# La coopération dans le dilemme du prisonnier

**Karine Fradet**

**karine.fradet@umontreal.ca**

Résumé

La tension présente dans un jeu du type du dilemme du prisonnier vient du fait que la solution optimale et la solution formant l'équilibre de Nash ne correspondent pas. Pour atteindre la solution optimale, les deux joueurs doivent choisir de coopérer, mais cette solution est contraire à la rationalité dans la théorie des jeux. Les caractéristiques du dilemme du prisonnier sont présentées afin de mieux explorer cette tension et des pistes de réflexion sont présentées et critiquées quant aux aspects pouvant permettre l'atteinte de la coopération réciproque.

En théorie des jeux, le dilemme du prisonnier est porteur d'un paradoxe : bien qu'il soit toujours plus payant pour un joueur de se défiler, la coopération réciproque est la solution optimale à la fois d'un point de vue individuel et d'un point de vue collectif. Cette caractéristique fait en sorte qu'il est classé parmi les dilemmes sociaux, soit des jeux où les intérêts individuels et collectifs entrent en conflit (Colman 2003). Ce coût associé à la décision de coopérer pose la question suivante : faut-il que les agents ne soient pas rationnels pour atteindre la solution optimale? D'un autre côté, le choix habituellement considéré comme rationnel mène à une solution très peu satisfaisante pour les deux joueurs.

Nous débuterons par présenter les caractéristiques du dilemme du prisonnier. Nous examinerons ensuite le cas des dilemmes du prisonnier itérés afin de considérer l'effet de la répétition sur le choix de la stratégie. Nous évaluerons finalement quelques solutions proposées pour expliquer comment les agents coopèrent dans une matrice du dilemme du prisonnier.

## La structure de base du dilemme du prisonnier

La structure de base d'un dilemme du prisonnier (DP) est représentée par la matrice suivante.

| | **C** | **D** |
|---|---|---|
| **C** | *R, R* | *S, T* |
| **D** | *T, S* | *P, P* |

Figure 1 : Matrice de la structure de base d'un DP.

Deux joueurs, en rangée et en colonne, doivent choisir entre deux alternatives : coopérer (C) ou se défiler (D). Chaque joueur recevra alors l'un des quatre gains suivants : si les deux joueurs coopèrent, ils recevront la récompense (*R*); s'ils se défilent tous les deux, ils recevront la punition (*P*); si l'un coopère alors que l'autre se défile, celui ayant coopéré recevra (ou sera) le sot (*S*) alors que celui s'étant défilé recevra la tentation (*T*) (Kuhn 2009). Pour que la matrice soit considérée comme un DP, la valeur de ces gains doit être ordonnée comme suit :

DP1[1]) $T > R > P > S$

Trois raisons poussent le joueur à choisir D (Howard 1988). Tout d'abord, dans une telle situation, le choix D domine strictement le choix C : peut importe le choix de l'autre joueur, il est toujours plus payant de choisir D. Ainsi, si l'adversaire choisit C, le fait de choisir D donne $T$ plutôt que $R$. Si l'adversaire choisit D, le fait de choisir D donne $P$ plutôt que $S$. Choisir D est donc plus avantageux dans les deux cas. Deuxièmement, D est une stratégie maximin : le choix D garantit d'avoir au moins $P$ alors que le choix C risque de rapporter $S$. Troisièmement, le résultat (D,D) est l'équilibre de Nash : aucun des deux joueurs n'a avantage à modifier son choix de façon unilatérale en sachant ce que l'autre a choisi.

La tension présente dans un DP vient du fait que (D,D) n'est pas une situation optimale[2]. Bien qu'il soit plus avantageux pour chaque joueur de choisir D, ils tireraient un bénéfice commun s'ils s'entendaient pour jouer C : le résultat (C,C) leur procurerait alors le gain $R$ plutôt que le gain $P$. Par exemple, considérons la matrice des gains suivante.

| | **C** | **D** |
|---|---|---|
| **C** | *3, 3* | *0, 5* |
| **D** | *5, 0* | *1, 1* |

Figure 2 : Matrice des gains d'un DP de base.

Il est plausible de considérer le choix C : le résultat (C,C) donnerait 3 à chacun des joueurs, ce qui est la situation optimale. Toutefois, la théorie des jeux suppose des joueurs rationnels qui cherchent à maximiser leur utilité (Luce and Raiffa 1989) : faire le choix D semble alors plus avantageux, la tentation du 5 étant alléchante. Le choix C semble également risqué, aucun des joueurs ne veut être le sot qui se retrouve avec 0 : mieux vaut choisir D si l'adversaire décide de se défiler. Chaque joueur sait que l'autre

[1] Pour des raisons de simplicité, les équations présentées considéreront des dilemmes du prisonnier symétriques. Les gains du joueur en rangée et de celui en colonne ne seront donc pas distingués. Les équations demeurent toutefois justes pour des jeux non symétriques à quelques adaptations près et la force du dilemme demeure la même pour un jeu non symétrique.

[2] Nous faisons référence ici au concept d'optimalité de Pareto qui est atteinte lorsqu'il est impossible d'améliorer la situation d'un des joueurs sans empirer celle d'un autre.

est rationnel et peut lui prêter le même raisonnement : le résultat (D,D) semble inévitable et chaque joueur recevra 1.

Comme mentionné plus haut, peu importe le choix de l'adversaire, le choix D est plus payant pour chacun des joueurs. S'ils arrivent à la même conclusion, ce à quoi on pourrait s'attendre avec des joueurs rationnels, ils obtiendront chacun 1. Ce résultat ne semble toutefois pas satisfaisant étant donné qu'ils auraient pu recevoir 3 en coopérant. Le dilemme semble donc le suivant : risquer la coopération ou se laisser tenter par la défection?

Bien qu'il soit plus avantageux d'un point de vue égoïste de jouer D, ce même égoïsme devrait nous porter à jouer C. Or, il semble que les joueurs doivent piler sur leur rationalité ou sur leur égoïsme afin d'atteindre cette solution, mais certains soutiennent que cette solution est possible. Qu'en est-il?

## Itération des DP

Une façon d'aborder le problème de la coopération dans le dilemme du prisonnier est de regarder des cas où le jeu est itéré. S'il semble logique de choisir D lorsqu'un DP est joué une seule fois afin d'éviter le gain *S*, cela n'est pas aussi évident pour les DP itérés (DPI). En effet, recevoir plusieurs fois le gain *R* en initiant une série de coopération réciproque est une solution plus avantageuse que de recevoir le gain *T*, ce qui sera vraisemblablement suivi par une série de gains *P*, l'adversaire réagissant de façon rationnelle. D'autre part, le risque de recevoir le gain *S* peut sembler moins grand puisqu'il est toujours possible de réajuster son choix face à un adversaire refusant de coopérer et ainsi obtenir plusieurs gains *P* par la suite. En d'autres mots, jouer C peut être récompensé aux tours suivants par la coopération de l'autre joueur alors que jouer D peut être puni par sa défection subséquente (Kuhn 2009). Le choix C semble donc moins risqué tout en étant potentiellement payant avec les répétitions et il semble que ce soit un choix rationnel dans le cas des DPI.

### *Itération finie*

La solution n'est toutefois pas aussi claire et cela dépend en partie des paramètres de répétition du DPI. Lorsque le DPI est joué pendant un nombre fini de tours, un joueur rationnel peut faire la réflexion ci-dessus dans un premier temps et se dire qu'il coopèrera à chaque tour (si son adversaire coopère) afin d'obtenir un maximum de gains *R*. Ce type de DPI est toutefois soumis au problème de l'induction rétrograde. Au dernier tour, le joueur se retrouve devant une situation identique à un DP à un seul tour : il peut se permettre de jouer D sans nuire à ses gains futurs. Il sait toutefois que son adversaire est rationnel et qu'il a probablement songé lui aussi à jouer D au dernier tour. Puisque son adversaire jouera vraisemblablement D au dernier tour, il est inutile de jouer C à l'avant-dernier tour : il peut donc jouer D à ce tour sans compromettre une éventuelle coopération dans le futur. L'autre joueur ayant pu penser à la même chose, ils joueront donc tous les deux D à l'avant-dernier tour. Il est donc inutile de coopérer à l'avant-avant-dernier tour et il peut jouer D pour aller chercher le gain *T*. L'argument peut se poursuivre ainsi jusqu'au tout premier tour et on peut en conclure que les deux joueurs choisiront D à tous les tours.

Cet argument de l'induction rétrograde n'est toutefois pas définitif et il ne semble pas s'appliquer lorsque le nombre de tours prévu est suffisamment grand. En effet, pour un nombre de tours donné, il existe un tour où un seuil est atteint. Avant ce seuil, il est plus avantageux de miser sur la coopération. Toutefois, à partir de ce seuil, l'argument s'applique et il est à prévoir que les deux joueurs commenceront à jouer D (Kuhn 2009). La valeur attendue de C et de D varie donc au fil des tours étant donné que l'adversaire réajustera sa stratégie en conséquence.

### *Itération infinie*

Une autre façon d'éviter l'argument de l'induction rétrograde est de jouer le DPI pendant un nombre de tours infini. Puisqu'il n'existe pas de dernier tour, l'induction rétrograde ne peut pas s'amorcer. Ce type de modélisation nous semble toutefois peu intéressant pour deux raisons. Tout d'abord, il est très peu réaliste et il est peu probable que cette modélisation puisse s'appliquer à des situations concrètes, même les ordinateurs possédant une capacité de calcul finie. Deuxièmement, un DPI infini admet par définition des gains infinis. L'utilité attendue de ce type de jeu est donc elle aussi infinie. Bien que certains soutiennent qu'il soit toujours possible d'évaluer la valeur d'une stratégie en évaluant la moyenne des gains (Binmore 1998), cette moyenne est peu utile en comparaison de gains infinis et l'apport de ce type de modèle est restreint à des considérations théoriques.

### *Itération indéfinie*

Une troisième façon d'envisager les DPI est de jouer pendant un nombre de tours indéfini. Le dernier tour n'étant pas connu d'avance, il est impossible d'initier l'induction rétrograde. De plus, le nombre de tours n'étant pas infini, la valeur de chacun est potentiellement significative pour le joueur.

Une façon de faire est d'établir une probabilité $p$ que le jeu continue à chaque tour. La stratégie à adopter dépend alors de cette probabilité. Si $p$ tend vers 1, il est fort probable que le jeu continue et il est avantageux d'établir une coopération réciproque entre les deux joueurs. Toutefois, si $p$ tend vers 0, chaque tour ressemble davantage à un DP à un seul tour puisqu'il est peu probable que le jeu continue. Il devient alors plus rationnel de jouer D comme dans le cas d'un DP à un seul tour.

Les DPI semblent ouvrir la porte à ce qu'il soit rationnel de coopérer dans un dilemme du prisonnier. Nous étudierons cette question dans la prochaine section.

## La coopération dans le DP

Les données empiriques semblent montrer une plus grande occurrence de la coopération que ce que suggère une analyse formelle du DP, et ce, peu importe qu'il s'agisse d'un DP à un seul tour ou d'un DPI (Colman 2003). Ces résultats peuvent être surprenants étant donné que D est le seul choix rationnel.

### *Consultation et coercition*

Les deux solutions qui viennent habituellement le plus rapidement à l'esprit sont celles de la consultation et de la coercition. Celles-ci ne sont toutefois pas satisfaisantes.

Lorsqu'une personne entend parler du DP pour la première fois, il suffit habituellement de lui présenter la version classique de Tucker (Luce and Raiffa 1989) pour qu'elle résolve le dilemme : les deux prisonniers se tairont puisqu'ils ont sûrement fait un pacte avant. L'idée d'une consultation semble effectivement faciliter la coopération, que ce soit avant ou même pendant la période de réflexion précédant la prise de décision. Toutefois, en l'absence de « lien d'intérêt » entre les deux joueurs, le dilemme est simplement reformulé : tenir parole ou ne pas tenir parole? La consultation ne peut pas en elle-même faire en sorte que les deux joueurs coopéreront puisqu'il est plus avantageux pour chacun de tromper l'autre.

Suite à ce contre-argument, notre auditeur propose souvent la solution suivante : les deux prisonniers se tairont sans quoi un mafieux les attendra pour leur casser les jambes. L'idée d'une force de coercition externe semble effectivement un moyen efficace de parvenir à la coopération, que ce soit le Léviathan de Hobbes, une instance juridique faisant respecter les contrats ou une forme moins noble d'autorité. Toutefois, la présence d'une autorité a pour effet de dénaturer le DP en ajoutant un coût supplémentaire à la défection. Ceci a pour effet de changer l'ordre des gains : le gain *R* se retrouve alors à être supérieur aux trois autres. La situation a alors changé et nous ne sommes plus dans un DP, les gains ou utilités attendues ne correspondant plus à l'équation DP1.

Parallèlement, la présence d'un « lien d'intérêt » entre les deux joueurs (affectif, moral, économique, etc) dénature le DP. Le fait de ne pas vouloir décevoir l'autre, de ne pas vouloir le pénaliser ou simplement de faire un choix moralement supérieur ajoute une valeur au gain attendu de la coopération. De même, la culpabilité d'avoir trompé l'autre diminue la valeur du gain attendu de la défection. Ces changements à l'utilité changent également l'ordre de préférence des gains, ce qui ne correspond plus à l'équation DP1.

Le fait de permettre la communication entre des joueurs neutres ne semble donc pas satisfaisant pour que des joueurs coopèrent. L'ajout d'un élément externe comme une autorité politique, un principe moral ou un lien affectif entre les joueurs modifie l'utilité attendue des différents choix de sorte que les joueurs ne se retrouvent plus devant un DP. La coopération dans de telles situations ne peut donc pas être interprétée comme de la coopération dans un dilemme du prisonnier.

### *Types et signaux*

Certains auteurs ont proposé que la reconnaissance des types serait une façon de permettre la coopération. Par exemple, Howard (1988) propose un tournoi se rapprochant d'un DP à un tour où les programmes informatiques ne pourraient pas reconnaître les individus, par exemple en ne pouvant pas se rappeler des interactions passées. Toutefois, ces programmes pourraient percevoir comment leurs adversaires pensent s'ils peuvent reconnaître des types et se rappeler des résultats des interactions avec chaque type. Chaque programme pourrait donc « lire » le programme de son adversaire et décider de la meilleure stratégie. Toutefois, la meilleure stratégie demeure D puisque peu importe ce que jouera l'adversaire, il vaut mieux jouer D. Par contre, si le programme reconnaît que son adversaire est du même type que lui, il jouera C.

Dans un cas où le joueur peut reconnaître des individus, Rapoport propose que la réputation aurait un effet sur la coopération (Rapoport *et al.* 1995). Ce qu'il propose pourrait également être appliqué aux types de Howard, certains types acquérant une certaine réputation. Selon Rapoport, les humains peuvent facilement utiliser de l'information plutôt que des expériences passées. Ainsi, la réputation d'un joueur peut aider un autre joueur à faire le meilleur choix. Inversement, chaque joueur tient à préserver sa réputation afin d'inciter les autres joueurs à coopérer avec lui. Dans l'optique où les joueurs ont de l'information sur les interactions passées d'un joueur, même s'il s'agit d'une interaction unique entre deux joueurs, ils peuvent prendre une décision selon la réputation de leur adversaire. Ainsi, jouer C s'accompagne d'un « bonus », soit un signal indiquant aux autres joueurs qu'il est sécuritaire de coopérer avec lui. Inversement, jouer D s'accompagne d'un signal négatif indiquant aux autres joueurs qu'il est risqué de coopérer lui. Les joueurs tiennent à préserver leur réputation afin de faciliter les interactions (C,C) par rapport aux interactions (D,D). La tentation de jouer D face à un joueur coopératif est diminuée par la tache que cela ferait à la réputation.

Toutefois, ce « signal » change l'utilité de chaque choix. Ainsi, jouer C ajoute une certaine utilité « montrer un bon dossier » aux gains $R$ et $S$; jouer D ajoute une valeur négative « mettre une tache sur le dossier » aux gains $T$ et $P$. Cette utilité peut être associée à la probabilité que les prochains adversaires décident de coopérer ou de se défiler lors des prochaines interactions. Ceci est analogue à ce dont nous avons discuté à propos d'une autorité extérieure à la section précédente à l'exception que cela est ici autorégulé. Ainsi, de nouveau, nous n'avons pas de la coopération dans un DP puisque la valeur des quatre gains possibles est modifiée de telle sorte que les joueurs ne se retrouvent plus devant un DP.

## Coopération dans les jeux itérés

Comme le font remarquer Luce et Raiffa (1989), l'analyse des DPI suppose une hypothèse implicite, soit que l'utilité d'une séquence de gains est égale à la somme des utilités des gains à chaque tour. Or, la théorie de l'utilité ne suppose rien de la sorte. Par exemple, Poundstone (1992) mentionne des expériences où des sujets étaient soumis à un DPI. Les gains étaient répartis comme à la figure 2 et les chiffres représentaient un montant gagné. Or, les joueurs ne jouaient pas pour maximiser le montant total gagné mais pour gagner plus que l'autre joueur. Ils préféraient jouer D et « gagner » le jeu en ayant plus de gains que leur adversaire plutôt que coopérer, ce qui leur aurait conféré un plus grand montant d'argent à la fin de l'expérience. Dans une telle situation, avoir plus que l'autre joueur (l'utilité de la séquence) était préféré à remporter un plus grand montant d'argent (le total de l'utilité de chaque tour). C'est un aspect auquel on doit prêter attention lorsqu'on considère des résultats, particulièrement avec des joueurs humains.

Calculer l'utilité de la séquence comme étant le total des utilités à chacun des tours est l'hypothèse habituellement utilisée pour discuter des DPI. Par exemple, dans des tournois avec des agents informatiques tels que celui d'Axelrod (2006), le total des gains accumulés à chaque tour est réellement l'aspect que l'on cherche à maximiser afin de remporter le tournoi.

Les tournois d'Axelrod sont particulièrement connus dans la littérature et leurs résultats montrent que les programmes facilitant la coopération performent mieux que les autres. Axelrod (1980) identifie trois propriétés[3] qui sont communes aux meilleurs programmes. La propriété la plus importante qui explique le succès d'une stratégie est la gentillesse, soit la propriété de ne pas être le premier à jouer D (à l'exception des quelques derniers tours). Parmi les programmes gentils, la seconde propriété est la

[3] Kuhn indique une quatrième propriété, la clarté, mais nous ne l'avons pas retrouvée dans les articles d'Axelrod consultés.

capacité de faire des représailles, c'est-à-dire à réagir à la défection des autres programmes. La troisième propriété est la capacité à pardonner, c'est-à-dire de pouvoir coopérer de nouveau malgré une défection de l'adversaire. Ces trois propriétés réunies donnent donc des programmes efficaces pour atteindre la coopération sans qu'ils ne se laissent exploiter par des programmes utilisant la défection.

On peut toutefois se demander comment ces programmes réagiraient dans d'autres contextes. Comme le mentionne Axelrod, « We have seen that the nice rules did well in the tournament largely because they did so well with each other, and because there were enough of them to substantially raise each other's average score. » (Axelrod 1980: 14) Si leur environnement était différent et si les programmes « gentils » étaient plus rares, réussiraient-ils à bien se classer? Plusieurs variations ont été testées et plusieurs travaux en théorie des jeux évolutive se sont penchés sur la question mais nous n'aborderons pas cet aspect ici.

Toutefois, nous devons mentionner que dans les DPI, le point d'intérêt n'est plus tant le choix des joueurs à un tour donné que leur stratégie, soit l'ensemble des choix effectués. Un exemple de stratégie pourrait être le suivante : jouer C tant que l'adversaire joue C. Jouer D dès que l'adversaire joue D et pour tous les tours suivants. Parmi toutes les séquences possibles, une telle stratégie élimine des stratégies dominées comme celles où l'on continue à jouer C malgré que l'adversaire ait commencé à jouer D.

Parler de stratégie plutôt que de choix a également l'effet que la notion d'équilibre n'est plus nécessairement équivalente entre un DP et un DPI. Deux stratégies forment un équilibre lorsqu'elles sont la meilleure réponse l'une à l'autre à tous les tours du jeu (Kuhn 2009). Le point d'équilibre d'un jeu en forme normale (c'est-à-dire tel que présenté ici avec les matrices) n'est pas nécessairement le même que pour le même jeu en forme extensive (c'est-à-dire présenté sous forme d'arbre développé pour tous les choix à tous les tours) (Selten 1975). Il peut donc être rationnel pour un joueur de choisir la coopération étant donné que cette stratégie peut être en équilibre avec celle de son adversaire. Si ces deux stratégies permettent également d'atteindre un résultat optimal, il est donc possible de répondre à la fois à l'optimalité et à l'équilibre malgré que le jeu soit sous la forme d'un DP.

Il semble donc rationnel de jouer C dans des DPI, peut-être à l'exception des derniers tours. Toutefois, les explications mises de l'avant semblent sujettes à la même critique quant à la valeur de l'utilité attendue de chaque choix. Selten (1975) soutient que les jeux itérés doivent être analysés sous forme extensive plutôt que sous forme normale. Or, les arbres de la forme extensive sont formés à partir des valeurs de la matrice de la forme normale. Ces deux façons de représenter un jeu ne sont donc pas indépendantes et ne restent que deux représentations du même jeu. Toutefois, nous avons l'impression que la forme extensive permet peut-être de mettre en évidence quelque chose qui est implicite sous la forme normale mais qui est probablement formalisable. Sur un arbre, à chaque point, nous repérons des stratégies dominées et nous les éliminons, ce qui élimine toutes les branches suivant ce nœud. De même, en supposant l'adversaire rationnel, nous pouvons éliminer les branches des stratégies dominées de l'adversaire.

| | **C** | **D** |
|---|---|---|
| **C** | *3, 3* | *1, 4* |
| **D** | *4, 1* | *2, 2* |

Figure 3 : Matrice des gains d'un DP dont la différence est uniforme.

Par exemple, prenons un DPI à quatre tours. Voici quatre branches de ce jeu ainsi que le total des points pour le joueur 1 en utilisant les valeurs de la figure 3 dont la différence entre les gains a été uniformisée afin de faire ressortir certains aspects lors de l'analyse.

| # | **Déroulement** | **Gains** | **Total** |
|---|---|---|---|
| B1 | (D,C), (D,C), (D,C), (D,C) | T+T+T+T | 16 |
| B2 | (C,C), (C,C), (C,C), (D,C) | R+R+R+T | 13 |
| B3 | (C,C), (C,C), (C,C), (C,C) | R+R+R+R | 12 |
| B4 | (D,C), (D,D), (D,D), (D,D) | T+P+P+P | 10 |
| B5 | (D,D), (D,D), (D,D), (D,D) | P+P+P+P | 8 |
| B6 | (C,D), (D,D), (D,D), (D,D) | S+P+P+P | 7 |

Considérons d'abord le cas où l'adversaire serait prêt à coopérer. Même si la branche B1 offre un gain total intéressant, pour le joueur 1, soit quatre fois *T*, celui-ci peut l'éliminer étant donné que le joueur 2, s'il réagit de façon rationnelle, cessera de coopérer. La branche B4 est un scénario plus réaliste étant donné la défection du joueur 1 au premier tour : le joueur 2 cessera de coopérer. Le joueur 1 aura donc eu le gain *T* au premier tour suivi du gain *P* aux trois tours suivants. À la branche B3, le joueur 1 coopère dès le premier tour. Si nous comparons les branches B3 et B4, le joueur 1 fait 1 point de plus au premier tour en se défilant dans la branche B4. Toutefois, il fait un point de moins aux trois tours suivants puisqu'il a le gain *P* plutôt que *R*. Après les quatre tours, il aura fait 2 points de moins avec la branche B4 par rapport à B3. Si sa première défection survient plus loin comme c'est le cas avec la branche B2, il peut accumuler plus de points qu'avec une défection hâtive. Le joueur 1 peut donc aller chercher jusqu'à 3 points supplémentaires s'il ne se défile pas au premier tour. La défection au premier tour semble donc avoir une utilité négative si on regarde ses répercussions aux tours subséquents.

Dans le cas où l'adversaire voudrait se défiler, la défection au premier tour comme dans la branche B5 permet d'éviter le gain *S*. Toutefois, tenter la coopération au premier tour ne fera perdre qu'un seul point comme on peut le constater en comparant B5 avec la branche B6. Dans ce cas, tester la coopération au premier tour semble donc peu risqué, particulièrement si on compare aux avantages que cela peut engendrer.

Encore une fois, il semble que la valeur de l'utilité doive être réajustée dans la matrice. La valeur d'un gain semble flexible en fonction du tour où l'on se situe. Ainsi, choisir la coopération est bel et bien un comportement rationnel, mais de nouveau, nous ne sommes plus dans un DP : l'ordre des gains est désormais modifié.

## Conclusion

Ce bref survol de la littérature concernant la coopération dans un DP ne permet pas de dégager des voies de solution claires. La plupart des solutions proposées pour expliquer qu'il est rationnel de coopérer semblent se buter au même problème, soit des utilités implicites qui ne sont pas inscrites dans la matrice des gains mais qui transforment la nature du jeu, de sorte que l'on ne se retrouve plus devant un DP. Cela semble particulièrement clair dans le cas des DPI si l'on suppose des joueurs rationnels : tout manquement au « pacte » de la coopération sera puni par des défections de l'adversaire. Les gains subséquents tendront vraisemblablement vers *P* plutôt que *R*. Cette baisse des gains attendus aux tours suivants diminue ou relativise la valeur du gain *T* que le joueur cherchait à obtenir avec la défection. La coopération est donc un choix rationnel puisqu'il maximise l'utilité attendue pour les tours futurs. Toutefois, il semble que nous ne sommes plus dans un DP.

L'état de nos recherches actuelles nous amènent à conclure qu'il n'est pas possible de coopérer dans un dilemme du prisonnier, mais qu'il est plutôt possible de sortir du dilemme du prisonnier pour coopérer.

## Bibliographie


Axelrod, R. (1980). "Effective choice in the prisoner's dilemma". *Journal of Conflict Resolution* **24**(1): 3-25.

--- (2006). *The Evolution of Cooperation*. Cambridge, Basic Books.

Binmore, K. G. (1998). *Game Theory and the Social Contract: Just Playing*, The MIT Press.

Colman, A. M. (2003). "Cooperation, psychological game theory, and limitations of rationality in social interaction". *Behavioral and brain sciences* **26**: 139-198.

Howard, J. V. (1988). "Cooperation in the Prisoner's Dilemma". *Theory and Decision* **24**(3): 203-213.

Kuhn, S. T. (2009). Prisonner's Dilemma. Stanford Encyclopedia of Philosophy. E. N. Zalta. Stanford, The Metaphysics Research Lab

Kuhn, S. T. and S. Moresi (1995). "Pure and Utilitarian Prisoner's Dilemmas". *Economics and Philosophy* **11**: 333-352.

Luce, R. D. and H. Raiffa (1989). *Games and decisions: Introduction and critical survey*, Dover Pubns.

Poundstone, W. (1992). *Prisoner's dilemma: John von Neumann, game theory and the puzzle of the bomb*. New York, Anchor Books.

Rapoport, A., A. Diekmann, et al. (1995). "Experiments with social traps IV: reputation effects in the evolution of cooperation". *Rationality and Society* **7**(4): 431.

Selten, R. (1975). "Reexamination of the perfectness concept for equilibrium points in extensive games". *International journal of game theory* **4**(1): 25-55.