

Dynamique de l'interaction

Samuel Delepouille^{1,2*}, Philippe Preux², and Jean-Claude Darcheville¹

¹ Unité de Recherche sur l'Évolution des Comportements et des Apprentissages (URECA), UPRES-EA 1059, Université de Lille 3, B.P. 149, 59653 Villeneuve d'Ascq Cedex, France, `lastname@univ-lille3.fr`

² Laboratoire d'Informatique du Littoral (LIL), Université du Littoral Côte d'Opale, UPRES-JE, B.P. 719, 62228 Calais Cedex, France, `lastname@lil.univ-littoral.fr`

Résumé Nous étudions l'évolution d'une interaction entre organismes adaptatifs. Loin d'être fixe, une interaction évolue au cours du temps. Nous nous appuyons sur la loi de l'effet qui décrit la sélection du comportement d'un organisme vivant pour explorer la dynamique d'une interaction. Notre argumentation s'appuie sur des expériences de psychologie sur des sujets humains. Nous montrons l'importance du temps dans la dynamique d'une interaction. Nous tirons diverses conclusions concernant notamment la vie artificielle et la psychologie sociale où la notion d'interaction est importante.

1 Introduction

Imitation, coopération, trahison, parasitisme, exploitation, utilisation ... qualifient la nature d'interactions entre des agents, vivants ou non. Quand deux agents vivants entrent en interaction, la nature de celle-ci change au cours du temps. Ce changement est lié à la nature adaptative des organismes vivants. De l'évolution de la nature des interactions entre agents naissent les relations et les structures sociales au sens très large du terme.

L'adaptation mise en jeu concerne ici le comportement des agents. Nous ne nous intéressons pas ici à l'évolution sur plusieurs générations des interactions. Ce type d'adaptation du comportement a été largement étudié depuis le début du XX^e siècle, notamment par des éthologues et des psychologues pratiquant l'analyse expérimentale du comportement. Ainsi, Thorndike a proposé un modèle simple résumant ses observations sur l'animal concernant l'émission d'un comportement. Ce modèle indique que pour un organisme vivant donné, l'un de ses comportements voit sa probabilité d'émission augmenter si son émission est suivie de conséquences positives et que cette probabilité diminue si l'émission du comportement est suivie de conséquences négatives; c'est la « loi de l'effet » [28,29,4]. Nous ajouterons que la probabilité d'émission d'un comportement n'est jamais tout à fait nulle à partir du moment où l'organisme peut effectivement (physiquement) l'émettre. Ce modèle de la « sélection du comportement par ses conséquences » a été observé, expérimenté, testé et vérifié à de

* durant la réalisation de ce travail, Samuel Delepouille a été financé par le Conseil Régional Nord-Pas de Calais, France, sous le contrat n 97 53 0283

très nombreuses reprises, au point où l'issue d'expériences peut être prédite¹. Des comportements éventuellement très complexes [12] peuvent ainsi être appris à des organismes vivants : le dressage de chiens d'aveugles ou d'animaux de cirque ne fait pas appel à d'autres principes, de même que l'apprentissage de la bicyclette chez l'humain et même, l'apprentissage de la résolution d'un système d'équations différentielles, chez l'humain également. Les exemples sont innombrables. De nombreux comportements « naturels » (observés chez les animaux vivants dans leur milieu naturel et n'ayant jamais été dressés) s'expliquent par leur acquisition au cours de leur existence *via* ce même principe de la sélection du comportement par ses conséquences [25,26,11]. Aussi, ce principe constitue la règle de base régissant le comportement des agents discutés dans cette contribution.

Aujourd'hui, un courant de recherches pluridisciplinaires alliant aussi bien la psychologie, l'éthologie et l'informatique propose d'étudier la dynamique comportementale des systèmes vivants en utilisant l'outillage mathématique des systèmes dynamiques non linéaires. Différents travaux ont ainsi mis en évidence que l'évolution de certains comportements peut être décrite avec des équations différentielles que l'on rencontre classiquement en physique. Déjà ancienne en éthologie [10], cette approche est aujourd'hui utilisée en psychologie [18,27,5,20,7], mais aussi en neurophysiologie [19,15]. Ces travaux montrent clairement qu'une description dynamique du comportement d'un organisme vivant et de son évolution est envisageable, possible et pertinente. Cette dynamique ressemble à celle des systèmes physiques loin de l'équilibre dont on a proposé qu'elle décrive la dynamique du vivant [22]. Des phénomènes de bifurcation et d'hystérésis ont été mis en évidence à diverses reprises [18,5].

Aussi, plus que la nature même d'une interaction à un instant donné, nous voulons souligner l'importance de l'histoire de l'interaction et l'étudier. Depuis de nombreuses années, des sciences humaines, en particulier la sociologie, ont emprunté le vocabulaire de la cybernétique pour se donner une assise scientifique. Avec les travaux réalisés depuis une dizaine d'années sur la dynamique du comportement animal, nous pouvons maintenant aller bien au-delà du discours : des modèles dynamiques non linéaires ont été proposés afin de décrire la dynamique du comportement. Lorsqu'il n'est plus possible d'utiliser un modèle analytique, des modèles informatiques de type multi-agents sont envisageables [8,9].

Dans cette contribution, nous synthétisons nos réflexions en nous appuyant sur des expérimentations de psychologie sur des sujets humains permettant l'étude de l'évolution de l'interaction. Nous montrons ainsi comment les conditions sous lesquelles est réalisée l'expérience sélectionne et contraint la dynamique de l'interaction. De ce travail, nous tirons des conséquences sur le problème du dilemme itéré des prisonniers, célèbre modèle censé modéliser les situations de

¹ afin d'éviter les mauvaises interprétations, nous insistons immédiatement sur le fait que ce modèle ne relève en aucune sorte d'un schéma stimulus-réponse ; ce modèle porte également les noms de conditionnement opérant, conditionnement instrumental, ou de sélection du comportement par ses conséquences

coopération, et soulevons une question qui n'a pas été envisagée à ce jour. Malgré cet état de fait, cette question est cependant capitale et touche à la pertinence du dilemme itéré des prisonniers pour analyser des situations de coopération réelle, telles celles qu'envisagent Axelrod lui-même.

2 Organisation temporelle des comportements au cours d'une interaction

Nous avons réalisé une expérience qui consiste à placer deux sujets humains dans une situation de coopération minimale, dite « contrôle mutuel du destin ». Cette expérience relevant de la psychologie sociale est inspirée de celle initialement réalisée et décrite dans [24]. L'expérience est réalisée à l'aide de deux ordinateurs reliés l'un à l'autre et pouvant communiquer. Ces deux ordinateurs sont situés dans deux pièces différentes ; un sujet humain est placé devant chacun des ordinateurs sans connaître l'existence de l'autre sujet ni de l'autre ordinateur. Chaque sujet interagit avec son ordinateur au travers d'une interface très simple : 2 boutons sur lesquels on lui indique qu'il peut cliquer avec la souris et un nombre est affiché (voir fig. 1). Initialement à 100, ce nombre est incrémenté à chaque fois que l'autre sujet appuie sur l'un des deux boutons, décrémente lorsqu'il appuie sur l'autre. Bien entendu, chaque sujet devant son ordinateur ne connaît pas l'effet de ses clics sur l'autre sujet ; lui-même voit un nombre s'incrémenter et se décrémente sur son écran sans raison apparente. De par sa nature, nous appellerons ce nombre un « score ». Il est alors indiqué aux sujets que leur objectif est de maximiser leur score (il est également vérifié que les sujets ont compris ce qui leur est demandé et qu'ils sont capables de cliquer sur les boutons). Les deux sujets sont ensuite laissés à eux-mêmes durant 30 minutes. Nous ne décrivons pas ici plus avant le protocole expérimental ni l'ensemble des résultats obtenus ; tout cela est disponible dans la thèse de doctorat de psychologie de S. Delepouille [7]. Nous nous concentrerons ici sur un point-clé des résultats de cette expérience. Ce point est le suivant : dans tous les cas où les sujets arrivent à s'augmenter mutuellement leurs scores, leurs comportements sont synchronisés : ils cliquent aux mêmes instants (*cf.* une définition plus précise plus loin) ; pour être précis sur ce point très important, nous insistons sur le fait que les deux sujets agissent initialement sans aucune contrainte : ils cliquent quand ils le désirent et chacun de leurs clics est transmis immédiatement à l'autre ordinateur, donc à l'autre sujet sous la forme d'une variation de son score. Aussi, les comportements des deux sujets ne sont en aucun cas synchronisés au début de l'expérience. Cependant, dès que leurs comportements se synchronisent, les deux sujets s'augmentent mutuellement leur score, ce qui, par feedback positif, augmente encore leur probabilité d'être synchronisés.

Une remarque doit être faite à ce stade : on pourrait imaginer qu'une fois synchronisés, les sujets ne vont pas forcément s'augmenter mutuellement leurs scores, mais les diminuer, ou encore, l'un des scores va augmenter pendant que l'autre diminuera. Hors cela n'est pas observé et s'explique par le principe de la sélection des comportements par leurs conséquences ET par le fait que les sujets

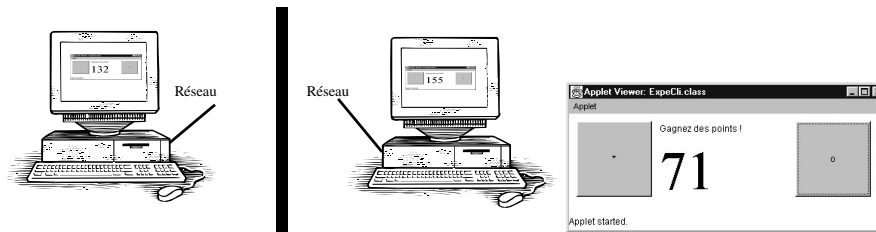


FIG. 1 – La partie gauche de la figure indique la situation des sujets humains durant l'expérience : ils sont localisés dans deux salles différentes, chacun étant situé devant un ordinateur ; les deux ordinateurs sont reliés de manière à ce que les transmissions entre les deux ordinateurs soient immédiates. À droite, on voit l'interface affichée sur chacun des deux ordinateurs ; on y distingue les deux boutons banalisés et le score du sujet.

se sont synchronisés. Nous montrons rapidement que sous ces deux hypothèses, la dynamique du système composé des deux sujets ne peut que converger rapidement vers une situation où chaque sujet incrémente le compteur de l'autre (sans le savoir bien entendu) ; nous dirons alors que les deux sujets « coopèrent ».

Soit :

- I le comportement consistant pour un sujet à cliquer sur le bouton qui incrémente le score de l'autre sujet ;
- D le comportement consistant pour un sujet à cliquer sur le bouton qui décrémente le score de l'autre sujet.

Il y a 3 cas différents :

1. les deux sujets cliquent initialement sur D. Les conséquences étant négatives, ils changent et émettent I. Les conséquences sont positives et ils continuent à cliquer sur I tous les deux ;
2. le sujet 1 clique sur I alors que le sujet 2 clique sur D (ou inversement). Le sujet 1 va donc changer pour D et le sujet 2 va cliquer à nouveau sur D. Nous retrouvons alors le premier cas ;
3. les deux sujets cliquent sur I. Les conséquences sont positives, ils n'ont donc pas de raison d'émettre D.

Dans les 3 cas, les sujets en viennent rapidement à cliquer tous les deux sur I s'ils se comportent en suivant le principe de sélection du comportement par ses conséquences. Cette dynamique est résumée sur le schéma de la figure 2.

Notons également que si les sujets ne suivent pas strictement les conséquences de leur comportement², l'émission par chacun des sujets du comportement I reprend toujours rapidement le dessus puisque c'est la seule configuration stable du système.

² le principe de sélection du comportement par ses conséquences s'exprime avec des probabilités

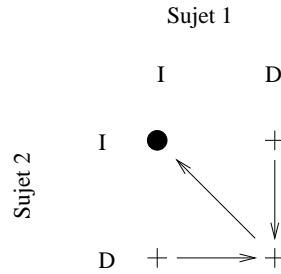


FIG. 2 – Portrait de phase pour la situation de contrôle mutuel du destin. L'espace des phases (configurations possibles) comprend quatre points, chaque sujet ayant le choix entre deux comportements. Quels que soient les comportements initialement émis par les sujets, le système converge vers l'émission par chacun des sujets du comportement I (qui entraîne l'incrémement du score de l'autre sujet). Le point (I, I) de l'espace des phases est donc un point fixe stable, noté par un disque plein. Les flèches indiquent les trajectoires dans l'espace des phases.

Aussi, cette expérience montre l'importance capitale de la synchronisation des comportements : cette synchronisation entraîne irrémédiablement l'issue de l'expérience dès lors que le comportement des sujets suit le modèle de la sélection du comportement par ses conséquences. Par ailleurs, d'autres expériences se sont intéressées à la même situation dès lors que la synchronisation des comportements est forcée [17]. Dans ce cas, les comportements coopératifs apparaissent beaucoup plus rapidement.

Pour être précis, par « synchronisation », nous signifions que les fréquences d'émission de comportement par les sujets, ν_A et ν_B , sont telles qu'il existe un rationnel positif non nul q tel que $\nu_A = q \nu_B$ et que les deux sujets sont en phase ; les sujets sont en quelque sorte en résonance.

3 Le dilemme itéré des prisonniers

Nous montrons ici les conséquences de ce qui précède sur le problème du dilemme itéré des prisonniers (DIP). Nous commençons par rappeler la définition du dilemme itéré du prisonnier introduite par R. Axelrod [2].

Dans ce problème, deux sujets ont le choix entre deux comportements, coopérer ou trahir. Si les deux sujets coopèrent, ils obtiennent chacun un gain R (que l'on suppose numérique). Si les deux sujets trahissent, ils obtiennent chacun un gain P . Si l'un des deux sujets coopèrent pendant que l'autre trahit, celui qui a coopéré obtient un gain S tandis que l'autre obtient un gain T . Les gains doivent respecter deux inégalités :

$$T > R > P > S$$

$$2R > T + S$$

Ces inégalités entraînent qu'il est plus avantageux pour les deux sujets de coopérer que de trahir ($R > P$); par contre, quand un sujet coopère alors que l'autre trahit, celui qui coopère obtient le gain minimal (S) alors que celui qui trahit obtient un gain maximal (T).

Le DIP consiste alors à se faire rencontrer les deux sujets plusieurs fois (à l'infini). Dans le DIP que nous qualifions de classique (celui proposé par Axelrod et étudié dans la littérature), la séquence d'actions suivantes est effectuée virtuellement à l'infini :

1. les deux sujets émettent chacun un comportement ;
2. un arbitre confronte leurs comportements et leur verse leurs gains ;
3. on itère à l'étape 1.

Dans une situation naturelle où le comportement des agents suit la loi de l'effet, la trahison l'emporte généralement : des expériences ont été réalisées avec des pigeons [13] et des humains [6]. On montre facilement que pour 2 agents synchronisés sélectionnant leur comportement en vertu de la loi de l'effet, la trahison réciproque est le seul attracteur stable.

Soit :

- C le comportement de coopération des sujets ;
- T le comportement de trahison des sujets.

Il y a à nouveau 3 cas :

1. les deux sujets émettent C. Dans ce cas, leurs scores augmentent et les sujets continuent d'émettre C. Aussi, (C, C) se comporte comme un point fixe stable. Cependant, si l'un des sujets émet T alors que l'autre sujet continue d'émettre C, le système perd sa stabilité et passe dans le cas suivant. Le système ne reviendra plus dans l'état (C, C) : ce point est fixe mais instable ;
2. si le sujet 1 émet C alors que le sujet 2 émet T (ou inversement), le sujet 1 va ensuite émettre T et le système atteint le cas suivant ;
3. si les deux sujets émettent T, leurs scores augmentent et ils continuent chacun d'émettre T.

Le point (T, T) est donc un point fixe stable du système. Cette dynamique est résumée sur le schéma de la figure 3.

À nouveau, notons que si les sujets ne suivent pas strictement les conséquences de leur comportement, l'émission par chacun des sujets du comportement T reprend toujours le dessus puisque c'est la seule configuration stable du système.

Pour pouvoir coopérer, il faut que les agents ne suivent pas la loi de l'effet ou il faut changer un élément du dilemme itéré des prisonniers. En effet, nous notons que les agents sont synchronisés dans cette définition du DIP. Selon ce qui vient d'être dit à la section 2, cela peut contraindre la dynamique du DIP. Il est donc intéressant de proposer une version désynchronisée du DIP, la synchronisation des comportements devenant une conséquence éventuelle du comportement des agents, plutôt qu'un pré-requis.

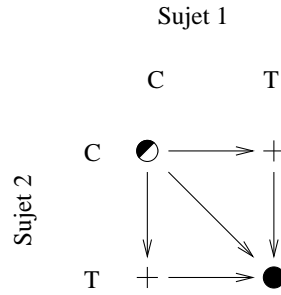


FIG. 3 – Portrait de phase pour le dilemme itéré des prisonniers. Quels que soient les comportements initialement émis par les sujets, le système converge vers l'émission par chacun des sujets du comportement T. Le point (T, T) de l'espace des phases est donc un point fixe stable, noté par un disque plein. Le point (C, C) est également intéressant : c'est un point fixe instable : tant que les deux sujets émettent C, ils continuent à émettre C ; dès que l'un des sujets émet T alors que l'autre émet C, le système quitte la configuration (C, C) pour converger vers (T, T). Graphiquement un point fixe instable se représente par un disque à moitié plein.

Aussi, nous avons montré qu'il est possible de formuler le DIP sous une forme désynchronisée [7]. Sous cette forme, le problème est défini par deux matrices, chacune correspondant à un comportement, et la valeur de ses éléments donnant les conséquences de l'émission du comportement par l'un et l'autre des agents sur lui-même et sur l'autre agent. Cette formulation inclut le cas synchronisé comme cas particulier. Nous avons alors montré qu'un DIP donné par un quadruplet (T, R, P, S) donne lieu à une infinité de versions désynchronisées différentes. De plus, nous avons montré que ces DIP désynchronisés (DDIP) issus du même DIP peuvent avoir des dynamiques qualitativement différentes pour des agents dont le comportement est sélectionné en vertu de la loi de l'effet : les attracteurs des comportements sont différents : dans un cas, l'attracteur est la coopération mutuelle tandis que dans l'autre cas, c'est la trahison mutuelle ou la trahison de l'un et la coopération de l'autre qui est l'attracteur.

Sous forme désynchronisée, tous les cas de figure sont possibles pour des agents sélectionnant leurs comportements en fonction de leurs conséquences. Aussi, ceci ne contredit pas le fait que des situations de DIP aient été observés dans la nature qui mènent à la coopération. En effet, cette désynchronisation des comportements initiaux est bien plus proche des situations réelles décrites par Axelrod lui-même : des puissances antagonistes ne lancent pas leurs missiles en même temps et ceux-ci n'atteignent pas leurs cibles en même temps non plus : les comportements réels ne sont jamais synchronisés.

4 Discussion

Dans cet article, nous avons présenté une modélisation simplifiée de la notion d'interaction. Cette modélisation met l'accent sur la sélection des comportements

en fonction de leurs conséquences. Les conséquences portent non seulement sur l'agent qui émet le comportement mais aussi sur celui avec lequel il interagit. Bien que simple (dans son expression), ce modèle permet de reformuler des problèmes connus. Nous pouvons ainsi explorer les situations utilisées par la psychologie expérimentale (coopération minimale) mais aussi les problèmes formulés par la théorie des jeux tel que le dilemme itéré des prisonniers. Ce modèle est donc compatible avec les formulations existantes de la coopération. De plus, en ne supposant pas que les interactions sont synchronisées, il est possible d'étendre l'analyse à des situations en temps réel. À cet égard, la majorité des interactions naturelles sont caractérisées par un mode de réponse désynchronisé. Nous avons montré que l'hypothèse de synchronisation des comportements n'est pas pertinente pour l'analyse des situations naturelles. Une fois désynchronisée, la dynamique de l'interaction entre les agents est riche et variée. Sur ce point, notre travail ouvre de nombreuses perspectives de travaux dans le domaine de la vie artificielle et propose un éclairage renouvelé de certains points en psychologie sociale. D'un point de vue plus « technologique », nous pensons que la désynchronisation des actions des agents, loin de poser de nouveaux problèmes ou de compliquer ceux qui existent, en rendant une certaine souplesse aux agents, donc des capacités accrues d'adaptation, suit l'esprit des recherches dans le domaine de l'intelligence artificielle basée sur les comportements (voir par exemple [1,3,21]). Il nous semble que les problématiques de négociation entre agents pourraient être simplifiées en utilisant des agents adaptatifs, plutôt qu'en essayant de faire communiquer des agents à l'architecture et aux comportements figés.

La conception de l'évolution des comportements permet de comprendre comment une interaction est essentiellement caractérisée par sa dynamique. Elle repose sur la description d'un système constitué de deux agents (humain, animal, société, ...). Ce système, que l'on désigne fréquemment sous le nom de dyade, peut être considéré comme l'élément de base de phénomènes sociaux plus complexes [14,23]. L'étude formelle ou empirique des interactions dyadique ouvre donc la porte à une analyse comportementale de la réalité sociale [16].

Références

1. R.C. Arkin. *Behavior Based Robotics*. MIT Press, 1998.
2. R. Axelrod. *The evolution of cooperation*. Basic Books, 1984.
3. R.A. Brooks. *Cambrian Intelligence*. MIT Press, 1999.
4. C. Catania. Thorndike's legacy : learning, selection, and the law of effect. *Journal of the experimental analysis of behavior*, 72 :425–428, 1999.
5. C. Clément. *L'ontogénèse du contrôle temporel du comportement humain envisagé comme un système dynamique non linéaire*. PhD thesis, Université de Lille 3, URECA, Villeneuve d'Ascq, October 1999. Thèse de doctorat de Psychologie.
6. S. Delepouille. La coopération entre agents adaptatifs. Master's thesis, Université de Lille 3, Villeneuve d'Ascq, July 1997. mémoire de DEA de psychologie.
7. S. Delepouille. *Coopération entre agents adaptatifs ; étude de la sélection des comportements sociaux, expérimentations et simulations*. PhD thesis, Université de

Lille 3, URECA, Villeneuve d'Ascq, October 2000. Thèse de doctorat de Psychologie.

8. S. Delepouille, Ph. Preux, and J-C. Darcheville. Evolution of cooperation within a behavior-based perspective : confronting nature and animats. In *Artificial Evolution'99*, volume 1829 of *Lecture Notes in Computer Science*, pages 204–216. Springer-Verlag, 2000.
9. S. Delepouille, Ph. Preux, and J-C. Darcheville. Simulation of social behaviors : why and how ? *Mexican journal of behavior analysis*, 26(2) :191–209, 2000.
10. J.L. Deneubourg. Application de l'ordre par fluctuations à la description de certaines étapes de la construction du nid chez les termites. *Insectes Sociaux*, 24 :117–130, 1977.
11. J.W. Donahoe. Edward L. Thorndike : the selectionist connectionist. *Journal of the experimental analysis of behavior*, 72 :451–454, 1999.
12. J.W. Donahoe and D.C. Palmer. *Learning and Complex Behavior*. Allyn and bacon, 1994.
13. L. Green, P.C. Price, and M.E. Hamburger. Prisoner's dilemma and the pigeon : control by immediate consequences. *Journal of the Experimental Analysis of Behavior*, 67 :1–17, 1995.
14. D.F. Hake and D.R. Olvera. *Cooperation, competition, and related social phenomena*, pages 208–245. Irvington Publisher inc, New-York, 1978.
15. R.P. Ingvaldsen and H.T.A. Whiting. Modern view on motor skill learning are not representative. *Human Movement Science*, 16 :705–732, 1997.
16. F.S. Keller and W.N. Schoenfeld. *Principles of psychology*. Appleton-Century Inc, 1950.
17. H.H. Kelley, J.W. Thibaut, R. Radloff, and D. Mundy. The developpement of cooperation in the "minimal social situation". *Psychological Monographs*, 76 :1–19, 1962.
18. S. Kelso. *Dynamic Patterns, the self-organization of brain and behavior*. MIT Press, 1995.
19. R. King, J.D. Barchas, and B.A. Huberman. Chaotic behavior in dopamine neurodynamics. In *Proc. National Academy of Science, Neurobiology*, volume 81, pages 1244–1247, 1984.
20. L. Madelain. *Approches dynamique et opérante de l'ontogénèse du comportement de poursuite visuelle lisse chez le nouveau né et de ses dysfonctions chez le sujet schizophrène*. PhD thesis, Université de Lille 3, URECA, Villeneuve d'Ascq, May 2000. Thèse de doctorat de Psychologie.
21. P. Maes. Behavior-based artificial intelligence. In *Proc. of the Second Conference on Adaptive Behavior*, pages 2–10. MIT Press, 1993.
22. G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems*. Wiley, 1977.
23. E. Ribes-Iñesta. Functional dimensions of social behavior : theoretical considerations and some preliminary data. In *VIth Biannual Symposium on the Science of Behavior held at Guadalajara (Mexico)*, Février 2000.
24. J.B. Sidowski, B. Wyckoff, and L. Tabor. The influence of reinforcement and punishment in a minimal social situation. *Journal of Abnormal Social Psychology*, 52 :115–119, 1956.

25. B.F. Skinner. Selection by consequences. *Science*, 213 :501-514, 1981.
26. J.E.R. Staddon. *Adaptive behavior and learning*. Cambridge University Press, 1983.
27. E. Thelen and L.B. Smith. *A dynamic systems approach to the development of cognition and action*. MIT Press, 1994.
28. E.L. Thorndike. Animal intelligence : An experimental study of the associative process in animals. *Psychology Monographs*, 2, 1898.
29. E.L. Thorndike. *Animal Intelligence : Experimental Studies*. Mac Millan, 1911.