

Chapitre 4

Agents et dilemmes sociaux

4.1. Dilemme et stratégies

L'étude de la coopération et de l'évolution de la coopération entre agents est souvent faite via un modèle mathématique issu de la théorie des jeux. L'utilisation de ce jeu introduit par M.Flood et M. Dresher dans [FLO 52] s'est généralisé après la publication du livre de R. Axelrod [AXE 84]. Le Dilemme Itéré du Prisonnier Classique (Classical Iterated Prisoner's Dilemma ou CIPD) est un jeu simultané non-coopératif, à somme non-nulle, impliquant deux joueurs. Chacun des deux joueurs a le choix entre deux cartes : **C** (pour Coopération) ou **D** (pour Defection, c'est-à-dire Trahison en français). Les gains de chacun des joueurs sont définis dans le tableau 1. Pour qu'il y ait dilemme, qui repose sur le fait que l'intérêt individuel diffère de l'intérêt collectif, il suffit que l'inéquation $S < P < R < T$ soit respectée. Ce modèle a été très largement utilisé et étudié et est un des modèles les plus particuliers dans sa classe de jeu, voir [RAG 66].

	Cooperate	Defect
Cooperate	R=3	S=0
Defect	T=5	P=1

Tableau 1 Matrice de gain du Dilemme Itéré du Prisonnier Classique

Le dilemme simple est résolu en Théorie des jeux via l'équilibre de Nash qui consiste en la trahison des deux joueurs. Le modèle est donc étendu : dans le

2 Titre de l'ouvrage

dilemme itéré du prisonnier les joueurs se rencontrent plus d'une fois sans savoir quand aura lieu la dernière rencontre. Le gain d'un joueur est alors la somme des gains qu'il a emporté lors de chacune des rencontres. De manière à favoriser la coopération mutuelle par rapport à la trahison et pour conserver la différence entre intérêt individuel et intérêt collectif l'inéquation suivante est alors respectée : $S+T < 2R$. Les valeurs classiques des gains sont données dans le tableau 1.

Avec un tel jeu ce que l'adversaire a fait dans le passé peut influencer la manière dont un joueur va choisir le coup à jouer à un instant donné. En d'autres termes il est alors facile d'étudier les comportements des joueurs (nommés stratégies en Théorie des Jeux). Il existe deux grandes familles de stratégies :

- les stratégies pures sont complètement déterministes. Ce qu'un agent fait lors d'une itération particulière est complètement prédéfini par l'historique de la rencontre des deux agents. Le comportement est complètement calculable. Le hasard n'intervient pas.
- les stratégies mixtes utilisent des distributions de stratégies pures. Face à deux situations identiques, c'est-à-dire deux historiques deux rencontres semblables, un agent sera capable de choisir deux coups différents. Le hasard peut intervenir.

Le dilemme itéré du prisonnier simple est un jeu à une étape. Deux stratégies seulement sont disponibles. On nomme ces deux stratégies, la coopération et la trahison, mais toute autre appellation peut être utilisée. On dit souvent être gentil ou être méchant, partager ou ne pas partager, etc. Tout n'est donc qu'une question de vocabulaire.

Dans le cas du dilemme simple, les stratégies sont au nombre de deux alors que dans le cas du dilemme itéré le nombre de stratégies est, a priori, indénombrable. Il faut donc choisir un mode de représentation. Le choix de description via le vocabulaire courant et la description des comportements humains n'est donc pas moins acceptable que le choix, par exemple, de la description extensive de tous les cas qu'une stratégie peut rencontrer et de la carte à jouer pour chacun de ceux-ci. Les représentations intensionnelles des stratégies sont préférables aux représentations par extension.

4.1.1. Les comportements modélisés par des stratégies

Voici quelques exemples de stratégies en commençant par des stratégies pures.

- **all_c** « Je coopère toujours, quel que soit le comportement de mon adversaire. » Cette stratégie est souvent nommée gentille.

- **all_d** « Je trahis toujours, quel que soit le comportement de mon adversaire. » Cette stratégie est souvent nommée méchante.
- **per_ccd** « Je coopère puis je coopère puis je trahis, puis je coopère, puis je coopère puis je trahis, puis ... » Cette dernière stratégie fait partie des stratégies qui seront dites périodiques. Elles jouent cycliquement la même suite de cartes. Cette suite est appelée la période. Ici la période est composée de deux coopérations suivie d'une trahison. On note ce comportement : (CCD) *. Un nombre infini de stratégies périodiques peut être décrit. Nous nous contenterons de celle-ci.

Les stratégies qui suivent sont réactives. Dans au moins un cas, elles se comportent différemment après un changement de comportement de l'adversaire.

- **spiteful** « Je coopère jusqu'à ce que mon adversaire ait trahi, après quoi je trahis toujours quel que soit son comportement suivant. » Cette stratégie est souvent nommée **rancunière**, **trigger**, ou encore **grim**.
- **easy_go** « Je trahis jusqu'à ce que mon adversaire ait coopéré, après quoi je coopère toujours quel que soit son comportement suivant. »
- **tit_for_tat** « Je coopère au premier coup, ensuite à chaque coup, je joue la carte jouée par mon adversaire au coup précédent. » Cette stratégie est souvent nommée **donnant_donnant**, **oeil_pour_oeil**, ou surnommée **tft**. Elle a été créée par Anatol Rapoport lors du concours organisée par Robert Axelrod.
- **mistrust** « Je trahis au premier coup, ensuite à chaque coup je joue la carte jouée par mon adversaire au coup précédent. » Cette stratégie est parfois nommée **méfiant**e, ou **supiscious_tit_for_tat**.
- **soft_majo** « Je joue la carte que mon adversaire a majoritairement jouée dans l'histoire passée de la partie. S'il a joué autant de fois C que D, alors je coopère. » Cette stratégie est souvent nommée **majo_mou**. Les stratégies utilisant ce principe de prise de décision en fonction du coup joué le plus souvent par l'adversaire sont dites majoritaires.
- **pavlov** « Je coopère au premier coup, puis je coopère uniquement si mon adversaire et moi avons joué la même carte au coup précédent, sinon je trahi. » Cette stratégie a été introduite et étudiée par Kraines et Kraines dans [KRK 89] puis Nowak et Sigmund dans [NOS 93]. Elle est aussi nommée **P1**, ou encore **simpleton**.

4 Titre de l'ouvrage

- **prober** « Je commence par jouer la séquence DCC, puis si mon adversaire a coopéré aux coups 2 et 3 alors je joue comme **all_d**, sinon comme **tit_for_tat**. » Cette stratégie est souvent appelée sondeur. L'idée de base de cette stratégie est de tester son adversaire dès le début de la partie et d'adapter son comportement à celui de l'adversaire. Les stratégies utilisant ce principe sont appelées des sondeuses.

Voici maintenant quelques stratégies mixtes, que l'on nomme plus fréquemment des stratégies probabilistes, du fait qu'elles font appel au hasard.

- **ipd_random** « Je coopère et trahis avec une probabilité de 0,5 à chaque coup. » Ce comportement peut se résumer au lancer d'une pièce de monnaie, en attachant la décision du jeu de la carte C à une des faces et de la carte D à l'autre. Cette stratégie est parfois nommée **lunatique**, **random** ou encore **hasard**. Une longue série de stratégies probabilistes peut ainsi être définie. Chaque stratégie est alors parfaitement décrite par la probabilité p de coopérer. La probabilité de trahir est alors $1-p$.

Cette idée de probabilité, et d'intervention du hasard dans le comportement d'un individu peut être adaptée aux différentes stratégies que nous avons décrites précédemment. Ainsi la stratégie suivante est une adaptation probabiliste de **tit_for_tat**.

- **hard_joss** « Je coopère au premier coup, ensuite si l'adversaire a trahi au coup précédent alors je trahis sinon je trahis avec une probabilité de 10%, et coopère avec une probabilité de 90%. »

D'autres comportements probabilistes sont envisageables. En effet, l'appel au hasard peut se faire en conjonction avec d'autres paramètres, comme le numéro du coup par exemple. Considérons que ce dernier est nommé TURN.

- **worse_and_worse** « Je trahis avec une probabilité égale à $\text{TURN}/1000$, c'est-à-dire de plus en plus. »

4.1.2. Une ébauche de taxonomie

Toutes les stratégies décrites peuvent être caractérisées par certains traits. Certaines des caractéristiques importantes qui peuvent être utilisées dans la composition d'un comportement sont :

- **bienveillance**. Une stratégie est dite bienveillante quand elle ne prend jamais l'initiative de la trahison. On dit aussi souvent que c'est de la gentillesse.

- **agressivité.** Une stratégie est dite agressive quand, au contraire, elle prend l'initiative de la trahison. On dit aussi que c'est de la méchanceté.
- **réactivité.** Une stratégie est dite réactive quand elle change son comportement après un changement de comportement de son adversaire. Le changement de comportement de la stratégie peut intervenir à n'importe quel moment après celui de l'adversaire, et pas forcément au coup juste après.
- **indulgence.** Une stratégie réactive est dite indulgente quand sa réaction n'implique pas que son comportement soit définitivement modifié. Plus exactement la période de réaction ne doit pas être de taille infinie.

Le tableau 2 caractérise l'ensemble des stratégies qui ont été décrites.

Stratégies	Bienveillance	Agressivité	Réactivité	Indulgence	Probabiliste
All_c					
All_d					
Per_ccd					
spiteful					
Easy_go					
tit_for_tat					
mistrust					
soft_majo					
pavlov					
prober					
ipd_random					
Hard_joss					
worse_and_worse					

Tableau 2 Caractérisation des stratégies exemples

4.2. Évaluations des stratégies

Maintenant que quelques stratégies ont été décrites, l'étape naturelle suivante est de tenter de les comparer, de les évaluer.

4.2.1. Tournoi et évolutions

La première évaluation possible d'une stratégie est la comparaison de son score à celui d'une autre stratégie. Plus une stratégie a un score élevé, mieux on la considère.

Il semble évident que, pour évaluer des stratégies, il faut avant tout pouvoir évaluer leur score. Pour évaluer les stratégies d'un panel donné, on peut imaginer faire la somme des scores de chaque stratégie de ce panel face à toutes les stratégies de ce panel. On imagine avec cette méthode faire une sorte de championnat sportif entre stratégies, on appelle souvent cela un tournoi. Le score d'une stratégie est alors la somme des scores de cette stratégie face à toutes les stratégies du panel (y compris elle même). A la fin du tournoi il est alors possible de classer les stratégies du panel en fonction de leur score. Les bonnes stratégies en tournoi sont bien adaptés à leur environnement, mais sont souvent peu robustes aux modifications de cet environnement. Même si les tournois sont un bon moyen de comparaison de stratégies, une évaluation sérieuse ne peut se contenter de ce manque de robustesse des résultats.

Une seconde méthode d'évaluation des stratégies est alors de simuler les principes de la sélection naturelle : moins un individu est fort, moins il a de chance de survivre. Cette simplification des mécanismes d'évolution est utilisée ici dans le cadre d'évolution de populations d'agents utilisant les stratégies d'un panel.

L'idée est de faire en sorte que chaque agent choisisse une stratégie dans le panel et rencontre un à un tous les autres agents. Il y a donc plusieurs représentants d'une même stratégie. Une fois que toutes les rencontres ont eu lieu chacun est alors capable de connaître la valeur de la stratégie qu'il utilise en cumulant les scores obtenus face à tous les autres. Les agents ayant choisi une stratégie ramenant peu de points ont alors tendance à disparaître, alors que ceux ayant choisi une stratégie rapportant beaucoup de points vont au contraire *croître et multiplier*. Le cycle recommence alors et on peut voir fluctuer la représentation des différentes stratégies dans la population globale.

Cette méthode est liée à l'étude des dynamiques de population. Soit une population de n joueurs adoptant chacun une stratégie d'un panel. Au début on considère que chaque stratégie du panel est représentée par des sous-populations de taille égale. Un tournoi impliquant tous les joueurs est alors effectué. Les stratégies ayant obtenues un bon score sont favorisées alors que les mauvaises sont défavorisées par une redistribution proportionnelle des sous-populations. Le processus de redistribution, aussi appelé génération, est répété jusqu'à stabilisation, c'est-à-dire jusqu'à ce qu'il n'y ait pas de modifications des sous-populations entre deux générations. Une bonne stratégie est alors une stratégie qui reste présente dans la population le plus longtemps possible et avec la plus grande taille de sous-population possible. Ce type d'évaluation est appelé évolution écologique.

4.2.2. Classes complètes

Dans [BDM 98] cette méthode est étendue à une échelle beaucoup plus large. Une méthode descriptive de définition de stratégies est mise au point. Des structures, aussi appelées *génotype*, pouvant être décodées en un comportement particulier, un *phénotype*, sont utilisées. Une manière de définir un stratégie est alors simplement de remplir une telle structure. Obtenir un grand nombre de stratégies de manière objective se ramène alors simplement à considérer toutes les façons de remplir une structure, autrement dit de prendre en compte tous les individus basés sur un génotype particulier. L'ensemble des stratégies décrites par un génotype particulier est alors appelée une *classe complète*. Pour évaluer une stratégie il suffit alors de calculer une évolution écologique impliquant celle-ci ainsi que l'ensemble des stratégies d'une classe complète.

Un certain nombre de classe complète ont été définies en se basant sur des idées simples telle la taille de la mémoire des stratégies (de combien de coups une stratégie peut se souvenir pour choisir la carte à jouer au coup suivant). Il est alors aisé de comparer deux stratégies en comparant leur rang à la fin d'une évolution écologique.

Voici une liste des grandes familles de génotypes utilisés :

- **memory** Les stratégies de cette famille utilisent une mémoire limitée du passé. Cette mémoire est différente pour les coups qu'elles ont joués et pour les coups que leur adversaire a joué. Chaque stratégie est définie complètement par les cartes à jouer au début du jeu et par les cartes à jouer dans chacune des configurations du passé possible.

Considérons, par exemple, toutes les stratégies ne pouvant utiliser que leur dernier coup et le dernier coup joué par leur adversaire. Une de ces stratégies peut être définie de cette manière :

au premier coup je joue **C** ensuite

Si j'ai joué **C** et qu'il a joué **C** alors je joue

C

8 Titre de l'ouvrage

Si j'ai joué **C** et qu'il a joué **D** alors je joue

Si j'ai joué **D** et qu'il a joué **C** alors je joue

Si j'ai joué **D** et qu'il a joué **D** alors je joue

D
D
D

Le génotype de cette stratégie est noté :

C	C	D	D	D
---	---	---	---	---

Avec un tel génotype $2^5=32$ stratégies peuvent être définies, y compris certaines stratégies classiques. De façon à simplifier l'identification des stratégies toutes les informations concernant le génotype d'une stratégie sont utilisées pour nommer celle-ci. Des préfixes particuliers sont utilisés pour identifier la famille de la stratégie. Pour la famille **memory** c'est le préfixe **mem** qui est utilisé. La stratégie décrite en exemple est nommée **mem_1_1_ccddd**, c'est la même que **spiteful**.

- **binary** Les stratégies de cette famille utilisent le même concept que celle de la famille **memory** mais utilisent en plus un indicateur déterminant si l'adversaire a plus souvent trahi ou coopéré depuis le début de la rencontre. Le préfixe utilisé pour cette famille est **bin**.
- **moore** Les stratégies de cette famille sont simplement des automates de Moore. Le préfixe utilisé pour cette famille est **moore**.
- **mealy** Les stratégies de cette famille sont simplement des automates de Mealy. Le préfixe utilisé pour cette famille est **mealy**.

4.2.3. Quelques résultats

Dans une série d'articles repris dans [AXE 84], Robert Axelrod rapporte les résultats de tournois et évolutions écologiques qu'il a menés grâce à des simulations informatiques. C'est, après ses travaux et la publication de son livre, que le dilemme du prisonnier a connu une très grande renommée. Les expérimentations d'Axelrod, mais aussi son immense talent de présentation du dilemme du prisonnier et surtout de ses applications et implications ont convaincu un grand nombre de l'intérêt de ce jeu comme modèle d'étude de la coopération. Depuis lors, l'intérêt soulevé par le dilemme du prisonnier n'a jamais faibli. On peut par exemple citer une bibliographie établie par Axelrod et D'Ambrosio pour la période de 1988 à 1994, dans laquelle ils citent plus de 200 références sur le dilemme du prisonnier.

L'ouvrage d'Axelrod, bien que peu technique et formel, capture bien l'essence du dilemme du prisonnier et livre vraiment certaines clés sur l'étude de la coopération à tel point qu'il est souvent considéré comme le livre à la base d'une théorie de la coopération.

Les conclusions les plus importantes qu'Axelrod tire de ses travaux sont que pour être bonne une stratégie doit satisfaire quatre critères :

- la gentillesse, ou encore la bienveillance
- la réactivité
- l'indulgence
- la simplicité

Il exhibe la stratégie `tit_for_tat` comme exemple parfait de stratégie comportant ces caractéristiques.

Il paraît surprenant que, dans un modèle où la communication est très difficile, la coopération ne soit favorisée que par des actes simples. La difficulté de communiquer, et donc de comprendre ou de se faire comprendre par son adversaire fait plutôt penser, qu'il faut savoir interpréter des comportements très différents, et que cette capacité n'est pas favorisée par la simplicité des comportements.

4.3. Comportements adaptatifs

4.3.1. *gradual*

La stratégie la plus souvent mise en valeur par les différents travaux existants est sans aucun doute `tit_for_tat`. Les qualités que lui accorde Axelrod sont du même coup considérées comme absolues. La force de `tit_for_tat` a été remise en cause par certains auteurs, [NOS 93] et [BOL 87] par exemple. Plus précisément une de ses caractéristiques est remise en doute : la simplicité. Ce point de vue avait déjà été discuté par Delahaye et Mathieu dans [DEM 95] à propos d'une variante du dilemme du prisonnier avec renoncement. Dans ce but, les capacités des stratégies connues sont étendues en les modifiant peu à peu et en vérifiant que les modifications améliorent les stratégies, tant en tournoi, qu'en compétition écologique. Durant ces nombreux essais un grand nombre de stratégies ont été créées dont la stratégie `gradual`, introduite dans [BDM 96], dont le comportement est rappelé ici :

« Je coopère jusqu'à la première trahison de mon adversaire. Ensuite pour chaque trahison je punis mon adversaire par n trahison(s), le calme par 2 coopérations et continue à coopérer. n est le nombre de fois que mon adversaire m'a trahi. »

Cette stratégie est apparue très forte par rapport aux stratégies classiques, et notamment par rapport à `tit_for_tat`.

10 Titre de l'ouvrage

Elle se comporte exactement comme `tit_for_tat`, excepté lorsqu'il s'agit de punir son adversaire et de se souvenir du passé. Elle coopère au premier coup et ne commence jamais à trahir avant que son adversaire ne l'ait fait. Tout comme `tit_for_tat` elle est bienveillante.

Après la première trahison de l'adversaire, elle se met à trahir, et change donc son comportement. Tout comme `tit_for_tat` elle est réactive. Après une certaine période de punition, elle retrouve son comportement initial de coopération. Tout comme `tit_for_tat` elle est indulgente.

A priori elle possède les trois premières caractéristiques qu'Axelrod préconise.

Possède-t-elle la quatrième caractéristique ? Il semble que non. Il est évident que quantifier la complexité dans l'absolu est relativement difficile. Cependant si on la quantifie en termes de nombres d'actions effectuées par la machine de Turing correspondante, il est particulièrement clair que `tit_for_tat` est beaucoup plus simple que `gradual`. En effet `tit_for_tat` n'a besoin de connaître que la situation de jeu précédente pour déterminer son coup suivant, alors que `gradual` a besoin non seulement de cette entrée mais également d'autres informations, ne serait-ce que le nombre de fois que l'adversaire a déjà trahi. Il semble donc clair que le comportement de `gradual` est plus complexe à décrire que celui de `tit_for_tat`. De là vient assez naturellement qu'il est aussi plus complexe de le comprendre que celui de `tit_for_tat`.

La caractéristique que `gradual` possède en plus des trois dernières est la gradualité dans la réaction, autrement l'évolutivité de la punition. Cette caractéristique lui permet d'essayer de faire comprendre à son adversaire qu'il vaut mieux éviter d'être agressif. Plus l'adversaire l'est plus `gradual` le sera aussi.

Les inspirations de `gradual` sont assez naturelles. La nouvelle caractéristique qu'elle exhibe peut se retrouver dans la vie quotidienne. Ce comportement est utilisé par exemple par des créanciers avec leurs débiteurs. On peut par exemple imaginer le comportement du gouvernement et des contribuables, ou des compagnies de téléphone ou d'électricité avec leurs clients.

Le comportement de `gradual` peut cependant être interprété de deux manières différentes :

- le joueur est très offensif. Il veut forcer son adversaire à coopérer. Il lui montre donc clairement qu'il a tout intérêt à ne pas le trahir.
- le joueur est très défensif. Il ne veut pas se faire exploiter. Il évite de plus en plus la coopération avec son adversaire, en se rapprochant de plus en plus de la solution de l'équilibre de Nash du dilemme simple.

Ces deux explications sont deux possibilités d'interprétation du jeu, comme il y a deux approches (au moins) des relations avec les autres. Dans le premier cas on essaie d'expliquer à l'autre quel est le meilleur choix pour les deux participants, dans le second on essaie de se protéger. C'est une sorte de choix entre l'ouverture et la fermeture de cette relation. Il est clair que ce choix n'est pas simple. **gradual**, contrairement à **tit_for_tat**, offre ce type de complexité aux joueurs.

On peut noter qu'une stratégie semblable à **gradual** est présentée par Shubik dans [SHU 70].

4.3.2. *bad_bet*

Comme pour le cas de **gradual**, une autre stratégie a été inspirée par les stratégies classiques ou par les résultats des différents concours organisés, notamment par Delahaye et Mathieu. La stratégie **bad_bet**, introduite dans [BDM 01], a le comportement suivant :

- Tant que l'adversaire coopère, elle coopère.
- Dès qu'il a trahi une fois, elle suit ce principe :
 - pendant 4 coups elle joue la stratégie **tit_for_tat**.
 - puis pendant 4 coups elle joue **all_c**.
 - puis pendant 4 coups elle joue **spiteful**.
 - puis pendant 4 coups elle joue **per_ccd**.
 - elle compare les scores relatifs, c'es-à-dire limités à la période d'essai, obtenus lors de l'essai des 4 stratégies par chacune d'entre elles, et choisit de jouer la stratégie qui a obtenu le score le plus élevé, pendant 4 coups.
 - elle met à jour (simple remplacement) le score relatif obtenu par la stratégie testée.
 - elle retourne au point 5.

bad_bet essaie plusieurs réactions lorsqu'elle se retrouve face à un adversaire agressif. Elle vérifie continuellement si la réaction qu'elle utilise est la plus efficace face à son adversaire. Elle tente d'adapter sa réaction au type d'agressivité de son adversaire. C'est là son idée principale : être *réactive*, non seulement dans le temps (un peu à la manière de **gradual**) mais aussi dans l'espace (le comportement adopté ne sera pas le même face à deux adversaires différents).

Cette adaptation a, une fois de plus, un prix et quelle que soit la mesure de complexité utilisée **bad_bet** est moins simple que **tit_for_tat** mais obtient de bien meilleurs résultats.

12 Titre de l'ouvrage

Cette adaptation joue bel et bien un rôle dans les bons résultats de **bad_bet**. Le tableau 3 présente un exemple de trace du comportement de **bad_bet** dans un tournoi particulier. Pour chacune des stratégies pour laquelle une punition a été nécessaire, on peut y lire tout d'abord le numéro du coup où la punition a commencé, la date de début de punition, puis pour chaque période, le numéro de la stratégie utilisée. Les quatre premières périodes sont celles que toute stratégie **bad_bet** teste avant de commencer à faire son choix. Pour les stratégies utilisant des choix aléatoires, seule la trace pour la première des rencontres a été conservée.

Adversaire	Date	Comportement utilisé
all_d	1	012300000000000000000000000000 0
ipd_random	2	01230022222222222222222222222222 2
per_ccd	3	01230022222222222222222222222222 2
prober	1	01230000000000000000000000000000 0
mistrust	1	0123111111111111111111111111111111 1
hard_joss	29	0123111111111111111111111111111111
worse_an_worse	15	01230333333333333300011111111111
easy_go	1	0123333333333333333333333333333333 3

Tableau 3 Pistage du comportement de **bad_bet**. Les stratégies utilisées par **bad_bet** sont 0=**tit_for_tat**, 1=**all_c**, 2=**spiteful** et 3=**per_ccd**

De nombreuses traces de ce genre ont été calculées. A chaque fois, on s'aperçoit que non seulement toutes les stratégies de la gamme de punition de **bad_bet** sont utilisées, mais qu'en plus au cours d'une rencontre plusieurs stratégies différentes sont parfois utilisées pour punir un adversaire récalcitrant.

L'analyse de ces traces permet de découvrir de nouvelles caractéristiques, qui semblent être des propriétés intéressantes pour la qualité d'une stratégie. L'idée globale de **bad_bet** est donc celle d'une stratégie gentille, réactive mais pas forcément indulgente. Plus précisément **bad_bet** adapte son comportement à son adversaire. Elle considère que toutes les trahisons de l'adversaire ne sont pas

forcément des tentatives d'exploitation. **bad_bet** considère qu'une trahison est une tentative de communication. Elle essaie alors différents langages et utilise celui qui lui coûte le moins cher. Ce choix reste cependant dynamique durant toute la longueur du reste de la partie pour le cas où l'adversaire réussit à apprendre le langage que **bad_bet** a choisi.

4.3.3. Expériences et analyses

De très nombreuses expériences ont été menées notamment dans le cadre de [BEA 00]. Deux résultats particulièrement significatifs illustrant les forces des stratégies décrites ici et par la même la faiblesse de **tit_for_tat** sont donnés par la figure 1 et le tableau 4. Ils n'ont pas à eux seuls force de démonstration mais permettent une illustration flagrante des expériences menées.

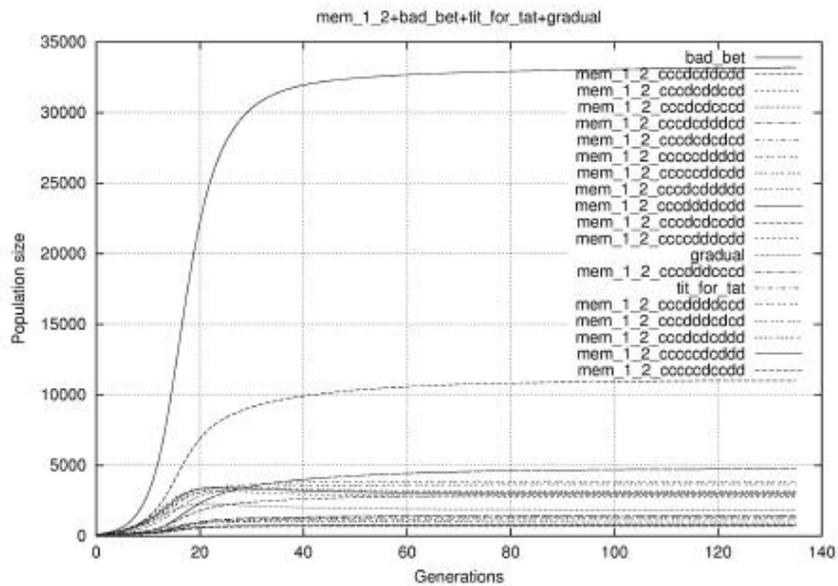


Figure 1 Evolution de la classe complète **mem_1_1** impliquant également **gradual** et **bad_bet**. La figure ne représente que les 20 meilleures stratégies au terme de l'évolution

Le tableau 4 présente l'évaluation de **gradual** et **bad_bet** dans un certain nombre de classes complètes. La première colonne présente le nom de la classe

complète utilisée, la seconde présente le nombre de stratégies définies par la classes complète, les deux dernières colonnes présentent le rang de chacune des deux stratégies évaluées à la fin d'une évolution impliquant toutes les stratégies de la classe complète utilisée plus la stratégie à évaluer. Les cases noircies représentent des cas où la stratégie évaluée arrive en deça des 20 premières stratégies. A chaque fois le mauvais score de **gradual** comme de **bad_bet** sont expliqués par la nature des comportements codés par la classe complète utilisée pour l'évaluation.

Classe	Taille	gradual	bad_bet
mem_0_1	8	1	1
mem_0_2	64	5	1
mem_1_1	32	2	1
mem_1_2	1024	10	1
mem_2_1	1024		
memd_0_2	128		1
memd_1_2	2048	15	1
memd_2_1	2048		
memld_0_3	4096		1
bin_0_1	32	1	1
bin_0_2	1024		1
bin_1_1	512	1	11
bind_0_2	2048		1
moore_0_1_2	128	3	1
moore_1_1_2	2048		1
mealy_0_1_2	1024	1	1

Table 4 Evaluation de **tit_for_tat**, **gradual** et **bad_bet** dans quelques classes complètes

Ce qui semble émerger très nettement des différentes expériences et de leurs résultats est qu'un critère semble indispensable à une stratégie si elle veut être efficace : l'*adaptation*. **bad_bet**, par exemple, s'adapte à son adversaire en

tentant d'utiliser quatre réactions différentes que l'on peut bien isoler. Ces quatre types de réponse à un agresseur sont :

- la punition simple. Dans ce cas **bad_bet** considère que son adversaire comprend bien les punitions, et qu'elles suffisent à ne pas lui faire perdre trop de points.
- le pardon ou l'oubli. Dans ce cas **bad_bet** ferme les yeux sur la trahison de son adversaire, parce qu'elle considère que tenter de le convaincre lui coûterait trop cher.
- la punition sévère. Dans ce cas **bad_bet** comprend que l'adversaire est vraiment trop agressif ou sournois et préfère rompre toute tentative de coopération.
- l'exploitation. Dans ce cas **bad_bet** essaie de profiter de la naïveté de son adversaire.

Dans tous les cas, **bad_bet** accumule des qualités exhibées par d'autres stratégies, comme la gentillesse et la réactivité de **tit_for_tat**, l'évolutivité de la réaction de **gradual**, et enfin l'adaptation du comportement, et plus précisément de la réaction.

Cela conforte l'idée que la simplicité n'est pas une bonne qualité, mais qu'au contraire il semble exister une hiérarchie de stratégies aux qualités et à la complexité croissantes.

4.4. Conclusions

Le dilemme itéré du prisonnier classique est et reste un des meilleurs modèles formels d'étude de l'interaction entre agents et plus particulièrement de la modélisation de la coopération. Sa simplicité apparente ne l'empêche pas de pouvoir représenter un très vaste éventail de comportements. L'utilisation de simulations informatiques permet d'évaluer l'influence d'un comportement sur l'évolution de la coopération dans un système multi-agents au sens le plus large.

Les résultats classiquement admis sont désormais clairement remis en cause par l'introduction de nouvelles stratégies faisant apparaître de nouveaux traits de caractères qui semblent impliquer non seulement des résultats beaucoup plus efficaces mais aussi et parallèlement un accroissement de la complexité des comportements.

Il semble notamment que la capacité pour un agent de savoir s'adapter à ses interlocuteurs soit primordiale. Cette adaptation doit être la plus large possible et ne doit pas se limiter à une dimension temporelle, mais doit savoir prendre en compte l'espace, c'est-à-dire la diversité de l'environnement dans lequel l'agent évolue.

Une bonne stratégie pour le dilemme itéré du prisonnier classique doit posséder deux caractéristiques essentielles : la *bienveillance* et la *réactivité adaptative* (dans laquelle on peut par exemple considérer la faculté de pardonner).

La faculté de s'adapter paraît donc être un facteur essentiel, sinon indispensable, à l'émergence et la stabilité de la coopération dans un monde composés de nombreux agents aux intérêts et donc aux comportements hétérogènes.

D'une manière plus large il ne semble pas complètement stupide de penser que cette faculté est sans doute une des composantes importantes de l'intelligence des agents.

Les différents logiciels utilisés pour les simulations dont il est fait référence ici, ainsi qu'un grand nombre d'informations sur le modèle sont disponibles sur le web à l'URL : <http://www.lifl.fr/IPD>

Bibliographie

- [AXE 84] AXELROD R., *The evolution of cooperation*, Basic Books, 1984
- [BDM 96] BEAUFILS B., DELAHAYE J.-P., MATHIEU P., *Our meeting with gradual, a good strategy for the iterated prisoner's dilemma*, Proceedings of the Artificial Life V conference pages 202-209, MIT Press, 1996
- [BDM 98] BEAUFILS B., DELAHAYE J.-P., MATHIEU P., *Complete classes of strategies for the classical iterated prisoner's dilemma*, Proceedings of the Evolutionary Programming VII conference, Lecture Notes in Computer Science, volume 1447 pages 33-41, Springer, 1998
- [BDM 01] BEAUFILS B., DELAHAYE J.-P., MATHIEU P., *Adaptive Behaviour in the Classical Iterated Prisoner's Dilemma*, Proceedings of the AISB'01 Symposium, 2001
- [BEA 00] BEAUFILS B., *Modèles et simulations informatiques des problèmes de coopération entre agents*, Thèse de l'Université des Sciences et Technologies de Lille, 2000
- [BOL 87] BOYD R., LOBERBAUM P., *No pure strategies is evolutionarily stable in the repeated prisoner's dilemma*, Nature, volume 327 pages 58-59, 1987
- [DEM 95] DELAHAYE J.-P., MATHIEU P., *Complex strategies in the iterated prisoner's dilemma*, dans Chaos and Society, IOS Press, 1995
- [FLO 52] FLOOD M., *Some experimental games*, Research Memorandum RM-789-1, The RAND Corporation, 1952
- [KRK 89] KRAINES D., KRAINES V., *Pavlov and the prisoner's dilemma*, Theory and decision, volume 26 pages 47-79, 1989
- [NOS 93] NOWAK M., SIGMUND L., *A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game*, Nature, volume 364 pages 56-58, 1993
- [RAG 66] RAPOPORT A., GUYER M., *A taxonomy of 2x2 games*, General Systems, volume 11 pages 203-214, 1966

[SHU 70] SHUBIK M., *Game theory, behavior, and the paradox of the prisoner's dilemma: three solutions*, Journal of Conflict resolution, volume XIV pages 181-193, 1970