

Comportement individuel et intelligence collective

Les bases expérimentales de PRISON

Bruno BEAUFILS

L'ÉQUIPE SMAC s'intéresse aux Systèmes Multi-Agents et à la Coopération. Elle comporte trois principaux projets de recherche autour de ces thèmes. L'un est plus particulièrement dédié à l'étude de la structure des systèmes multi agents (MAGIQUE), un s'intéresse à la négociation entre agents (ANTS), et enfin un troisième est consacré à l'étude de la coopération et de l'évolution de la coopération entre agents (PRISON). Les intersections entre ces projets ne sont pas vides.

Les débuts de SMAC

HISTORIQUEMENT, les thèmes abordés par le projet PRISON ont été les premiers à être approchés par des membres de l'équipe, suite à un article de Jean-Paul DELAHAYE, paru dans la revue POUR LA SCIENCE, consacré au Dilemme Itéré du Prisonnier.

À cette époque, les futurs membres de l'équipe SMAC appartenaient à l'équipe METHEOL, et leurs travaux se focalisaient sur les outils logiques pour l'Intelligence Artificielle (IA). L'approche IA ayant montré ses limites dans certaines situations, Jean-Paul DELAHAYE et Philippe MATHIEU se sont intéressés à ce qui a vite été appelée la Vie Artificielle (VA). L'idée de base de cette discipline nouvelle était qu'il ne fallait pas essayer de créer de l'intelligence en tentant de cloner le raisonnement humain, *i.e.* un gros agent très complexe; mais qu'il fallait plutôt essayer de voir comment l'intelligence pouvait émerger dans un système de petites entités, au comportement simple, interagissant non seulement entre elles, mais aussi avec leur environnement¹. Tandis que l'IA est une tentative de copie à un moment donné, la VA essaie plutôt de comprendre la manière dont l'intelligence peut émerger, via les mécanismes d'évolutions, et d'interactions dans des groupes d'individus.

De nombreux modèles correspondants à la VA ont été étudiés dans la littérature comme par exemple le jeu de la vie, les modèles proie-prédateur, ou le dilemme itéré du prisonnier. Ce dernier est devenu le modèle de VA choisi comme

support de recherche pour le futur projet PRISON. La Vie Artificielle, domaine de recherche naissant et très théorique, ne s'intéressait pas tant à la structure des groupes qu'à leur comportement global et leur émergence. Elle apportait pourtant rapidement de nombreux résultats intéressants parmi lesquelles on peut citer les algorithmes évolutifs, dont font partie les algorithmes génétiques, ou les automates cellulaires.

Philippe MATHIEU s'est intéressé aux structures des groupes d'individus devinant qu'elles devaient avoir un rôle important dans les comportements globaux, et dans les facultés d'adaptation de tels systèmes. De ces idées et des travaux qui étaient alors en cours sur la distribution de l'IA, via par exemple le projet BIVOUAC, est née l'idée d'un modèle Multi AGent hierarchIQUE (MAGIQUE). C'est alors tout naturellement que le projet ANTS, concernant les mécanismes de négociations entre agents, est venu se greffer aux deux précédents et que l'équipe SMAC a pu voir le jour de manière cohérente.

Aujourd'hui l'équipe comporte 7 membres, dont 3 permanents; développe 3 projets clairement identifiés; intervient dans les différentes formations informatiques de Lille, et entretient des contacts avec deux entreprises dans le cadre, notamment, de la conduite de stages.

Le Dilemme Itéré du Prisonnier

LE Dilemme Itéré du Prisonnier (DIP) est un modèle issu de la théorie des jeux de VON NEUMANN et MORGENSTERN, qui a été introduit par FLOOD et DRESHER à la RAND Corporation au début des années 1950. Il a été mis en avant avec l'arrivée de la Vie Artificielle, et surtout après la publication du livre *The evolution of cooperation*, de Robert AXELROD au début des années 1980. Ce modèle mathématique, *a priori*, très simple est principalement utilisé par les économistes pour étudier les comportements d'agents rationnels, et par les biologistes pour étudier les dynamiques de populations.

Au sens de la théorie des jeux, le DIP est un jeu à deux joueurs, non-coopératif, simultané, et à somme non-nulle. Chacun des deux joueurs a le choix entre deux possibilités, couramment nommées **Coopération** et **Trahison**. Les deux joueurs déclarent leur choix simultanément, *i.e* sans connaître le choix de l'autre. Si les deux joueurs coopèrent, ils obtiennent alors chacun une *Récompense* de R points, alors que s'ils trahissent tous les deux, ils n'obtiennent qu'une *Punition*, plus faible, de P points. Le dilemme apparaît lorsqu'un des deux joueurs trahit son adversaire qui coopère. Dans ce cas, le traître obtient la récompense de la *Tentation*, T points, alors que l'autre n'a que le *Salaire de la duperie*, S points. Le dilemme est alors caractérisé par l'inéquation $T > R > P > S$, de cette manière la trahison est favorisée par rapport à la coopération. Individuellement les agents ont intérêt à trahir, alors que collectivement ils ont intérêt à coopérer. La distribution classique des points est la suivante : $T = 5$, $R = 3$, $P = 1$, $S = 0$. Si les agents ne se rencontrent qu'une seule fois, la meilleure stratégie est de trahir. Mais s'ils doivent se rencontrer plusieurs fois, il n'y a pas de meilleure stratégie dans l'absolu. Dans ce cas, le dilemme est dit itéré ou répété. On ajoute le respect de l'inéquation $2R > T + S$, de manière à assurer que le gain total est le plus important dans le cas d'une coopération mutuelle.

Désormais les stratégies à la disposition des agents peuvent prendre en compte l'histoire du jeu, c'est à dire les coups précédemment joués. On peut alors dire que chaque agent peut avoir un comportement. En théorie des jeux, on parlera plus volontiers de stratégies. Le nombre de ces comportements est infini ; certains sont déterministes, on les dira *purs*, d'autres probabilistes, on les dira *mixtes*. Voici quelques unes de ces stratégies :

- ♦ **gentille** coopère tout le temps ;
- ♦ **méchante** trahit tout le temps ;
- ♦ **rancunière** coopère jusqu'à ce que son adversaire l'ait trahie, après quoi elle trahit tout le temps ;
- ♦ **per_ccd** joue périodiquement coopérer, coopérer, trahir ;
- ♦ **donnant_donnant** coopère au premier coup, puis joue ce que l'adversaire a joué au coup précédent. Cette stratégie, proposée par A. Rapoport, est considérée depuis la publication des travaux d'AXELROD, comme une des meilleures car elle possède les quatre qualités, qui selon lui sont indis-

pensables à une stratégie pour bien se comporter :

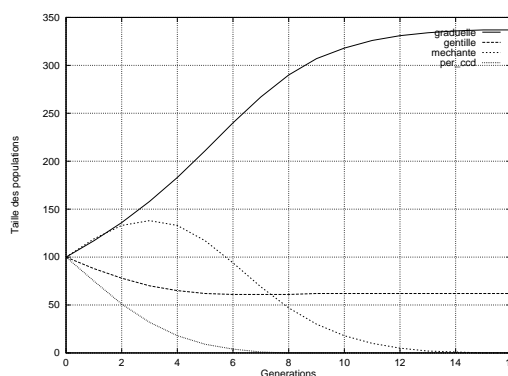
- ① *la gentillesse* : ne jamais commencer à trahir.
- ② *la réactivité* : réagir à une trahison.
- ③ *le pardon* : savoir pardonner un adversaire qui a trahi.
- ④ *la simplicité* : avoir un comportement simple, de façon à pouvoir être compris par son adversaire.

C'est ce dernier point sur lequel nous ne sommes pas d'accord avec AXELROD.

♦ **graduelle**, qui commence par coopérer, puis réagit à la première défection de son adversaire par une punition de un coup, suivi d'une période d'acalmie de deux coups, ..., à la n^e trahison de son adversaire par une punition de n coups, suivi d'une période d'acalmie de 2 coups. Nous avons introduit cette stratégie, de façon à montrer le bien fondé de nos idées sur la simplicité. **graduelle** est meilleure que **donnant_donnant** dans la majorité des cas, et sa complexité est plus importante (elle a besoin d'une mémoire potentiellement infinie).

Deux modes principaux d'évaluation des stratégies sont utilisés :

- ① les **tournois** qui correspondent à des championnats sportifs, le score d'une stratégie étant la somme des scores obtenus par la stratégie lors de toutes les rencontres ;
- ② les **évolutions écologiques** qui correspondent à l'évolution de la distribution des stratégies dans une population d'individus, via une redistribution proportionnelle fonction du score de chaque sous-population. On peut voir un exemple d'une telle évolution sur la figure suivante :



Le problème que pose le DIP est de savoir déterminer la meilleure stratégie à adopter dans

un contexte particulier. Plus généralement on voudrait connaître les traits de caractères importants pour qu'une stratégie se comporte bien, *i.e.* qu'elle ait le plus gros gain possible, dans le plus grand nombre de situations différentes.

Le projet PRISON

POUR atteindre ce but, le projet PRISON, dont le nom vient du premier simulateur écrit au laboratoire, se propose de collecter le plus d'informations possibles autour du dilemme itéré du prisonnier, des stratégies pour celui-ci, et plus généralement des modèles de coopération entre agents.

Une des principales idées défendues par le projet est que contrairement aux idées reçues et issues des travaux d'AXELROD, la simplicité n'est pas un bon critère de qualité pour les stratégies. Nous avons trouvé des dynamiques oscillatoires, quasi chaotiques, sur des populations de stratégies pures ; vérifié la robustesse du modèle aux variations des différents paramètres ; élaboré un protocole d'évaluation de stratégies ; tenté de rechercher automatiquement de bonnes stratégies via des algorithmes évolutifs ou via l'exploration systématique de grands espaces de stratégies (classes complètes de stratégies définies génétiquement) ; mis au point une variante du DIP où la coopération peut se faire sur deux niveaux. Dans ce modèle le changement d'une inéquation permet d'obtenir un nouveau dilemme dans lequel deux niveaux de coopérations sont possibles via une sorte d'entente entre les adversaires. Celle-ci ne peut cependant se décider que par le seul canal de communication disponible : le jeu, *i.e.* les coups joués dans le passé². Dans ce modèle la coopération est beaucoup plus liée à la communication, avec les problèmes que cela implique.

Les résultats de PRISON sont donc plus souvent basés sur l'expérimentation et l'analyse de résultats expérimentaux, plutôt que sur la démonstration formelle de résultats, malheureusement fréquemment impossible avec le DIP.

Un des buts du projet est de devenir une référence en ce qui concerne l'étude de la coopération. À cet effet, nous avons construit un site web, sur lequel nous mettons à disposition de tous les utilisateurs du DIP les différents résultats existants. Nous offrons en particulier nos différents

simulateurs, qui ont été réécrits de nombreuses fois pour coller à nos besoins, et dont les dernières versions sont en JAVA³ et qui permettent à tous de vérifier les expériences sur lesquelles se basent nos résultats. L'idée sous jacente est d'être le plus rigoureux possible dans nos expériences afin que quiconque s'y intéresse puisse aisément vérifier nos affirmations et résultats. Nous offrons également la possibilité de comprendre et de tester le modèle via quelques applets. Tous nos articles, ainsi qu'une bibliographie importante sont également disponibles sur le site, avec une grosse collection de liens sur le sujet.

Le futur

LES travaux en cours sont centrés sur le calcul d'évolution écologique sur un grand nombre de stratégies. Jusqu'à présent, et du fait de la limitation de la puissance de calcul et de stockage des ordinateurs, il était quasiment impossible de faire des simulations impliquant plus de 1000 stratégies. Grâce à l'explosion de la vitesse des processeurs, et au modèle de distribution de calcul récent nous espérons atteindre l'objectif d'un calcul impliquant plus de 10000 stratégies, soit 10 fois plus. Les bénéfices pour PRISON sont évidents car avec un tel nombre de stratégies, il semble clair que les évaluations de stratégies existantes seront plus fiables. Nous espérons donc pouvoir découvrir dans les stratégies ainsi révélées de nouveaux critères de qualité pour les comportements coopératifs. Cette phase de travail est effectuée en collaboration avec le projet MAGIQUE. La distribution du calcul s'effectue via le modèle hiérarchique d'agents de ce projet. Les éventuels résultats de ce travail pourront permettre de valider un peu plus le modèle MAGIQUE.

Pour plus d'informations on pourra consulter les différents sites de SMAC à partir de :

<http://www.lifl.fr/SMAC>

Notes

¹On parle souvent d'approche Haut-Bas (Top-Down) pour l'IA, et Bas vers le Haut (Bottom-Up) pour la VA.

²Le jeu est non-coopératif, c'est à dire que les joueurs ne peuvent communiquer autrement qu'en jouant.

³La politique de l'équipe est que pour une meilleure interopérabilité entre projets, et pour une meilleure diffusion, tous les développements de SMAC sont fait en JAVA.

