni la solution la plus populaire pour les jeux de coalitions, la « valeur de Shapley ». Dans ces deux contributions, les références à l'ouvrage *Theory of Games and Economic Behavior* sont explicites. On peut même pousser l'analyse plus avant en notant que la valeur de Shapley partage plusieurs caractéristiques avec les travaux de John von Neumann et Oskar Morgenstern :

- La valeur de Shapley est introduite de manière axiomatique, comme la théorie de l'utilité espérée.
- Alvin Roth (1977), colauréat avec Lloyd Shapley du prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel, a montré que la valeur de Shapley peut être vue comme une fonction d'utilité espérée de type von Neumann et Morgenstern.
- La valeur de Shapley tient compte du pouvoir de chaque individu dans les coalitions auxquelles il peut appartenir. Elle peut d'ailleurs être mobilisée pour mesurer le pouvoir d'un électeur dans une assemblée politique.
- Elle satisfait à des propriétés intuitives de justice distributive.
- Elle admet une représentation dynamique et marginaliste, deux aspects importants dans les thèses de l'école autrichienne au même titre que la prise de décision en environnement incertain. Précisément, pour chaque ordre de formation séquentielle de la grande coalition, la valeur de Shapley rémunère chaque individu à hauteur de sa contribution marginale à la coalition qu'il rejoint relativement à cet ordre.
- La valeur de Shapley est une solution opérationnelle pour des problèmes d'imputation de coûts, de répartition de profits, de mesure de pouvoir.

Plus généralement, ce dernier chapitre ne met pas suffisamment en avant l'influence des lignes de force ouvertes par l'ouvrage de 1944 sur les travaux des années 1950 et du début des années 1960. Malgré le sentiment mitigé laissé par ce chapitre, il faut souligner encore une fois l'excellent travail historique effectué par Robert Leonard. Il s'agit d'un ouvrage référence sur l'histoire de la théorie des jeux, dont je conseille la lecture sans aucune réserve.

Références

- Buchanan J. M. (2001), Game Theory, Mathematics, and Economics, *Journal of Economic Methodology*, 8 :27-32.
- Greenberg J., *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge University Press, Cambridge, 1990.
- Maynard Smith, J., Price G (1973), The Logic of Animal Conflict, *Nature*, 246:15-18.
- Roth, A. (1997), The Shapley Value as a von Neumann-Morgenstern *Econometrica*, 45 :657-664.
- Shapley L. (1953) A Value for n-Person Games, in *Contribution to the Theory of Games* vol. II (H.W. Kuhn and A.W. Tucker eds). Annals of Mathematics Studies 28, Princeton University Press.
- Shapley L. (1953) Stochastic Games. *Proceedings of the National Academy of Sciences*, 39 :1095-1100.

Publié dans laviedesidees.fr, le 31 décembre 2014

© laviedesidees.fr