

Rapport intelligence artificielle : Dilemme Itéré du Prisonnier

1. Une présentation du problème posé

Le but du projet est de créer un algorithme d'informatique avancé qui simule le dilemme itéré du prisonnier (DIP). Le dilemme du prisonnier classique explore les interactions stratégiques entre deux individus ou groupes d'individus, dans des situations où ils doivent choisir entre trahir ou coopérer. Le DIP étend ce concept en le plaçant dans un contexte répété. Dans le dilemme du prisonnier classique, deux personnes sont arrêtées, elles sont placées dans des cellules séparées et ont deux choix qui s'offrent à elles : coopérer ou duper. Les résultats possibles sont les suivants :

- Les deux trahissent, ils reçoivent tous les deux une peine de prison moyenne,
- L'un trahit et l'autre coopère, celui qui trahit est libéré tandis que l'autre reçoit la peine maximale,
- Les deux coopèrent, ils reçoivent tous les deux une peine de prison minimale.

Dans le dilemme itéré du prisonnier, cette situation est jouée à plusieurs reprises avec les mêmes joueurs/stratégies. À chaque itération, les joueurs peuvent se souvenir (ou non) des choix de leur adversaire. Cela permet des stratégies plus complexes, car les participants peuvent prendre en compte les conséquences à long terme de leurs actions. Des scores sont attribués à chaque joueur en fonction de leur performance.

De plus, le dilemme itéré du prisonnier permet de mettre en lumière les tensions entre intérêts individuels et intérêts communs dans des interactions répétées et d'explorer comment les stratégies évoluent dans de telles situations au fil du temps et quelles sont les meilleures stratégies. Le dilemme du prisonnier est une situation théorique ayant beaucoup d'applications dans la vie réelle. Par exemple, un pays doit-il lever des taxes douanières sur les produits importés de l'autre pays ? Deux entreprises concurrentes doivent-elles essayer de s'entendre pour se partager un marché ou se faire concurrence ? Deux espèces vivant sur un même territoire doivent-elles cohabiter ou se disputer la nourriture disponible ? On souhaite donc développer un programme qui va calculer et sélectionner la meilleure stratégie à ce dilemme. Dans ce programme, on va confronter divers joueurs, et les comparer afin de savoir lequel sera le plus efficace selon des critères bien précis et des environnements variés.

2. Description de notre approche et caractéristiques principales du système intelligent

Notre problème est de créer un système intelligent, capable d'apprendre et d'adapter son comportement face à différentes situations/environnement dans le cadre du DIP. La première chose à faire est de définir chaque stratégies/joueurs par rapport à ses caractéristiques. Certains joueurs sont dotés d'une mémoire à taille variables d'autres non, certains peuvent renoncer en plus de pouvoir coopérer ou duper. Il y a des stratégies plus complexes c'est à dire qu'elles possèdent une taille de mémoire plus importantes et sont plus réactives. Et d'autres stratégies sont qualifiées d'« intelligente » ou apprenantes. Une stratégie pour un joueur est une manière de jouer dans laquelle ses coups sont déterminés par les coups précédents. Une stratégie peut tout à fait obliger un joueur à faire de mauvais coups, cela reste quand même une stratégie.

Quelle stratégie faut-il adopter en fonction du comportement passé de l'adversaire ?

- **Qu'est-ce qu'un système intelligent (SI) ?**

La théorie des jeux possède une branche importante dans l'économie mathématique qui étudie les situations d'interaction stratégique entre des agents économiques supposés rationnels. La rationalité ici fait référence au concept qu'un agent économique est un « homo oeconomicus », c'est à dire intelligent, calculateur, parfaitement informé de tout ce qui l'entoure et doté d'une puissante mémoire. Il n'est donc pas influençable et est constamment à la recherche de ses propres profits. La théorie des jeux s'applique à de nombreux domaines économiques tels que l'oligopole, la négociation, les enchères, la concurrence imparfaite, la tarification, les stratégies d'entreprise, etc. Elle nous permet d'une part de modéliser et d'analyser les interactions entre différents acteurs économiques (entreprises, consommateurs, gouvernements, etc.) en tenant compte de leurs intérêts respectifs et des stratégies qu'ils peuvent adopter. Une fois la modélisation et l'analyse faite, il s'agit de déterminer l'équilibre de Nash, s'il existe ; c'est un concept fondamental dans lequel aucun joueur n'a intérêt à changer unilatéralement de stratégie, étant donné les stratégies choisies par les autres joueurs. Cet équilibre représente une situation stable. C'est ce concept qui a donné naissance au dilemme du prisonnier et à la notion de jeux coopératifs et non coopératifs. L'équilibre de Nash représente une situation où les intérêts individuels ne suivent pas forcément les intérêts de la collectivité. Dans le cas le plus simple c'est à dire à deux joueurs avec seulement 2 choix (trahir ou coopérer), la meilleure stratégie est toujours de **trahir**. On obtient alors l'équilibre suboptimal qui est un équilibre non parfait mais stable c'est ce qu'on appelle équilibre de Nash. Dans les jeux non coopératifs, les joueurs agissent de manière indépendante pour maximiser leurs gains propres tandis que dans les jeux coopératifs, les joueurs peuvent former des coalitions et coordonner leurs actions. Il y a aussi la notion de jeux répétés dans le temps qui ouvrent la voie à des stratégies plus complexes comme la coopération conditionnelle ou des menaces crédibles (l'exemple parfait est celui du dilemme itéré du prisonnier). La théorie des jeux s'applique en intelligence artificielle notamment par l'apprentissage par renforcement multi-agents. Cet apprentissage consiste à faire interagir plusieurs agents dans un même environnement comme dans un jeu où chaque agent cherche à maximiser sa récompense en fonction des actions des autres agents afin d'utiliser des concepts comme l'équilibre de Nash ou la méthode d'induction à rebours pour trouver des stratégies optimales pour les agents. Les techniques de théorie des jeux peuvent également être utilisées pour permettre à des agents de négocier de manière automatique des accords mutuellement avantageux, en modélisant la négociation comme un jeu et en analysant les différentes stratégies possibles.

Dans le domaine des sciences cognitives, l'intelligence est associée à la capacité à raisonner, à apprendre, à stocker des informations dans sa mémoire, à résoudre des problèmes complexes tels que problèmes mathématiques. Mais aussi à savoir s'adapter à de nouvelles situations et à interagir efficacement avec son environnement. Notre intelligence nous permet de prendre des décisions dans des contextes précis. L'intelligence est aussi la capacité à apprendre par observations ; par exemple le Drongo brillant, une espèce d'oiseau dite intelligente car il a appris à imiter le bruit de certains animaux, comme le cri d'alerte des Suricates afin de voler leur nourriture. Pour avoir la confiance des Suricates, le Drongo va la première fois prévenir d'un réel danger. Ensuite, ayant la confiance des Suricates, il va s'en servir pour imiter une seconde fois leur bruit et ainsi les faire fuir alors qu'il n'y a pas de danger et n'aura qu'à prendre dans les trous fait par les suricates les insectes. Le Drongo ne sait pas creuser de trous où les insectes se cachent. Ainsi chaque espèce contribue à sa propre survie en s'adaptant à son environnement et en utilisant une stratégie efficace. Dans le monde des vivants, on peut distinguer l'intelligence individuelle de l'intelligence collective. L'intelligence collective n'est pas la somme des

intelligences individuelles, elle correspond au fait de collaborer collectivement dans un but commun, comme le font les abeilles ou les sociétés humaines. L'objectif d'une espèce intelligente est donc la survie. Dans le domaine de l'intelligence artificielle, un système est généralement considéré comme intelligent s'il est capable de simuler certaines des fonctions cognitives humaines, telles que l'apprentissage, la résolution de problèmes, la reconnaissance de motifs, la perception, le langage naturel et la prise de décision.

En résumé, l'intelligence peut être définie comme la capacité à résoudre des problèmes, à apprendre, à prendre des décisions, à s'adapter et à interagir efficacement avec l'environnement. Un système intelligent, qu'il soit humain ou artificiel, est capable de démontrer certaines de ces capacités de manière adaptative et contextuellement appropriée.

- **Les récompenses :**

Les récompenses varient selon les actions choisies : si les joueurs coopèrent, ils reçoivent une récompense identique (R), s'ils trahissent, ils subissent une punition commune (P). Si l'un coopère et l'autre trahit, celui qui trahit reçoit une grande récompense (T) tandis que l'autre subit une perte (S). L'ordre de priorité des récompenses est défini comme suit : $T \text{ (trahir)} > R \text{ (coopérer)} > P \text{ (punition)} > S \text{ (perte)}$. Pour éviter l'alternance entre coopération et trahison, une inégalité est imposée : $2R > T + S$.

Dans une variante, une troisième action "Renoncer" est introduite, avec un score identique pour les joueurs, noté A. L'ordre des récompenses devient alors $T > R > A > P > S$. Une autre variante permet d'introduire des messages bruités, où un joueur peut mal interpréter l'action de l'adversaire (pensant avoir été dupé alors qu'il a coopéré). Une autre possibilité de variation consiste à modifier le système d'évaluation pour les rencontres impliquant plus de deux joueurs.

- **Les stratégies :**

Nos stratégies reposent sur des ensembles de règles logiques qui définissent le comportement du joueur. Les règles sont généralement de la forme "si X, alors Y". Ces systèmes sont souvent utilisés dans les systèmes intelligents où ils peuvent être utilisés pour prendre des décisions en fonction de certaines conditions.

Le jalon1 réunit les stratégies « simples » c'est à dire sans mémoire ou avec une mémoire de taille 1 maximum. De plus, le comportement de ses stratégies est automatique (**Gently**, gentille, joue toujours la coopération, **Bad**, méchant, trahit toujours, **Fool**, dupe, joue aléatoirement C ou D). Et des stratégies dont le comportement est basé sur les récompenses reçues au tour précédent comme **Pavlov** qui coopère puis joue la même chose qu'au tour précédent si la récompense a été R ou P et qui joue l'opposé si la récompense a été T ou S. Le premier jalon nous a permis de nous familiariser avec la notion d'héritage qui nous permet d'utiliser ou de redéfinir les propriétés et méthodes d'une classe parente ici **Strategy** dans une autre classe enfant et d'y ajouter nos spécificités.

Dans le jalon2, nous avons défini deux types de stratégies plus complexes et plus réactives que celles du jalon1. Il réunit des stratégies dites adaptatives. Une première catégorie de joueurs déterministes c'est à dire que leur comportement est prévisible et répétitif, mais peuvent être efficace dans certains contextes, surtout si les adversaires n'adaptent pas leur comportement en conséquence. La stratégie **Periodic** consiste à répéter de manière cyclique une séquence d'actions définie par l'utilisateur. La stratégie **Majority** analyse le comportement passé de l'adversaire pour déterminer s'il coopère ou

trahit plus souvent. En fonction de cette analyse, elle ajuste son propre comportement en imitant celui de l'adversaire : si l'adversaire coopère plus souvent, elle coopère aussi, et si l'adversaire trahit plus souvent, elle trahit aussi. La stratégie **Gradual** réagit aux trahisons adverses en devenant progressivement plus hostile, en alternant entre coopération et trahison en réponse au comportement de l'adversaire. Une seconde catégorie de joueurs non déterministes avec la stratégie **Markov** qui se base sur des probabilités conditionnelles pour décider de coopérer ou de trahir, en analysant les récompenses passées dans des situations spécifiques. La stratégie **Stochastic** considère un ensemble plus large de probabilités pour adapter son comportement en fonction des récompenses reçues, permettant ainsi une plus grande flexibilité dans ses décisions. Bien que ces stratégies présentent des caractéristiques d'adaptabilité et de prise de décision, elles ne possèdent pas les attributs essentiels de la conscience ou du raisonnement propre à l'intelligence humaine. Par conséquent, il est plus approprié de les considérer comme des agents de décision plutôt que comme des agents "intelligents" au sens humain du terme.

Dans le jalon3, nous avons regroupé les stratégies apprenantes, toutes les stratégies héritent de la classe abstraite 'AbstractLearner' sauf **Automaton** qui hérite de **Strategy**. L'apprentissage fait référence à la capacité du système à améliorer sa performance sur une tâche donnée à partir de l'expérience. Notre système utilise la technique de l'apprentissage par renforcement où le système apprend par essais et erreurs en interagissant avec son environnement. La classe **Automaton** fonctionne en attribuant une séquence de règles à suivre pour prendre des décisions. Ces règles sont définies par une chaîne de caractères fournie lors de la création de l'instance de la classe. La classe **Mime** agit comme un observateur, son comportement consiste à identifier les schémas dans les actions de l'adversaire et à ajuster ses propres actions en conséquence. Elle conserve une mémoire des actions adverses passées et utilise cette information pour prendre des décisions. La classe **Motif** agit également comme un observateur mais elle se concentre sur l'identification des motifs récurrents dans les actions de l'adversaire. Son comportement consiste à rechercher les schémas dans les actions passées de l'adversaire et à ajuster ses propres actions en fonction des motifs identifiés. Elles peuvent être considérées comme intelligentes dans la mesure où elles utilisent des informations passées pour influencer leurs décisions futures.

Le comportement de la classe **Shannon** consiste à utiliser ses mémoires pour prédire les actions futures de l'adversaire. En fonction des informations stockées dans ses mémoires, la classe Shannon peut choisir de maintenir une stratégie ou de modifier son comportement pour s'adapter à l'évolution de la situation. En utilisant plusieurs mémoires et en analysant différents aspects du comportement de l'adversaire, la classe Shannon est capable de prendre des décisions stratégiques plus élaborées et est capable d'apprendre en mettant à jour ses mémoires en fonction des résultats des matchs. Cela lui permet d'ajuster ses prédictions et ses stratégies au fur et à mesure des parties.

- **Quel(s) critères utiliser pour évaluer le SI ?**
 - **Les évaluations :**

Pour évaluer les différentes stratégies, nous avons mis en place différentes approches :

- L'évaluation en rencontre directe, un joueur en affronte un autre lors d'une simulation sur un nombre fini de coups.
- L'évaluation par tournoi : chaque stratégie est confrontée à toutes les autres, y compris à elle-même, puis un classement est établi pour déterminer la meilleure.

- L'évaluation écologiques : les joueurs appartiennent à différentes "espèces" avec un effectif et participent à des tournois. Le nombre de joueurs par espèce évolue à chaque étape.
- L'évaluation écologique spatiale : Les stratégies sont des "habitants" d'une grille discrète où elles interagissent selon des règles spécifiques (miroir, symétrie, ...). La simulation continue jusqu'à ce qu'il n'y ait plus d'évolution dans la population ou jusqu'à ce qu'un nombre maximum d'itérations soit atteint.

Ces confrontations peuvent se dérouler dans un monde "parfait" ou "bruité", où une action perçue peut différer de l'action réelle. Le nombre de stratégies étant important, il est difficile de réaliser chaque évaluation pour chaque stratégie. Nous allons donc poser des hypothèses déterminées par nos critères d'évaluations de l'intelligence du système. Afin d'étudier quelles stratégies est la 'meilleure' et pour déterminer quels sont les critères les plus pertinents pour évaluer l'intelligence du système nous allons confronter nos hypothèses.

• Critères d'évaluations :

Pour évaluer l'intelligence du système dans le cadre du DIP, les critères pertinents pourraient être :

- Marquer le plus de points dans un affrontements direct
- Score moyen obtenues aux différentes évaluations
- Le classement à l'issue du tournoi
- La « survie » de l'espèce dans un environnement écologique
- Ne pas être gêné par du bruit dans telle ou telle situation
- Capacité à prendre des décisions adaptatives en fonction des situations rencontrées
- Capacité à apprendre et à s'adapter aux comportements adverses au fil du temps
- Capacité à maximiser les gains dans différentes conditions, y compris en présence de bruit

Pour mettre en valeur l'aspect intelligent du système et déterminer les meilleures stratégies dans le cadre des évaluations, nous pouvons adopter plusieurs approches :

Capacité d'adaptation : On peut choisir d'évaluer la capacité des stratégies à s'adapter aux changements dans l'environnement de jeu, tels que l'introduction de nouvelles actions ou la présence de bruit dans les communications. Les stratégies qui peuvent ajuster leur comportement en fonction de ces changements peuvent être considérées comme plus intelligentes.

Apprentissage : Nous pouvons aussi mesurer la capacité des stratégies à apprendre des interactions passées avec les adversaires. Les stratégies qui peuvent analyser et tirer des enseignements des résultats des matchs précédents pour améliorer leurs performances futures peuvent être considérées comme plus intelligentes.

Performance globale : Ainsi nous pouvons baser notre analyse sur les résultats aux performances globale des stratégies en termes de score moyen, de classement dans les tournois et de survie dans des environnements écologiques. Les stratégies qui parviennent à maximiser leurs gains dans différentes conditions et à maintenir une position concurrentielle favorable peuvent être considérées comme plus efficaces.

Prévisibilité des adversaires : Nous pouvons aussi évaluer la capacité des stratégies à prévoir les actions probables des adversaires en fonction de leur historique de jeu. Les stratégies qui peuvent anticiper les mouvements adverses et ajuster leur comportement en conséquence peuvent être

considérées comme plus intelligentes. Le compteur **good_guess** nous permet d'établir le nombre de bonne prédiction effectuer par les stratégies.

En utilisant ces approches, nous pouvons identifier les stratégies qui démontrent les comportements les plus intelligents et les plus efficaces dans diverses conditions de jeu. Les stratégies qui peuvent s'adapter, apprendre et anticiper les actions adverses seront susceptibles de se démarquer comme les meilleures dans le cadre du Dilemme du Prisonnier.

Évaluations générales des stratégies du jalon1 :

Pour effectuer une évaluation par rencontre directe, on utilise la méthode **adaptive_evaluation()**, les résultats montrent que pour les stratégies du jalon1, celles qui obtiennent les meilleures scores et qui sont imbattables sont **Bad** et **BadSulky**. On choisit de ne pas détailler ces résultats parce qu'il ne reflète pas le coté intelligent du système. Cependant on choisit de détailler les résultats en tournois car les stratégies les plus fortes à l'un contre un ne le sont pas forcément lors d'un tournoi, comme le montre les résultats suivants.

Nous effectuons un tournoi de tous ces joueurs. Pour cela, on utilise la méthode d'évaluation en tournoi et on choisit les stratégies que l'on souhaite confronter. Le tournoi se fait sur 5 matchs et soit dans un environnement sans bruit soit avec bruit.

Voici les commandes à entrer dans le terminal après avoir lancer le fichier évaluation et avoir défini les stratégies (x = Gentle(), y = Bad(), etc) que l'on souhaite évaluer :

>>tournament((x,y,z,b,c,d,e,f,g),5,100,101) : sans bruit

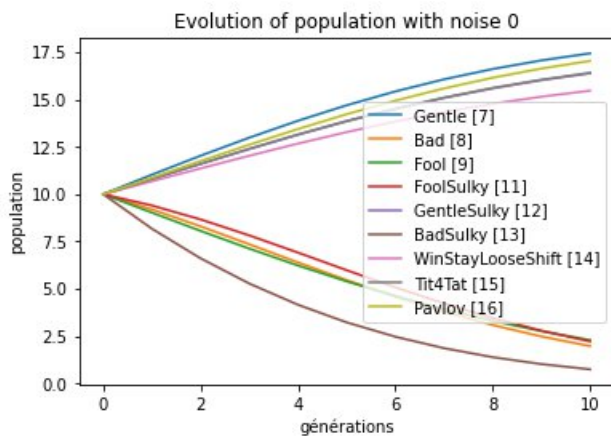
>>tournament((x,y,z,b,c,d,f,g),5,100,101, epsilon = 0.3) : avec bruit

Joueurs	Gentle	Bad	Fool	FoolSulky	GentleSulk y	BadSulk y	WSTS	T4T	Pavlov
Scores sans bruit	37,5	31,2	30,26	32,126	37,45	27,65	35 ,9	37,6	37,3
Score avec Bruit	28.48	31.77	29 ,68	25.63	25.63	25.86	30.31	30.31	30.5

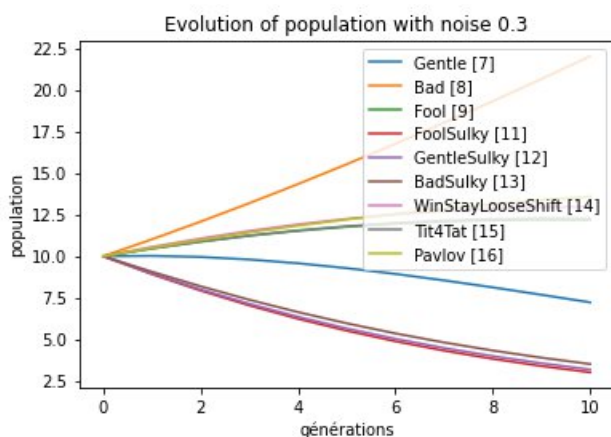
Les résultats montrent que les stratégies qui obtiennent les meilleurs scores dans un environnement non bruité sont **Gentle**, **GentleSulky**, **Pavlov** et **Tit4Tat**. Alors qu'en ajoutant du bruit, ce sont **Bad**, **WinStayLooseShift**, **Tit4Tat** et **Pavlov** qui obtiennent une meilleure place au classement.

• Évaluation écologique des stratégies du jalon1 :

Cette approche s'inspire de la théorie de l'évolution biologique pour résoudre des problèmes complexes. Elle consiste à utiliser des algorithmes évolutionnaires pour évaluer les conséquences à long terme des choix de chaque population de stratégie.



Ce graphique montre l'évolution des différentes populations des joueurs du jalon1. L'évaluation se fait dans un environnement sans bruit. On obtient des résultats similaires à ceux du tournoi, les populations de Gentle, GentleSulky, Pavlov et Tit4Tat augmentent tandis que ceux de Bad, BadSulky, Fool, FoolSulky etc diminuent jusqu'à s'éteindre.



Ce graphique nous montre l'évolution des différentes populations des joueurs du jalon1. L'évaluation se fait dans un environnement bruité (0.3). On obtient des résultats similaires à ceux du tournoi, l'espèce Bad augmente exponentiellement, les populations de Pavlov, Tit4Tat semble stagné et les autres diminuent jusqu'à l'extinction.

- **Capacité d'adaptation :**

Pour évaluer la capacité d'adaptation des différentes stratégies dans le contexte du DIP. Nous allons déterminer quelles stratégies s'adaptent aux changements de conditions dans l'environnement de jeu.

Pour commencer, nous identifions les différents scénarios représentant des situations où de nouveaux éléments sont introduits dans la simulation. Nous pouvons ajouter du bruit plus ou moins élevés à l'environnement ce qui peut entraîner de la confusion chez certains joueurs, nous allons tenter d'identifier lesquels. On peut aussi ajouter une troisième action : celle de pouvoir Renoncer (model = m2).

Nous sélectionnons les stratégies que nous souhaitons évaluer, en privilégiant les stratégies apprenantes du jalon3 que nous allons modéliser à partir des stratégies du jalon1, ainsi que celles adaptatives du jalon2. Par exemple, nous pourrions envisager de représenter la stratégie **Mime** en la dérivant de la stratégie **Gentle()**.

- **1er Scénario :**

On choisit les stratégies qui ont obtenues les meilleurs résultats aux derniers tournois effectué (avec ou sans bruit). On commence par dériver toutes ces stratégies avec Mime : (Gentle, GentleSulky, Pavlov, Tit4Tat, Bad et WinStayLooseShift). Le choix d'utiliser la dérivation des stratégies à partir de Mime pour évaluer leur performance dans différents contextes permet de comparer les performances des différentes stratégies dans des conditions équivalentes.

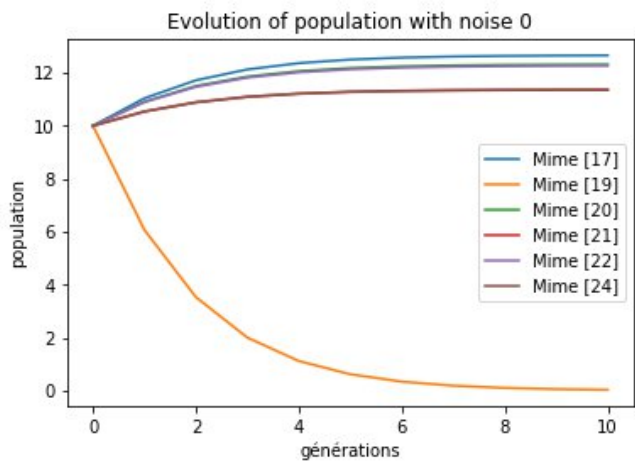
On lance les tournois avec ou sans bruit entre ces différents joueurs à partir de ces commandes :


```
>> tournament((x,y,z,w,u,v), 5, 100, 101,model = m2) : sans bruit
>> tournament((x,y,z,w,u,v), 5, 100, 101,model = m2, epsilon = 0.3) : avec bruit
```

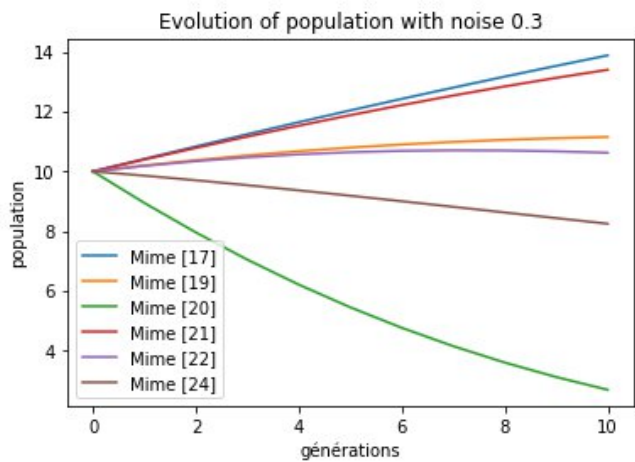
Joueurs	Mime(Gentle)	Mime(Bad)	Mime(GentleSulky)	Mime(Pavlov)	Mime(WSL)	Mime(T4T)
Score sans bruit	28.458	15.39	27.98	27.00	27.72	27.00
Score avec bruit	20.5	20.32	18.04	20.27	20.37	20.66

Dans un environnement sans bruit, les stratégies qui obtiennent les meilleures résultats sont Mime(gentle), Mime(GentleSulky) et Mime(WinStayLooseShift). Tandis que dans un environnement bruité, ce sont Mime(Bad), Mime(Tit4Tat) et Mime(WinStayLooseShift). On remarque que Mime(Bad) s’adapte à son nouvel environnement et améliore ses performances. Regardons ce qui se passe lorsque l’on effectue une évaluation écologique avec la méthode ‘eco_evol()’. Les commandes à lancer dans le terminal sont :

```
>> eco_evol(10, (x,y,z,w,u,v)) : sans bruit
>> eco_evol(10, (x,y,z,w,u,v), epsilon =0.3) : avec bruit
```



Ce graphique indique que seul la population de Mime(Bad) décroît jusqu’à disparaître totalement de l’environnement . Tandis que les autres populations de joueurs atteignent un palier et se stabilisent à un certain niveau.



Dans un environnement bruité, l’évolution est totalement différente. On peut voir que Mime(Gentle) et Mime(Pavlov) augmentent de façon linéaire. Mime(GentleSulky) ne s’adapte pas au bruit et décroît linéairement jusqu’à l’extinction, tandis que les autres populations ne montrent pas d’évolutions significatives.

Les joueurs ont des résultats différents en fonction du bruit ajouté. Les résultats présentés montrent que les performances varient selon la présence ou l'absence de bruit, ce qui confirme l'importance de l'adaptabilité des stratégies. On suppose donc que certains joueurs s'adaptent mieux à ces conditions comme Mime(Gentle), Mime(Bad), Mime(Pavlov), Mime(Tit4Tat), Mime(WinStayLooseShift).

Certaines stratégies, comme Mime(Bad) ou Mime(Tit4Tat), sont capables de s'adapter rapidement aux perturbations causées par le bruit. Elles reconnaissent et ajustent leur comportement en réponse aux signaux bruités, ce qui les avantage dans des environnements dynamiques.

D'autres stratégies, comme Mime(Pavlov) ou Mime(Tit4Tat), intègrent des mécanismes d'apprentissage ou de mémorisation. Elles apprennent des interactions passées et ajustent leur comportement en conséquence, ce qui les rend efficaces face aux perturbations dans l'environnement bruité.

Certaines stratégies, comme Mime(WinStayLooseShift), réagissent de manière dynamique aux actions des adversaires. Elles détectent les changements dans le comportement des adversaires et modifient leur propre stratégie en conséquence pour optimiser les gains, ce qui les rend adaptées aux environnements bruités.

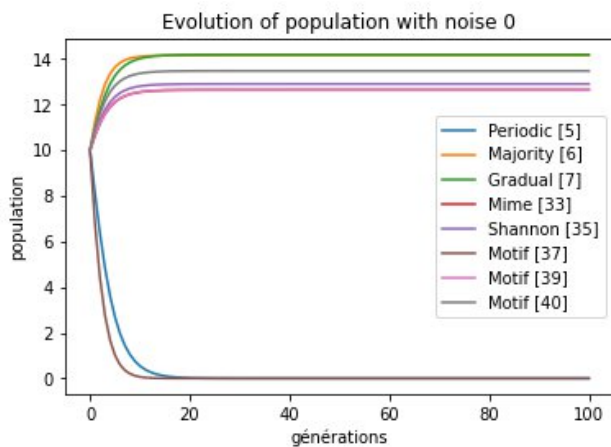
2ème Scénario :

On choisit de sélectionner d'autres joueurs. Nous allons évaluer tous les joueurs du jalon2 ainsi que ceux du jalons 3 dérivés des stratégies que l'on juge les plus pertinentes du jalon1 d'après notre précédente analyse ; c'est à dire Gentle, Bad, Tit4Tat, WinStayLooseShift et Pavlov.

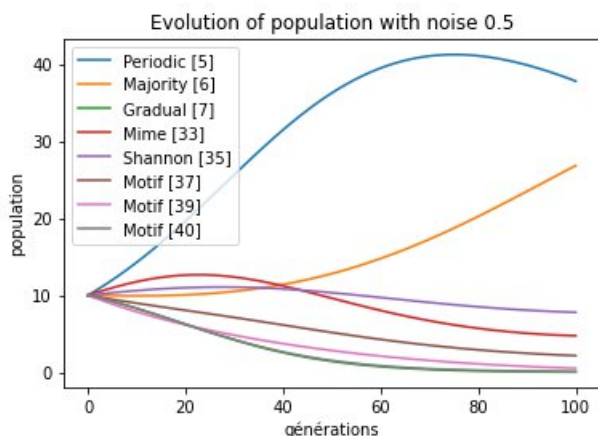
Joueurs	Periodic ('DC')	Majority	Gradual (True,1,3, M2)	Mime (Pavlov)	Shannon (Gentle)	Motif (Bad)	Motif (T4T)	Motif (WSLS)
Score sans bruit	26.61	37.47	36.5	35.488	35.558	21.632	35.492	36.23
Score avec bruit	28.364	28.267	27.889	28.229	27.918	27.736	28.132	27.832

On remarque que Periodic et Motif(Bad) voient leur performance augmenter dans un environnement bruité. Tandis que les autres joueurs ont de moins bon résultat causé par le bruit. On peut en déduire que ces joueurs rencontrent des difficultés face à cette perturbation.

On remarque aussi que dans l'environnement bruité toutes les stratégies ont des scores à peu près équivalents. La différence de performance selon le score moyen obtenu à l'issue du tournoi est plus marqué dans un environnement sans bruit.



Ce graphique montre que dans un milieu non bruité, les populations de Periodic et Motif(Bad) décroissent jusqu'à s'éteindre avant la 20^{ème} génération. Les autres populations augmentent jusqu'au seuil de 14 environ avant de se stabiliser.



Dans l'environnement bruité, on observe une croissance rapide de la population de la stratégie Periodic, suivie toutefois d'un léger déclin. La stratégie Majority, quant à elle, semble particulièrement bien s'acclimater aux conditions bruitées, sa population affichant une tendance haussière soutenue. En revanche, les autres stratégies peinent à s'adapter et voient leurs effectifs se réduire progressivement.

Dans ce deuxième scénario, la capacité à s'adapter à un environnement bruité se révèle cruciale. Les stratégies Periodic et Motif (Bad) tirent leur épingle du jeu dans ce contexte, tandis que les autres montrent des difficultés face au bruit. Dans un environnement bruité, toutes les stratégies affichent des performances similaires, mais les différences entre elles sont plus prononcées en l'absence de bruit. Les résultats de l'évaluation écologique confirment cette tendance, montrant des dynamiques de population distinctes selon l'environnement.

- **Une conclusion critique de votre travail, et une éventuelle ouverture**

Il convient de noter que toutes les méthodes d'évaluation proposées n'ont pas été entièrement utilisées, notamment les tournois direct impliquant une interaction avec une stratégie humaine (Human). Cette évaluation n'a pu être réalisée en raison des contraintes liées à la saisie manuelle des actions à chaque tour, ce qui aurait été long... De plus, cette analyse est concentrée principalement sur le critère clé de la capacité d'adaptation aux changements d'environnement, ce qui n'est pas toujours le cas de façon générale comme vu dans la partie définition sur l'intelligence artificielle. Mais ce point est très important pour nous dans la définition de l'intelligence, d'où ce choix.

A travers des méthodes variées telles que les rencontres directes, les tournois et les évaluations écologiques, nous avons exploré les capacités des stratégies à s'adapter à différentes conditions de jeu, y compris en présence de bruit. La capacité d'adaptation s'est révélée cruciale, avec des stratégies telles que Mime(Bad) et Periodic dont le comportement s'ajuste en fonction des perturbations dans l'environnement de jeu. De plus, on remarque que les stratégies intégrant des mécanismes d'apprentissage (Mime(Pavlov) et Mime(Tit4Tat)) ont montré une capacité à apprendre des interactions passées pour améliorer leurs performances futures. Certaines stratégies ont mieux performé dans des

environnements bruités, tandis que d'autres ont montré des difficultés à maintenir leur efficacité dans de telles conditions.

En comparant les stratégies du dilemme du prisonnier aux intelligences artificielles existantes comme ChatGPT ou DALL·E, on peut voir des points communs et des différences intéressantes. Tout d'abord, les intelligences artificielles comme ChatGPT et DALL·E sont conçues pour résoudre des problèmes complexes en utilisant des modèles d'apprentissage profond et des algorithmes avancés. Elles sont capables de comprendre et de générer du texte ou des images de haute qualité en fonction des données qu'elles ont analysées. De la même manière, certaines stratégies du dilemme du prisonnier, comme Mime(Pavlov) ou Mime(Tit4Tat) intègrent des mécanismes d'apprentissage pour s'adapter aux comportements des adversaires et améliorer leurs performances. Cependant, les intelligences artificielles comme ChatGPT et DALL·E sont souvent utilisées dans des contextes sans compétition directe. Leur objectif principal est de produire des résultats utiles ou créatifs pour ceux qui les utilisent. De plus, ces IA sont généralement évaluées en fonction de critères tels que la qualité du texte généré ou la fidélité visuelle des images produites. Tandis que l'évaluation des stratégies du dilemme du prisonnier se concentre sur des aspects tels que l'adaptabilité aux changements de règles, la capacité à apprendre des erreurs passées et à anticiper les actions des adversaires. En comparant ces différents types d'intelligences artificielles, on comprend les forces et les limites de chaque approche afin de développer des systèmes plus performants et plus polyvalents dans le futur.

En conclusion, notre évaluation des stratégies dans le DIP met en évidence la diversité des approches utilisées par les agents décisionnels pour maximiser leurs gains dans un environnement compétitif et incertain. Les stratégies les plus intelligentes sont celles qui peuvent s'adapter rapidement aux changements, apprendre des interactions passées et anticiper les actions probables des adversaires. En continuant à explorer ces aspects, nous pouvons affiner notre compréhension de l'intelligence artificielle et développer des systèmes de prise de décision plus efficaces et robustes dans une variété de contextes.

- **Annexe :**

Sitographie :

https://www.iro.umontreal.ca/~aimeur/cours/ift6261/cours_2016/Cours_IFT6261-Intro-H16.pdf

<https://www.ibisc.univ-evry.fr/~hutzler/Documents/These.pdf>

<https://www.cril.univ-artois.fr/~konieczny/enseignement/DIP.pdf>

<https://ecopsycho.gretha.cnrs.fr>