

# L'Économie des Conventions à l'école des institutions

HERVÉ DEFALVARD

*Université de Marne-la-Vallée et Centre d'études de l'emploi*

DOCUMENT DE TRAVAIL

N°02

juillet 2000

Ce texte a été soumis au **comité éditorial** du CEE, composé de : Gabrielle Balazs, Jean-Claude Barbier, Pierre Boisard, Bruno Courault, François Eymard-Duvernay, Jérôme Gautié, Michel Gollac, Françoise Laroche, Marie-Thérèse Letablier, Martine Lurol, Emmanuèle Reynaud, Bernard Simonin, Marie-Madeleine Vennat, Serge Volkoff.

# L'Économie des Conventions à l'école des institutions

Hervé Defalvard

## Résumé

Développé depuis le milieu des années 1980, le programme de recherches de l'Économie des Conventions connaît aujourd'hui une transformation majeure, liée à sa récente rencontre avec les institutions. Cette prise en compte des institutions est allée de paire avec la mise à jour des limites des notions de convention et d'incertitude, se révélant insuffisantes pour explorer sans réductionnisme les mécanismes de la coordination sociale. D'autre part, cette référence aux institutions a permis à l'Économie des Conventions de développer une critique renouvelée de la rationalité standard et de proposer une nouvelle modélisation de l'action individuelle. À la suite de l'inventaire de cette (r)évolution théorique, nous proposons un modèle de jeux à institutions qui s'inscrit dans cette nouvelle perspective.

**Mots-clefs :** Coordination sociale, institutions, conventions.

**Classification JEL :** B41, C72

## *Economics of Conventions discovering institutions*

*Hervé Defalvard*

### *Abstract*

*Born in the middle of eighties, the research programm of Economics of Conventions is today the subject of an important transformation. That one is linked on the tour of institutions, which went hand in hand the critique of convention and uncertainty notion's, inadequate to think the social coordination. In relation to institutions, Economics of Conventions today writes a new critique of standard rationality hypothesis and develops a new theory of individual actions, set in collective environments. After brought out the merits of this new Economics of Conventions, we propound in same perspective a model of game with institutions.*

**Key words :** *Social Coordination, Institutions, Conventions.*

**JEL Classification :** *B41, C72*



Voici un peu plus de dix ans que naissait, en France, un programme de recherches en économie, appelé l'Économie des Conventions<sup>1</sup>. Il s'agissait, en faisant porter l'accent sur les conventions, d'étudier une forme singulière de coordination, marchande ou non, qui était « à appréhender à la fois comme le résultat d'actions individuelles et comme un cadre contraignant les sujets »<sup>2</sup>. En accord avec ce projet, l'Économie des Conventions revendiqua un « individualisme sophistiqué » qui, tout en réduisant les conventions aux comportements individuels, maintenait l'incomplétude des individus. Dans une étude consacrée à sa posture aussi bien ontologique que méthodologique, nous avons alors conclu que, pour devenir une véritable alternative à l'économie néoclassique<sup>3</sup>, l'Économie des Conventions aurait à penser l'incomplétude de l'individu autrement que dans une référence à l'incertitude, probabilisable ou non ; le défi théorique consistait, selon nous, à « voir » l'incomplétude individuelle en relation avec la nature collective des mondes, au sein desquels tout individu agit<sup>4</sup>.

Une évolution récente et convergente des travaux de l'Économie des Conventions marque une étape décisive dans cette direction<sup>5</sup> ; telle est, en tout cas, la thèse centrale de ce papier, étudiant le glissement de l'école des conventions vers l'analyse des institutions. Ce glissement n'est pas mineur. Il est même à l'origine d'une nouvelle orientation, explicitée par Orléan en ces termes : « les réflexions du présent article trouvent leur origine dans un programme de recherches qui milite pour *une approche formalisée des formes institutionnelles* »<sup>6</sup>, ou par Favereau quand il écrit que : « au terme de cet examen de la question de l'interdépendance des niveaux de décision (IND), émerge un programme de recherches non standard, *traitant les institutions d'une façon qui simultanément échappe et emprunte aux méthodologies traditionnelles* : alors que la tradition individualiste ne voit, derrière les institutions, que le choc des décisions, ce programme hétérodoxe se rapproche de la tradition holiste, sans s'y perdre, en donnant enfin à voir, derrière le choc des décisions, le poids des situations »<sup>7</sup>. Cette évolution nous paraît cependant plus spontanée que programmée ; en effet, les quatre textes imprimant cette nouvelle direction à l'Économie des Conventions, ne forment pas un réseau de références croisées : seul le texte de Salais (1998) contient un renvoi à l'article de Favereau (1998). Ils ne constituent pas non plus un corpus homogène : si les trois premiers, par ordre chronologique de publication, sont programmatiques, la dernière étude de Eymard-Duvernay et alii (1999), sur les institutions du recrutement en France et en Grande-Bretagne, se développe

---

<sup>1</sup> On peut prendre comme acte officiel de sa naissance, le numéro spécial de la *Revue Économique* de mars 1989, intitulé « L'économie des conventions ». Si ce programme a rejoint des développements théoriques menés ailleurs qu'en France, dont on trouvera trace dans l'ouvrage *L'analyse économique des conventions*, publié sous la direction de A. Orléan (1994), il ne marque pas moins une spécificité française aussi bien dans la critique de la théorie dominante, en économie, que dans la conduite d'une alternative à celle-ci.

<sup>2</sup> Introduction collective, dans *Revue Économique*, mars 1989, p. 145.

<sup>3</sup> Plutôt que de parler de théorie standard et de théorie standard étendue, nous nous référons, selon une vue plus historique des choses, à l'économie néoclassique, dont nous avons par ailleurs proposé une histoire raisonnée dans Defalvard (2000a).

<sup>4</sup> Defalvard (1992, p. 141-142). Cette voie de recherches nous a également conduit à intégrer les institutions pour poser la question de la coordination sociale, voir Defalvard (1999) et (2000b).

<sup>5</sup> Nous nous référons aux études de Orléan (1997), *Jeux évolutionnistes et normes sociales* ; de Salais (1998), *À la recherche du fondement conventionnel des institutions* ; de Favereau (1998), *Décisions, situations, institutions* ; de Eymard-Duvernay et alii (1999), *Les institutions du recrutement*. L'étude plus antérieure de Thévenot (1995), *L'action en plan*, peut être vue rétrospectivement comme un point de passage intermédiaire vers l'intégration des institutions pour penser l'action ; le plan n'étant déjà plus seulement une convention pour l'action sans être tout à fait une institution.

<sup>6</sup> Op. cit., p. 197, souligné par nous.

<sup>7</sup> Op. cit., p. 166, souligné par nous.

déjà à l'intérieur de ce nouveau programme de recherches ; pour cette raison, elle n'occupera qu'une place marginale dans notre enquête.

Dans une première partie, nous montrerons comment sa rencontre avec les institutions a conduit l'Économie des Conventions à dépasser les deux concepts de base sur lesquels elle fondait sa différence : les conventions et l'incertitude. Puis, dans une deuxième partie, nous synthétiserons les avancées de son nouveau programme de recherches, aussi bien du point de vue de sa critique de la rationalité standard que de sa modélisation de l'action individuelle en présence d'institutions. Enfin, dans une troisième et dernière partie, nous proposerons une voie possible de développement de cette nouvelle Économie des Institutions, en construisant un modèle de jeu à institutions.

## 1. LES FAIBLESSES DE L'ANCIEN PROGRAMME

Dès son départ, l'ambition affichée par l'Économie des Conventions a fait l'objet de réserves quant à la portée véritablement alternative de son programme. L'impression dominante était qu'elle restait au milieu du gué, faisait du neuf avec de l'ancien<sup>8</sup>. Mais c'est sans aucun doute l'étude de Postel (1998) qui a le mieux fait ressortir le maillon faible de l'Économie des Conventions, relativement à l'exigence de fonder les règles et les institutions nécessaires à la réussite de la coordination sur une dynamique historique, et non pas naturelle : « il faut, écrit-il, pour satisfaire cette troisième exigence doter les acteurs d'une capacité à créer ou à faire évoluer les règles et les institutions qui leur permettent de se coordonner. Or, la prise en compte de cette capacité implique d'élargir considérablement l'hypothèse standard de rationalité individuelle qui est strictement instrumentale. (...) cet élargissement représente un obstacle théorique que l'économie des conventions a eu tendance à contourner » (p. 1475).

Notre thèse est que l'inflexion récente de l'Économie des Conventions vers les institutions lui a déjà permis de commencer les travaux théoriques nécessaires à cet élargissement ; non pas d'ailleurs en réponse aux réserves qu'elle a pu susciter, mais en découvrant, par voie interne, que l'étude des conventions et de l'incertitude ne menait pas, par elle-même, à rompre avec l'approche réductionniste de la coordination sociale<sup>9</sup>.

### 1.1. L'insuffisance des conventions

La mobilisation des conventions, dans le but d'en faire le cheval de Troie d'une alternative à la théorie néoclassique, a vite tourné court pour une raison simple mais robuste : sa reprise à l'intérieur même du paradigme dominant en tant que solution à la multiplicité des équilibres dans les jeux non coopératifs ou évolutionnistes<sup>10</sup>. C'est, d'une certaine manière, ce constat qui a conduit aussi bien Orléan (1997) que Salais (1998) à proposer de dépasser la notion de convention. Nous suivrons,

---

<sup>8</sup> Cette vue a notamment été exprimée dans les études de Quéré (1993), de Rallet (1993) ou, encore, de Defalvard (1992).

<sup>9</sup> Une façon complémentaire de présenter l'évolution de l'Économie des Conventions est de soutenir que l'entrée dans son programme de recherches par les sciences sociales autres que l'économie a fini par imposer à la seconde entrée à partir de l'économie orthodoxe, une sortie du paradigme néoclassique. On trouvera, dans Favereau (1995, p. 104), cette référence à une double entrée dans le programme de l'Économie des Conventions.

<sup>10</sup> Voir Sugden (1986), cité par Orléan (op. cit., p. 190), ou encore Chiappori (1994) et Young (1996).

ici, l'article de Orléan, qui démontre, de manière approfondie, les raisons de cette insuffisance des conventions.

Afin de repérer le point de sortie de l'analyse néoclassique, Orléan se réfère au traitement que la théorie des jeux évolutionnistes (TJE) réserve aux conventions. Notons C une convention (celle, par exemple, de rouler à droite) ; son émergence est expliquée par la TJE par le fait que son adoption par tous les individus améliore l'utilité de chaque individu. Formellement, cela s'écrit :

$$u_i(e, C) \geq u_i(e) \quad \forall i \quad \forall e \in E$$

où E désigne l'ensemble des états du monde possibles, et où  $u_i(e, C)$  désigne l'utilité de l'individu i quand l'état du monde est e et quand est présente la convention C. L'émergence d'une convention repose généralement sur un phénomène de croyance intersubjective, illustré par Orléan sur le cas de figure appelé « l'effet Reagan ». Dans une déclaration rapportée par le *New York Times* du 12 novembre 1987, Reagan affirmait que le dollar avait suffisamment baissé. Notons Q cette proposition de Reagan, dont on a pu observer en les interrogeant que les opérateurs financiers ne lui accordaient aucun crédit. Chacun d'eux soutenait, en fait, la proposition P selon laquelle il ne fallait accorder aucun crédit à Q. Or, tous les cambistes ont néanmoins acheté du dollar car si chacun d'eux croyait P, chacun pensait aussi que les autres cambistes pourraient accorder un crédit à Q, d'où leur intérêt à acheter finalement du dollar, par anticipation du fait que les autres allaient en acheter. On peut formaliser les deux niveaux de croyance à l'origine de « l'effet Reagan » de la manière suivante :

$$C_i P \quad \forall i \quad (C_i P \text{ se lisant } i \text{ croit } P)$$

$$C_i C_h Q \quad \forall h \neq i \text{ et cela } \forall i \quad (C_i C_h Q \text{ se lisant } i \text{ croit que } h \text{ croit } Q)$$

La leçon, ici, est que la croyance intersubjective domine toujours la croyance individuelle. Ainsi, un Anglais visitant la France peut croire qu'il est préférable de rouler à gauche, mais supposant que tous les autres pensent qu'il faut rouler à droite, alors notre Anglais roulera à droite. Toutefois, l'intégration des croyances intersubjectives à l'origine de l'émergence et de la stabilité des conventions n'est pas suffisante pour sortir des rets du paradigme dominant<sup>11</sup>. Pourquoi ?

La raison, mise en avant par Orléan, est que les conventions, même fondées sur des croyances intersubjectives, n'en demeurent pas moins des équilibres, dont la stabilité repose toujours sur un calcul d'utilité. Cette option, selon lui, en reste à une version étroite des conventions, car « la stabilité des institutions suppose la mobilisation de jugements qui vont bien au delà du seul calcul des utilités : il s'agit d'en reconnaître la légitimité, c'est-à-dire le bien fondé au regard des valeurs partagées par le groupe » (p. 190). Il critique alors la position de la TJE qui traite toutes les institutions à la manière d'un équilibre alors même « qu'il convient de distinguer conventions, conventions légitimées et normes sociales » (p. 191)<sup>12</sup>. La convention légitimée étant une convention devenue ce que Douglass

---

<sup>11</sup> Sur ce point, Orléan (p. 184-188) montre néanmoins une hésitation qui, selon nous, est liée à son absence de distinction entre les croyances intersubjectives et les croyances collectives.

<sup>12</sup> Nous avons déjà noté cette confusion entre convention et institution, propre à la théorie des jeux, dans une étude précédente, (Defalvard, 1999), qui par d'autres chemins participe également à l'élaboration de ce nouveau programme de recherches.

appelle « une institution sociale légitime »<sup>13</sup>, appartenant au domaine des institutions, et non plus seulement au champ de la pure coordination. En empruntant un raccourci, Salais rejoint également cette conclusion. Ainsi, en introduction de son texte, il souligne l'insuffisance de la notion de convention du fait qu'elle reste associée aux seuls problèmes de coordination, faisant l'impasse sur d'autres problèmes : « non plus seulement de coordination, mais d'action collective : ce déplacement impose la présence des institutions au sein de la coordination » (p. 256). Comme tout à l'heure, c'est à l'aune des institutions que se révèle l'insuffisance des conventions en tant que solution à un problème de coordination. Pour autant, Salais ne délaisse pas la notion de convention, mais l'intègre dans le champ de l'économie des institutions où elle devient, non une « convention légitimée » à la manière de Orléan, mais la « convention de l'État ».

## 1.2 Les limites de l'incertitude

Liée à celle de convention, la notion d'incertitude constitua le second pôle théorique à partir duquel l'Économie des Conventions choisit, au départ, de bâtir son programme de recherches. Si cette notion est restée implicite dans les analyses d'Orléan, en arrière-plan de la multiplicité des équilibres, elle occupe une place centrale dans la contribution de Favereau, vers laquelle nous nous tournons à présent.

À l'occasion d'une réflexion sur l'interdépendance des niveaux de décision (IND)<sup>14</sup>, Favereau reprend sa critique des théories standard et standard étendue, mais en lui donnant un nouveau tour. Ainsi, dans un premier temps, on retrouve la critique usuelle de la théorie néoclassique, lui reprochant son « postulat fondamental de dénombrement parfait des états de la nature ». Ce dernier, en effet, lui fait exclure de son champ, « les surprises potentielles (Shackle) et les contingences imprévues (Kreps) » (p. 158). Cette critique ouvre traditionnellement la voie, depuis Simon au moins, à une analyse menée en termes de rationalité limitée. Or, nous savons aujourd'hui que, loin d'être un objet non identifié par la théorie néoclassique, la rationalité limitée fait partie de son horizon théorique<sup>15</sup>. Nous retrouvons ainsi le même type d'impasse que celui où nous a conduits la notion de convention ; l'incertitude, même accompagnée de rationalité limitée, appartient toujours au cadre de la théorie dominante, elle n'en est pas une porte de sortie. C'est là qu'intervient la nouveauté dans les analyses de Favereau, qui ne se prolongent pas par un énième développement sur la rationalité limitée, mais par la mise en place des notions de situation et d'institution<sup>16</sup>.

La construction de ces notions s'avère en fait nécessaire pour rendre compte des décisions des individus dans des mondes où existe l'interdépendance des niveaux de décisions ; de tels mondes, en simplifiant, se caractérise par des individus I de niveau inférieur et des individus S de niveau supérieur, où « l'agent S prend sa décision avant l'agent I, l'antériorité étant ici logique plutôt que

---

<sup>13</sup> Citée par Orléan.

<sup>14</sup> Cette réflexion fut collective ; elle a donné lieu à un colloque réunissant toutes les équipes de FORUM, dont une partie des actes a été publiée dans Vinokur (1998).

<sup>15</sup> La rationalité limitée est intégrée au paradigme néoclassique par la porte des contrats incomplets dont la théorie est, selon Salanié (1994, p. 118), « indissociable d'une théorie de la rationalité limitée dont nous sommes encore loin de disposer ». Dans un travail remettant en cause l'hypothèse dite de « petit monde » propre à la théorie de Savage, MacLeod (1996) développe des premiers jalons en direction d'une théorie néoclassique de la rationalité limitée.

<sup>16</sup> Ce trajet, de la raison procédurale (versus raison instrumentale) vers les institutions, De Munck (1999) tente, dans son ouvrage *L'institution sociale de l'esprit*, d'en montrer la nécessité interne.



chronologique : les décisions de S délimitent le domaine du possible des décisions de I » (p. 154). De ce point de vue, le postulat de dénombrement parfait des états du monde n'est plus seulement une limite parce qu'il exclut les surprises ou les contingences, mais parce qu'il n'autorise « aucune incertitude sur la position même du décideur » (p. 159). Penser une telle incertitude impose d'introduire les institutions, car ce sont elles qui « servent à utiliser/canaliser/neutraliser l'IND pour fabriquer de l'ordre à partir de ce qui ne serait autrement que du désordre » (p. 160). Autrement dit, l'incertitude est ici conçue en relation avec les institutions qui ordonnent les niveaux de décision. Sur ce point, Favereau est rejoint par Eymard-Duvernay et alii<sup>17</sup>.

Ces derniers, en effet, soutiennent également que la prise en compte de la seule incertitude est insuffisante ; pour eux, elle ne peut conduire « qu'à comprendre comment les agents peuvent conclure des contrats efficaces dans de tels environnements » (p. 2). Aussi, ils ajoutent à leurs recherches un second volet qui, en plus de l'incertitude sur la qualité, vise à intégrer « le rôle des contextes institutionnels » (p. 3).

Au total, l'Économie des Conventions converge aujourd'hui pour dépasser son binôme de départ, qui croisait les conventions et l'incertitude, au profit d'une analyse de la coordination des actions situées dans des espaces institutionnels.

## **2. DE L'ACTION SITUÉE DANS UN MONDE D'INSTITUTIONS**

Si le point de départ du nouveau programme de recherches de l'Économie des Conventions reste lié à une critique de la théorie néoclassique, ce point d'appui est largement renouvelé par l'intégration des institutions dans le traitement de la coordination des actions. Par ailleurs, l'inflexion commune donnée à leur programme se traduit, selon les auteurs, par des entrées différentes dans le champ de l'institutionnel : les normes sociales pour Orléan, le choix social pour Salais, l'IND pour Favereau. Il est alors remarquable qu'à partir de ces entrées distinctes, ces auteurs aboutissent à des modèles convergents de l'action individuelle en société.

### **2.1 Une critique renouvelée de la rationalité standard**

Des décennies de critique de la rationalité standard auraient pu laisser penser que tout avait été dit sur ses limites et ses impasses. Ce n'est pas le moindre des mérites de l'Économie des Conventions que de faire progresser son analyse sur ce point précis en adoptant une approche institutionnaliste de la coordination.

#### **2.1.1. La critique de l'utilité des économistes**

---

<sup>17</sup> Sur la question de l'incertitude, la position de Salais est plus ambiguë. Ainsi, en introduction, il semble faire découler la nécessaire intégration des institutions de la question de l'incertitude, alors que, plus loin (p. 280), il considère, de manière plus fondée selon nous, l'incertitude en relation avec l'action collective et les institutions.

Comme nous l'avons déjà indiqué, Orléan s'attache d'abord à une critique de la TJE<sup>18</sup>, dont la limite principale est que, dans celle-ci, « les jugements individuels portent exclusivement sur l'utilité », et « qu'il n'est fait aucune place à l'évaluation normative des stratégies » (p. 178). Toutefois, il ajoute que cette limite dépasse le strict cadre de la TJE et, qu'à travers elle, « c'est la capacité de la théorie économique à proposer une analyse pertinente des institutions qui est alors posée » (p. 178). Ainsi, le défaut de la TJE, « faisant l'impasse sur toute notion de normativité », est-il commun à l'ensemble du paradigme néoclassique.

La TJE en économie étant, selon Orléan, une adaptation *a minima* de sa version initiale développée en biologie, il en découle que les comportements  $y$  sont totalement orientés par l'utilité des stratégies gagnantes, qui sont elles-mêmes fonction de la proportion de ceux qui, dans la population totale, les choisissent. Chaque agent  $i$ , après observation des  $u_i(p)$  où  $p$  est le vecteur des proportions, retient la stratégie qui maximise  $u_i(p)$ . Cette dynamique de la meilleure réponse conduit alors à des états stables qui sont des équilibres de Nash du jeu évolutionniste. Contre cette modélisation de la TJE, Orléan développe une critique qui pose toutefois un problème, celui de l'assimilation entre l'utilité d'une stratégie et son gain. Ainsi, dans le même temps où il reproche que « la seule question pertinente est de savoir *quelle stratégie produit les gains les plus élevés* sans que jamais n'intervienne une évaluation portant sur la pertinence, la justesse ou la conformité des conduites », il affirme « qu'une analyse satisfaisante de la formation d'une convention ou d'une norme sociale doit intégrer l'existence de jugements qui ne portent pas uniquement sur *l'utilité respective des stratégies*, mais qui en évaluent également la «normalité» » (p. 181, souligné par nous). Cette assimilation du gain d'une stratégie et de son utilité est plusieurs fois répétée dans le texte. Or, elle pose problème dans la mesure où l'utilité d'une stratégie n'est pas réductible à son gain, comme d'ailleurs Orléan finit par le reconnaître indirectement. Reprenant le modèle de Akerlof (1980), il est, en effet, conduit à distinguer « l'utilité classiquement prise en compte par les économistes » et « une utilité de nature spéciale qui modélise les effets sociaux et psychologiques particuliers provoqués par la norme » (p. 192). Retenons simplement, ici, que la critique de Orléan porte sur « l'utilité des économistes », en tant que celle-ci constitue une conception très réductrice du comportement humain, ramenant tous ses mobiles à la recherche du gain le plus élevé.

### **2.1.2. La critique de l'action rationnelle**

Si la théorie visée par Salais n'est plus la TJE mais la théorie du choix social, l'objet de sa critique demeure la notion de rationalité. En visitant à nouveaux frais les grands classiques que sont en la matière Buchanan, Hayek et Arrow, il est ainsi amené à conclure que « le débat État-marché (...) fait apparaître que pour progresser sur le chemin de l'institution, il faut s'éloigner de l'hypothèse standard de l'action rationnelle » (p. 263). Et si Rawls représente, sur ce chemin, un point d'ancrage possible, c'est qu'il réintroduit les valeurs collectives de la coordination économique, en faisant de celles-ci « non un instrument supplémentaire de manipulation des autres dans la coordination, mais une référence mutuellement attendue des actions, substituant ainsi à l'action rationnelle l'action douée de raisonabilité » (p. 257).

---

<sup>18</sup> Nous rappellerons que sur leur rapport à la TJE, les recherches de Orléan manifestent sinon une ambiguïté, du moins une dualité ; en effet, comme il le rappelle lui-même, certains de ses travaux appartiennent à la TJE, ceux notamment sur la convention fordiste.

Plus loin, Salais précise le sens de cette critique : « en refusant l'utilité néoclassique comme mesure sociale de toute action, Rawls rompt avec la rationalité économique standard. Mais il ne sacrifie pas pour autant à l'altruisme intéressé ou à l'irrationalité, ni ne fait de l'institution le garde-fou qui enchâsse l'opportunisme. Ce faisant, Rawls a franchi la grande ligne de partage, celle qui sépare ceux qui en restent à la rationalité individuelle opportuniste de ceux qui tiennent pour acquis la présence dans l'action individuelle d'une référence générale » (p. 265). En dressant une telle ligne de partage, Salais fait tomber du côté de la rationalité standard des auteurs qui, habituellement, relèvent plus ou moins d'une démarche hétérodoxe en économie ; c'est le cas de « Elster ou Granovetter, sans parler de Williamson » (n. 13, p. 266).

### **2.1.3. La critique du modèle du décideur**

Si, dans son article, Favereau prend pour cible, comme à son habitude, les théories standard et standard étendue, il change ici sensiblement l'angle de sa critique. Et sous l'angle de l'IND, c'est l'aspect monolithique du décideur rationnel qui apparaît poser problème : « étant donné qu'il n'y a qu'un seul modèle de décideur (rationnel), ou bien on va d'abord modéliser l'agent S, ou bien on va partir de l'agent I, mais chaque fois il s'avérera impossible de (re)constituer l'autre agent. Dans le premier cas, l'agent I va manquer à l'appel ; dans le second, on ne parvient pas à faire une place à l'agent S » (p. 155).

Désormais, ce n'est plus à l'aune de l'incertitude que se mesure la faiblesse du modèle néoclassique de l'agent rationnel, mais à celle des « situations ». Le contenu de la critique s'en trouve modifié : « à l'issue de ce rapide diagnostic des tensions antagonistes suscitées par l'intégration des situations dans le programme de recherches de l'individualisme méthodologique, le pronostic est que va s'estomper la figure romantique du décideur, seul face à la nature, triomphant, grâce à sa puissance de calcul, de l'obscurité et de l'adversité des temps à venir » (p. 164). La faiblesse de ce modèle est donc d'avoir développé une conception naturaliste de la rationalité de l'individu, oubliant que ce dernier est toujours en situation sociale, pour préciser un peu les choses. Ainsi, si la critique porte toujours sur l'incomplétude du modèle néoclassique, elle change de sens ; elle ne vise plus son incapacité à considérer l'incertain radical, mais son incapacité à considérer l'homme en société après l'avoir figé dans une relation « autiste » à la nature.

Les trois critiques de l'utilité des économistes, de l'action rationnelle et du décideur rationnel n'en font finalement qu'une seule ; elles instruisent toutes les trois le procès de la rationalité instrumentale au motif de son incapacité à considérer des jugements en vue de l'action, qui ne seraient plus seulement liés à ses gains, mais aussi à son mode social d'existence.

## **2.2. La dimension collective des contextes d'action individuelle**

Si les entrées dans l'institutionnel ne sont pas amorcées par les mêmes objets théoriques, elles n'en conduisent pas moins le nouveau programme de l'Économie des Conventions à toujours référer les institutions à la dimension collective des mondes dans lesquels elles interviennent. C'est là le point de départ commun aux analyses du nouveau programme de recherches de l'Économie des Conventions. Cette dimension collective reste toutefois peu développée chez Orléan où il est seulement souligné que les « croyances collectives » et les « sanctions » sont parmi les mécanismes

fondamentaux à la base des institutions<sup>19</sup>. A l'inverse, la dimension collective des contextes d'actions individuelles fait l'objet d'une enquête plus approfondie, chez Favereau aussi bien que chez Salais.

### **2.2.1. Le retour de l'acteur collectif**

C'est la question de la modélisation de l'IND<sup>20</sup> que se pose au départ Favereau qui l'amène finalement à prédire que « la théorie économique pourrait être conduite à réévaluer le rôle des institutions dans la coordination des activités et des comportements » (p. 153). Pour saisir ce nouveau rôle attribué aux institutions, il est nécessaire de comprendre ce qui se cache derrière l'IND.

L'IND désigne la relation imbriquée qui existe entre un décideur S de niveau supérieur et un décideur I de niveau inférieur. Dans cette relation, si le niveau supérieur est logiquement antérieur au niveau inférieur, ce dernier fait néanmoins retour sur le premier, à l'image de la relation de pouvoir décrite par Foucault (1984)<sup>21</sup>. Il est alors nécessaire, pour saisir la réalité de l'IND, de sortir du cadre de la théorie dominante pour laquelle la distinction des niveaux est soit méconnue, soit posée comme exogène. Cette sortie s'accomplit, chez Favereau, par l'intermédiaire d'un détour empirique qui se révèle productif dans la mesure où il « fait émerger la notion de "situation" *en tant qu'ensemble de déterminants contextuels de la position du décideur* » (p.161). En effet, cette notion rend possible « l'interrogation de la position même du décideur » (p. 160). Toutefois, en dépit d'un effort de définition<sup>22</sup>, la notion de situation reste assez floue jusqu'à ce que Favereau revienne sur le point décisif de son analyse, à savoir la possibilité d'endogénéiser la position S ou I du décideur : « Je parle ici de situations incluant un doute sur la position même de décideur. C'est par conséquent la théorie de l'acteur collectif qui est évoquée par cette référence à autrui : les agents S se différencient des agents I, en ce qu'ils participent d'un même acteur collectif qui reconnaît aux agents S des droits sur le domaine des actions possibles des agents I » (p. 164). Avec cette mise en forme de la notion de situation, nous glissons en fait dans le champ des institutions, « en entendant par là des systèmes de règles associés à des entités collectives clairement identifiables et généralement sanctionnés par le Droit » (p. 165). L'enjeu est alors de développer, enfin, une « vision positive » des institutions en économie.

### **2.2.3. Le bien commun**

La question du choix social étant une voie critique ancienne, Salais ajoute que, pour être une voie d'accès aux phénomènes institutionnels, elle doit sortir de son alternative habituelle qui oppose le marché à l'État. C'est seulement à cette condition qu'il devient possible de poser la question de savoir s'il y a « dans l'action individuelle présence de référents généraux tels le bien ou le juste » (p. 263), ce qui finalement conduit à la rencontre des institutions. En effet, une réponse positive à cette première question nous introduit dans le champ des institutions, par le relais d'une seconde question

---

<sup>19</sup> La raison de ce faible développement réside en partie dans l'absence claire de distinction entre les croyances intersubjectives et les croyances collectives, que nous avons pour notre part abordée dans Defalvard (1999). Là s'enracine, sans doute, la tentation chez Orléan de réduire le fondement des normes aux gains qu'elles permettent d'obtenir (voir plus bas).

<sup>20</sup> Interdépendance des niveaux de décision.

<sup>21</sup> Cité par Favereau, op. cit., p. 154-155.

<sup>22</sup> « Je définirai donc désormais le concept de situation comme appariement des caractéristiques de l'environnement interne d'un décideur avec celles de son environnement externe, appariement *indexé par l'ensemble des mondes possibles accessibles à partir du monde actuel* » (p. 162).

qui en découle. Car, « s'il en est ainsi, comment s'actualisent dans la situation d'action collective ces référents généraux. C'est le problème des institutions et de leur saisie en contexte d'action » (p. 263).

Il est alors intéressant de suivre la bifurcation que fait connaître Salais aux analyses de Rawls, après les avoir longtemps suivies, afin de prendre vraiment la direction des institutions, sans retomber dans leur réduction contractualiste. Si cette réduction est présente chez Rawls, celui en tout cas de *The Theory of Justice*, c'est que, dans sa construction, « les biens poursuivis en société ne sont que des biens privés et non des biens communs » (n. 17, p. 271)<sup>23</sup>. C'est, en définitive, la considération du bien commun qui ouvre, chez Salais, la route aux institutions : « Le propre de l'institution est, par ses règles, et l'interprétation en cours d'action à laquelle elles donnent lieu, d'autoriser l'engagement de ce processus de création de la "situation" (...) en traçant un corridor défini par le bien commun engagé dans la situation » (p. 280).

Quelle que soit l'entrée dans l'institutionnel retenue, elle conduit toujours à inscrire l'action individuelle dans un monde dont sa dimension collective est ce qui le définit. Comment alors, dans ces mondes à dimension collective, l'action individuelle est-elle modélisée ?

### 2.3. Modélisation de l'action individuelle

Cette dernière tranche du programme de l'Économie des Conventions correspond au chantier théorique ouvert à la suite de son inflexion vers les institutions. De loin la plus difficile à mener, en imposant de concevoir comment la dimension collective et le jeu des institutions orientent les actions individuelles, son état d'avancement est très différent selon les textes. Elle reste ainsi à l'état d'esquisse chez Favereau où il est seulement reconnu « la part sociale de l'identité individuelle » (p. 163), renvoyant à des travaux sur les groupes menés ailleurs qu'en économie<sup>24</sup>. Si Orléan, en reprenant le modèle de Akerlof (1980), dépasse ce stade de l'esquisse, son analyse pose le problème de la réduction individualiste de la norme sociale et, donc, du retour à la case départ. Enfin, nous trouverons, chez Salais, les développements les plus avancés en ce qui concerne la conceptualisation de l'action individuelle en société, plus précisément en présence d'institutions.

#### 2.3.1. Le risque de l'instrumentalisation des normes

Après s'être appuyé sur les travaux de Elster (1989) sur les normes, Orléan reprend le modèle développé par Akerlof sur la « *social custom* », en vue d'offrir une première modélisation du rôle des normes sociales. Cette dernière passe par une nouvelle écriture de la fonction d'utilité individuelle, qui ajoute aux déterminants habituels des « *nonindividualistic arguments* »<sup>25</sup>. En allant à l'essentiel, la nouvelle fonction d'utilité d'un travailleur s'écrit :

$$U = V - d^R m\bar{R} - d^R d^C \bar{C}$$

---

<sup>23</sup> Dans Defalvard (1992), nous avons déjà souligné le rôle de cette absence de prise en compte du caractère commun des biens, caractéristique de l'approche individualiste de la coordination sociale.

<sup>24</sup> Ceux de Turner (1987) en particulier.

<sup>25</sup> Akerlof, 1980, p. 755. Nous retrouvons la démarche usuelle de Akerlof, introduisant dans le modèle walrasien des marchés, des éléments sociaux au risque d'une difficile cohérence de l'ensemble.

où  $V$  est « l'utilité des économistes » qui réfléchit les préférences du travailleur relatives à ses consommations,  $d^R$  une variable muette valant 0 si le travailleur obéit à la norme du groupe<sup>26</sup>, et 1 s'il désobéit ;  $d^C$  est égal à 1 si l'individu croit au bien fondé de la norme et égal à 0 sinon ;  $m$  indique la proportion de ceux parmi les travailleurs qui croient à la norme et est donnée à court terme, auquel nous nous limiterons ; enfin,  $\bar{R}$  et  $\bar{C}$  sont des paramètres positifs. L'étude détaillée de cette fonction d'utilité permet de montrer que ses deux premiers termes,  $V$  et  $d^R m \bar{R}$ , se réduisent à l'intérêt, alors que le troisième terme ne s'y réduit pas, en tout cas pas nécessairement.

Si  $V$ , par définition de l'utilité des économistes, se réduit à l'intérêt, qu'en est-il du deuxième terme, c'est-à-dire du comportement d'obéissance envers la norme ? Orléan nous l'explique en ces termes : « un individu qui viole la règle ( $d^R = 0$ ) subit une perte d'utilité au travers des effets réputationnels négatifs qu'engendre sur sa personne le non-respect du code social. Par exemple, beaucoup de personnes choquées par sa conduite n'accepteront plus de transacter avec lui. Pour ces raisons l'activité économique de l'individu en question sera rendue difficile, *ce qui réduira d'autant ses gains* » (p. 193, souligné par nous). Sans l'ombre d'un doute, le respect de la norme dépend ici entièrement du gain net qu'il permet de réaliser ; le deuxième terme de  $U$  traduit une instrumentalisation de la norme que le troisième terme permet d'éviter, au moins potentiellement.

Lorsqu'un individu croit au bien fondé de la norme ( $d^C = 1$ ) et qu'il viole celle-ci ( $d^R = 1$ ), le terme  $d^R d^C \bar{C}$  mesure le « coût intérieur d'une transgression, intimement lié à notre culture » (p. 193). La norme ici n'est plus nécessairement instrumentalisée ; plus exactement, elle ne l'est que dans le cas d'une « culture » où le « coût intérieur » se confond avec la perte de gain liée au non-respect de la norme ; une culture donc où les normes et les règles ne sont plus fondées que sur les gains individuels que leur respect permet d'obtenir<sup>27</sup>. Ou encore où les croyances n'ont plus pour fondement que l'intérêt qu'elles permettent d'obtenir.

C'est du point de vue d'une telle culture que semble se placer Orléan lorsqu'il conclut que : « Puis les individus maximisent leur utilité et choisissent d'obéir à la norme, soit de lui désobéir, *en fonction du meilleur gain* » (p. 193, souligné par nous). Cela explique sans doute que lorsqu'il étend le modèle de Akerlof à des situations de type dilemme du prisonnier, seul l'effet réputationnel est modélisé, dans l'abandon du troisième terme de la fonction d'utilité. Ainsi, si comme Orléan nous pensons que l'utilisation de la fonction d'utilité peut permettre d'aboutir à des résultats intéressants, convient-il encore de ne pas réduire l'utilité à « l'utilité des économistes » ; piège que ne réussissent pas toujours à éviter ni Akerlof, ni Orléan<sup>28</sup>, et dont nous essaierons de sortir dans la troisième partie de ce papier.

### **2.3.2. De l'action individuelle dans le cours institué d'une action collective**

---

<sup>26</sup> Dans le modèle de Akerlof, cette norme consiste pour les travailleurs à se référer au salaire équitable fixé par le groupe des pairs, et non pas au salaire d'équilibre walrasien.

<sup>27</sup> Nous retrouvons ici l'opposition entre deux cultures ou entre deux éthiques : celle humaniste, où la morale est attachée à l'action désintéressée accomplissant la loi pour elle-même dont *Les fondements de la métaphysique des mœurs* de Kant offre l'illustration, et celle utilitariste où la morale ouvre sur des actions dont le fondement ultime est l'intérêt, accomplissant la loi par intérêt (cf. Ferry, 2000, p. 72-73).

<sup>28</sup> Ainsi, notre critique de modèle du « *fair wage* » de Akerlof et Yellen (1990) portait-elle sur la réduction, au total, du groupe au marché dans la mesure où le critère d'appartenance au groupe y était la productivité marginale du travail, telle qu'elle s'exprime dans le cadre du marché walrasien, à l'équilibre de la firme (voir Cordonnier et Defalvard, 1994, p. 128-129).

Avec les analyses de Salais, nous entrons plus encore dans le vif du sujet, dans la mesure où elles se centrent sur « l'articulation entre la personne et l'institution » (p. 257). Leur point de départ est l'hypothèse selon laquelle « les individus ont comme prémisses de leurs actions de savoir qu'ils vivent en société (et qu'ils contribuent ainsi à la fabriquer) » (p. 264) ; plus précisément encore « qu'elles (les personnes) savent qu'elles vivent en société et non seules » (p. 266). Ce « savoir vivre en société » va nous conduire à explorer la question si difficile de l'articulation entre l'individu et l'institution.

En allant directement au point d'arrivée, nous pouvons dire que ce « savoir vivre en société » apparaît au niveau des attentes des personnes où figure l'accord social grâce aux institutions qui, « par le fait d'être fondées sur cet accord, assurent à chacun qu'il est bien présent » (p. 266). Mais en allant directement au point d'arrivée, on donne l'impression malheureuse que les institutions tombent du ciel pour venir garantir, à travers les attentes des personnes, la présence de la société. La seule manière pour que ne tombent du ciel, pleinement constituées, ni les personnes ni les institutions, est de les aborder dans une démarche pragmatique, en contexte, où elles se révèlent ensemble. C'est ici le sens de la critique que Salais adresse à Rawls qui « n'intègre pas les épreuves de réalité auxquelles les institutions sont soumises dans l'action économique et qui développent à leur égard des capacités de jugement chez les acteurs » (p. 271). Nous sommes ainsi invités à abandonner la perspective trompeuse de la genèse des « bonnes institutions », raisonnant à partir de l'individu rationnel, pour ne plus considérer les institutions que comme « *practice* ».

En contexte, une action individuelle intègre ses différents moments : « délibération, décision, réalisation dans le cours d'une action collective » (p. 284). Autrement dit, si une personne, engagée dans le cours d'une action, sait qu'elle vit en société, c'est que le cours de ses actions participe d'une action collective. Cette dernière est alors caractérisée par le fait qu'elle engage une relation au bien commun où ce dernier, sans être un *a priori*, sous-tend l'action collective. De manière plus précise encore, le savoir vivre en société des personnes relève de capacités qui « ne se forment que dans l'apprentissage de la participation à la réalisation de biens communs. La prémisses de telles capacités fonde l'institution » (p. 289).

Loin de tomber du ciel platonicien des Idées, les institutions reposent sur des capacités dont la formation s'ancre dans l'apprentissage d'un agir ensemble pour un bien commun. En tant que règles, elles sont alors ce par quoi le bien commun est présent dans le cours singulier des actions individuelles.

Reprenant nombre des éléments du nouveau programme de l'Économie des Conventions, notre perspective des jeux à institutions se propose de développer plus en avant sa dernière tranche.

### **3. COORDINATION DES ACTIONS ET INSTITUTIONS**

En modifiant quelque peu le formalisme de la théorie des jeux, il est possible d'offrir une modélisation des actions individuelles en société, intégrant la part des institutions dans les conduites individuelles. À cette fin, nous commencerons par construire un jeu à institutions avant d'indiquer les équilibres culturels qui peuvent lui être associés.

#### **3.1. Les jeux à institutions**

Comme le remarque Salais, ce ne sont ni les définitions, ni les approches des institutions qui font défaut. Pour notre part, en nous appuyant sur les travaux récents de Searle (1995), nous développerons une approche institutionnaliste de la coordination sociale en termes de jeux, selon une démarche plus proche de la tradition de von Neumann et Morgenstern (1944) que de celle du programme de Nash<sup>29</sup>.

### 3.1.1. Les institutions en pratique

Nous partirons de la définition des institutions proposée par Searle (1995), en la transcrivant de la manière suivante :

*Une institution est un dispositif symbolique par lequel une entité X est comptée collectivement comme un Y dans un contexte C, où Y est une fonction-statut à laquelle sont attachés collectivement un ensemble de droits, négatifs et positifs.*

Cette définition reprend en fait de nombreux éléments du nouveau programme de recherches de l'Économie des Conventions, tout en permettant de dépasser certaines des difficultés que nous avons rencontrées à son propos.

Le premier point d'accord concerne l'approche pragmatique des institutions, en contexte ou en situation. Selon celle-ci, il n'existe pas d'institutions en général mais des institutions en particulier qui peuvent faire l'objet d'une schématisation, d'une abstraction ; ces dernières n'ayant aucune existence, mais seulement une essence. Le déficit de l'analyse économique en matière de recherches empiriques sur l'interaction sociale, que Manski (2000) a récemment déploré, trouve en partie son explication dans le fait que l'analyse économique accorde une existence aux abstractions qu'elle étudie ; elle s'y limite alors assez naturellement au prix d'une mauvaise métaphysique.

Le deuxième élément commun concerne la dimension collective des mondes dans lesquels les institutions interviennent : X est compté *collectivement* comme un Y, qui est une fonction-statut à laquelle sont attachés *collectivement* des droits et des devoirs. Pour Searle, de manière

très claire, les institutions ne créent pas *ex nihilo* du collectif, mais elles interviennent dans un monde dont le cours est déjà collectif. Qu'est-ce donc que ce « collectif » que les institutions transforment en droit ? À la lecture de *La construction sociale de la réalité*, nous dirons que ce « collectif » est un mode d'intentionnalité en acte, propre aux situations de coopération où les individus agissent ensemble dans un but commun : une intentionnalité collective où le sujet de l'action est ce « nous » lorsque nous agissons ensemble dans un but commun<sup>30</sup>. Les institutions introduisent alors, dans le monde de la coopération, une nouveauté radicale : celle du symbolique qui se traduit par la capacité à compter une entité X comme un Y, non pas en vertu des propriétés physiques de X mais d'un « accord collectif ». Ainsi, lorsque les 35 heures deviennent le nouvel horaire hebdomadaire légal, ce n'est pas en vertu d'une propriété physique des 35 heures mais à la suite d'un acte institutionnel dans

---

<sup>29</sup> En effet, le programme de Nash, en donnant la prééminence aux jeux non coopératifs, modélise les actions d'individus « *completely free* » (Nash, 1953), alors qu'à l'inverse la voie ouverte par *Theory of Games and Economic Behavior* comprenait les comportements individuels en termes de « *standards of behavior* », qui réfléchissaient les mondes sociaux des individus. Un des points de retournement de la théorie des jeux en faveur de la tradition de Nash est sans aucun doute l'article de Harsanyi (1966).

<sup>30</sup> Dans leur approche des institutions, Denzau et North (1994) tendent à réduire celles-ci à leur dimension cognitive, en les comprenant seulement comme des « *shared mental models* » (p. 4).



un contexte C, visant à ouvrir des droits aux entités X (qui sont toujours en définitive des individus) reconnues faire (ou faire faire) l'horaire légal.

Le troisième point commun porte sur la difficile question de l'articulation entre le micro et le macro. En général, cette articulation est comprise comme le passage du niveau des actions individuelles au niveau de la totalité de ces actions, qu'un holisme vulgaire mais répandu a tendance à traiter comme une action individuelle. Dans notre perspective institutionnaliste, l'articulation devient celle entre le niveau macro-institutionnel où les actions individuelles produisent des institutions et le niveau micro-institutionnel où les actions individuelles opèrent dans un contexte institutionnel donné ; le niveau macro est ainsi celui des acteurs institutionnels dont les actions portent sur les règles alors que le niveau micro est celui des acteurs tout court qui agissent dans le cadre de règles données. Cette distinction ne recoupe pas totalement celle entre les acteurs S et les acteurs I de Favereau, car un employeur S qui a autorité sur son employé I n'est pas, comme tel, un acteur institutionnel. Par contre, lorsque le principal écrit le contrat de travail dans un modèle de type principal-agent, c'est un acteur institutionnel dont le pouvoir est tel qu'il comprend toutes les actions des acteurs dans le contrat ; il est ainsi en position de dictateur. Enfin, si le niveau macro-institutionnel est logiquement antérieur au niveau micro-institutionnel, le niveau des actions des acteurs est chronologiquement antérieur, car il n'y a pas d'institutions « sans société » au préalable, sans une situation de coopération<sup>31</sup>.

Notre approche pragmatique des mondes institutionnels laisse ouvert le passage entre micro et macro, et dans les deux sens : si le cours des actions individuelles est toujours pris dans un réseau d'institutions, les institutions se transforment à leur tour dans le cours des actions. La question n'est donc pas celle du passage d'un monde sans institution à un monde avec des institutions optimales, mais plutôt celle de la dynamique des institutions. Nous sortons ainsi de la dichotomie néoclassique, dont le mode ancien opposait marché et État, et dont le mode actuel oppose jeux et Éthique selon la terminologie de Harsanyi (1966)<sup>32</sup>.

Il nous reste maintenant à traduire sans les trahir ces développements sur les institutions dans le cadre des jeux. Pour cela, nous modifierons ce dernier dans une direction qui nous paraît rejoindre l'inspiration des travaux pionniers de von Neumann et Morgenstern (1944).

### **3.1.2. Les jeux à institutions**

Commencée dans des travaux antérieurs (Defalvard, 1999 et 2000b), nous poursuivrons ici notre construction des jeux à institutions en distinguant la matrice des paiements au niveau macro-institutionnel et la matrice des gains au niveau micro-institutionnel.

Reprenant l'exemple, issu de Rousseau (1755), du jeu de la chasse au cerf, nous considérerons deux individus ( $i = 1, 2$ ), dont l'action coopérative se compose de l'action individuelle de rabattre le cerf, notée A, et de l'action individuelle de le guetter au poste, notée B. La dimension collective de

---

<sup>31</sup> Si l'institutionnel, au sens défini ci-dessus, est le propre des actions humaines, le mode coopératif ou social des actions individuelles ne l'est pas, comme nous l'apprennent toujours davantage les travaux sur les sociétés animales (cf. Joulain, 1998 ou Thierry, 1998).

<sup>32</sup> Signalons ici les travaux de Ponsard et alii (1996), qui participent de cet effort de nouer ensemble le niveau macro et le niveau micro, en renversant la perspective usuelle de la théorie des jeux. On regrettera que cet effort soit limité aux questions d'information, et donc de coordination, sans inclure les institutions. Dans une perspective très différente, les travaux de Baron et alii (1998) sur l'insertion contribuent également au traitement de cette question à partir de leur notion de finalité.

l'action individuelle de rabattre le cerf se traduit par l'écriture  $A_i/B_j$  ( $i \neq j$ ) : celle-ci exprime le point de vue de l'individu  $i$  sur son action de rabattre le cerf en tant qu'elle appartient à une action collective, que l'on pourrait transcrire par : « nous chassons le cerf ensemble, où je (i) rabats le cerf pendant que l'autre (j) le guette à son poste ».

Les règles instituant la chasse au cerf attribuent alors aux différentes actions individuelles des paiements  $P_i$ , dont voici la matrice à ce niveau macro-institutionnel (le premier nombre dans chaque parenthèse indique le paiement de l'individu 1 jouant en ligne, alors que le second indique celui de l'individu 2 jouant en colonne)<sup>33</sup> :

*Figure 1 : Le jeu macro-institutionnel*

	$A_2/B_1$	$B_2/A_1$
$A_1/B_2$		(4,6)
$B_1/A_2$	(6,4)	

Nous admettons ici que les règles instituant la chasse au cerf imposent à l'individu 1, en raison de sa position dans le système de parenté, le rôle du rabatteur alors qu'elles imposent à l'individu 2 le rôle du guetteur. Cette convention, au niveau macro-institutionnel, ne revient pas à choisir un équilibre de Nash dans une situation à équilibre multiple, car il ne s'agit pas ici d'un problème de pure coordination, mais d'une coordination sociale où l'institution ne permet pas le calcul des équilibres de Nash, en rendant certaines des cases de la matrice vides.

En situation de chasse, chaque individu a, en plus de ce que prévoit l'institution, la possibilité ou non de tirer un lièvre en sachant qu'au moins un lièvre passera à portée de fusil de chacun des individus. Cette institution n'est donc pas totalitaire, en ce sens qu'elle ne comprend pas la totalité des actions individuelles, comme c'est le cas des contrats optimaux. L'action individuelle de tirer le lièvre sera notée  $C_i$ . Au niveau micro-institutionnel, chacun a ainsi le choix entre chasser le cerf selon les règles du jeu et la convention ou chasser le cerf en tirant également le lièvre. Les gains  $G_i$ , au niveau micro-institutionnel, sont donnés par la nouvelle matrice suivante<sup>34</sup> :

*Figure 2 : Le jeu micro-institutionnel*

	$B_2/A_1$	$C_2$
$A_1/B_2$	(4,6)	(2,7)
$C_1$	(7,2)	(3,3)

Avant de considérer les équilibres de ce jeu, il est essentiel d'introduire une distinction entre, d'un côté, les gains du jeu micro-institutionnel et, de l'autre côté, les comportements des individus. Cette distinction n'est le plus souvent pas faite par la théorie des jeux, confondant les paiements associés à une stratégie et l'utilité de celle-ci.

<sup>33</sup> Nous sommes très loin de la situation imaginée par Rousseau où la chasse au cerf se déroulait avant toute institution, à l'état de nature (voir Defalvard, 1999).

<sup>34</sup> Les nouveaux gains témoignent que si pour chacun, qu'il soit guetteur ou rabatteur, le gain de chasser le cerf sans tirer le lièvre si l'autre le fait aussi est plus élevé que celui résultant de la situation où chacun tire le lièvre, le gain de tirer le lièvre est pour chacun plus élevé que celui de chasser le cerf sans tirer le lièvre, quelle que soit l'action de l'autre.

### 3.2. Les équilibres culturels du jeu

Contrairement à Harsanyi (1966), nous ne pensons pas qu'il existe un comportement formel universel, donné par la théorie de la décision dans le cadre de son individu isolé, et susceptible de se décliner sous des formes singulières diverses. Il n'existe que des formes singulières de comportement, c'est-à-dire en contexte. Et ces formes varient selon les contextes. Pour ne pas compliquer inutilement nos développements, on considérera que les deux individus peuvent être de deux types, I ou II. Lorsque l'individu  $i$  est de type I, son comportement est rendu par le programme suivant :

$$\text{Max } G_i$$

L'individu  $i$  de type I se conduit donc de telle sorte que son gain soit maximum. Lorsque l'individu  $i$  est de type II, son comportement est rendu par le programme suivant :

$$\text{Max min } G_h \quad h = 1,2$$

L'individu  $i$  de type II se conduit selon le principe du maximin<sup>35</sup>, revenant à choisir l'action qui assure le gain le plus élevé à l'individu ayant le gain le plus faible.

Nous définirons alors un *équilibre culturel* comme la situation qui émerge d'un jeu lorsque les joueurs sont de même type, et constituée des actions individuelles que chaque individu sélectionne à partir de son mode de comportement. Sur la base de la matrice des gains définie au niveau micro-institutionnel, il existe deux équilibres culturels. Le premier est l'équilibre culturel intervenant entre deux individus de type I ; il correspond à la situation où chaque individu chasse le cerf en tirant le lièvre car, pour chaque joueur, cette stratégie est alors dominante. Le second est l'équilibre culturel intervenant entre deux individus de type II ; il correspond à la situation où chacun chasse le cerf en ne tirant pas le lièvre. Il est important de remarquer que l'on aboutit ici à cet équilibre culturel en raison de l'institution qui garantit à chaque joueur qu'il participe à une chasse au cerf selon les règles, et à la croyance collective de chacun en l'institution<sup>36</sup>.

Si les deux individus chassant le cerf ne sont pas de même type, aucun équilibre culturel, par définition, ne peut intervenir entre eux. Dans l'exemple, les situations émergeant dans ce cas traduisent une situation de crise des institutions : la croyance en l'institution est remise en cause pour l'individu de type II qui agit selon cette croyance. La dynamique des institutions est alors synonyme de crise des institutions, dont l'origine mêle le changement des croyances et le changement des situations à l'épreuve du temps.

Pour finir, nous devons ajouter que les types de comportement ne sont pas indépendants des contextes institutionnels, bien que la relation entre les deux ne soit pas une bijection appariant chaque contexte institutionnel avec un seul type de comportement. Il y a plutôt congruence entre un contexte institutionnel et un type de comportement<sup>37</sup> ; cette congruence est assurée en grande partie par les

---

<sup>35</sup> Le maximin n'est pas ici, à la différence de son utilisation par Rawls, un principe de sélection des bonnes institutions, jouant donc au niveau macro-institutionnel ; il est une modalité de comportement au niveau micro-institutionnel parmi d'autres possibles.

<sup>36</sup> Sur les croyances collectives ou de groupe, on pourra se rapporter à Bar-Tal (1999).

<sup>37</sup> Ce constat a été fait par l'économie expérimentale des jeux (cf. Guth et alii, 1998). Il a été obtenu comme un résultat dans le modèle de Fehr et Schmidt (1999) sur la coopération, où ces derniers utilisent aussi une fonction de comportement faisant

croyances dont sont porteuses les institutions et à partir desquelles les comportements font sens. Ainsi, le type I de comportement est congruent avec un monde institutionnel où l'on croit à la main invisible d'Adam Smith, selon laquelle la meilleure façon de réaliser l'intérêt général est, pour chacun, de suivre son intérêt particulier, en cherchant à obtenir le gain le plus grand. À l'inverse, le type II est congruent avec un monde institutionnel où l'on croit qu'il est bien d'agir de sorte que la situation des plus faibles soit la meilleure. En posant clairement la distinction entre les paiements d'un jeu (dépendants des institutions en vigueur) et les types de comportement, nous avons pu envisager des comportements qui ne se réduisent pas au comportement nashien, sélectionnant les stratégies qui assurent le meilleur gain. Évidemment, les types de comportement en contexte ne se limitent pas à ceux retenus ici.

---

intervenir les gains de l'autre (sous la forme d'une différence) ; par contre, ce modèle en reste, conformément à la perspective individualiste, à un seul niveau d'action, le niveau micro-institutionnel.

## CONCLUSION

L'approche institutionnaliste de la coordination sociale, telle que la développe aujourd'hui l'Économie des Conventions et dans laquelle notre modèle de jeu à institutions s'est inscrit, relève-t-elle du holisme ou de l'individualisme ? Cette question ne veut pas rouvrir un débat souvent stérile, mais permettre de faire retour sur le problème de l'incomplétude de l'individu au terme de notre étude.

Loin d'introduire les célèbres totalités agissantes, connues sous les noms de l'État ou encore de l'Entreprise, dont on habille régulièrement le holisme afin de mieux s'en débarrasser comme on dit d'un chien encombrant qu'il a la rage, notre approche analyse la coordination sociale en référant celle-ci aux seules actions individuelles, au niveau micro-institutionnel comme au niveau macro-institutionnel. Cela en fait-il une approche individualiste ? À suivre Durkheim, dans son ouvrage *De la division du travail social* (1893), la réponse est clairement négative :

« En voilà assez, pensons-nous, pour répondre à ceux qui croient prouver que tout est individuel dans la vie sociale, parce que la société n'est faite que d'individus. Sans doute, elle n'a pas d'autre substrat ; mais parce que les individus forment une société, des phénomènes nouveaux se produisent qui réagissent sur les consciences individuelles, les forment en grande partie » (p. 342, n.2). Ainsi, la critique de Durkheim aux tenants de la méthode individualiste n'est pas incompatible avec la reconnaissance que les individus forment le substrat de la vie sociale. Au contraire, sa critique s'étaye d'abord sur un point d'accord avec eux, le conduisant finalement à écrire que : « C'est aux économistes que revient le mérite d'avoir les premiers signalé le caractère spontané de la vie sociale (..) et d'avoir montré qu'elle résulte, non d'arrangements extérieurs et imposés, mais d'une libre élaboration interne. À ce titre, ils ont rendu un important service à la science de la morale » (p. 380). Alors que ce trait du holisme durkheimien a été souvent occulté, il est cependant essentiel pour comprendre la nature de la critique que Durkheim adresse aux économistes. Celle-ci dénonce notamment la conception de la liberté individuelle des économistes, qui la considèrent comme « un attribut constitutif de l'homme, le déduisant logiquement de l'individu en soi, *abstraction faite de toute société* » (p. 380, souligné par nous). Le désaccord porte sur leur manière de considérer les actions comme étant celles « d'individus isolés et indépendants », car « d'individualités autonomes, comme celles qu'on imagine, il ne peut rien sortir que d'individuel, et par conséquent, la coopération elle-même, qui est un fait social, soumis à des règles sociales, n'en peut naître » (p. 263). Il apparaît ainsi que la pointe de la critique durkheimienne concerne la complétude de l'individu isolé, qui est à la base de toute l'économie néoclassique, selon l'architecture qu'en a proposée Harsanyi (1966). Alors que cette complétude de l'individu dans la théorie bayésienne de la décision laisse une place à l'incertitude (son environnement pouvant être risqué ou incertain), elle n'en laisse aucune à la relation aux autres ; l'individu isolé est une société à lui tout seul, auquel « l'action sociale n'a rien à y ajouter » comme le déplore Durkheim (p. 380). Contrairement à la thèse de Favereau, nous pouvons comprendre ici que ce n'est pas la méthode holiste qui impose d'assumer « l'assimilation des humains aux fourmis » (p. 162), mais bien la posture individualiste. En effet, les travaux récents sur les insectes sociaux ont montré que le mode de coordination de leurs actions était de même nature que celui des insectes solitaires. Dans les deux cas, les mécanismes de coordination reposent

sur les mêmes comportements d'émission de phéromone et de marquage, de nature physico-chimique<sup>38</sup>. Comme les insectes solitaires, l'individu isolé est une société à lui tout seul.

Pour l'approche institutionnaliste de la coordination sociale, l'individu isolé est une fiction trompeuse, voire dangereuse. L'individu est incomplet en tant que la société passe en lui, à travers l'apprentissage des règles sociales et sous la forme de croyances collectives donnant sens à ses actions situées dans les mondes collectifs au sein desquels il est plongé. Toutefois, le holisme que nous défendons n'interdit en rien de retrouver l'individu autonome, non plus comme le point de départ de la vie sociale, mais comme une production de celle-ci. En accord cette fois avec Favereau, il est donc possible d'écrire que l'intégration des institutions fait de l'individualisme une méthode « menacée, même si en un autre sens, la philosophie générale de l'individualisme semble définitivement validée » (p. 163). En effet, l'individualisme est à la fois, et sans que ceci soit une contradiction, menacé et validé. Là encore une référence à Durkheim n'est pas inutile afin de comprendre pourquoi. En effet les sociétés modernes, dans lesquelles « il (l'individu) change plus aisément de milieu, quitte les siens pour aller vivre une vie plus autonome » (p. 396), se caractérisent toujours, selon Durkheim, par une « conscience commune : ce culte de la personne, de la dignité individuelle dont nous venons de parler, et qui, dès aujourd'hui, est l'unique centre de ralliement de tant d'esprits » (p. 396). Ainsi, que l'individualisme soit menacé comme méthode d'investigation des faits de la vie sociale en raison de la nature institutionnelle de ceux-ci, n'obère donc en rien que l'individualisme soit la métaphysique des temps modernes, plutôt que leur « philosophie générale ». Plus, seule une méthodologie holiste est à même de l'établir, sans devoir pour cela réduire les hommes à des fourmis.

## BIBLIOGRAPHIE

- Akerlof G. (1980) : « A Theory of Social Custom, of which Unemployment May be one Consequence », *The Quarterly Journal of Economics*, juin, p. 749-775.
- Akerlof G. et Yellen J. (1990) : « The Fair Wage-Effort Hypothesis and Unemployment », *The Quarterly Journal of Economics*, mai, p. 255-283.
- Baron et alii (1998) : *Insertion : les feux follets de la démocratie*, éditions Charles Léopold Mayer.
- Bar-Tal D. (1999) : « Croyances, idéologie et construction du groupe », dans Deschamps J.-C. et alii, *L'identité sociale*, Presses Universitaires de Grenoble, p.43-67.
- Chiappori P.-A. (1994) : « Anticipations rationnelles et conventions », dans Orléan A., ss. la dir., *Analyse économique des conventions*, Puf, p. 61-77.
- Cordonnier L. et Defalvard H. (1994) : « L'hypothèse sociologique dans la théorie du salaire d'efficience : un examen de la méthode », dans *Les Cahiers Lillois d'Economie et de Sociologie*, n°23, p.109-132.
- Defalvard H. (1992) : « Critique de l'individualisme méthodologique revu par l'économie des conventions », dans *Revue Économique*, n°1, p. 127-143.
- Defalvard H. (1999) : « Coordination, anticipations et croyances : la part des institutions », dans *Économie Appliquée*, n°3, p. 7-39.
- Defalvard H. (2000a) : *La pensée économique néoclassique*, Les Topos, Dunod.

---

<sup>38</sup> Théraulaz et alii (1998, p. 89).

- Defalvard H. (2000b) : « Coordination sociale et croyances individuelles : à propos de quelques résultats récents en théorie des jeux », *L'Actualité Économique*, septembre (à paraître).
- Denzau A. et North D. (1994) : « Shared Mental Models : Ideologies and Institutions », dans *Kyklos*, n°1, p. 3-31.
- Durkheim E. (1986) : *De la division du travail social*, Quadrige, Puf.
- Eymard-Duvernay F. et alii (1999) : « Les institutions du recrutement. Approche comparative France/ Grande-Bretagne », *La Lettre du CEE*, n°60.
- Elster J. (1989) : *The Cement of Society*, Cambridge University Press.
- Favereau O. (1995) : « L'économie des conventions. Politique d'un programme de recherches en sciences sociales », dans *Actuel Marx*, n°17, p. 103-113.
- Favereau O. (1998) : « Décisions, situations, institutions », dans Vinokur A., textes réunis par, *Décisions économiques*, Economica, p.153-168.
- Fehr E. et Schmidt K. (1999) : « A Theory of Fairness, Competition, and Cooperation », dans *The Quarterly Journal of Economics*, August, p. 817-868.
- Ferry L. (2000) : « L'éthique évolutionniste et ses critiques », dans Ferry L. et Vincent J.-D., *Qu'est-ce que l'homme ?*, éditions Odile Jacob, p.69-112.
- Güth W. et alii (1998) : « Equilibration et dépendance du contexte », dans *Revue Économique*, mai, p. 785-794.
- Harsanyi J. (1966) : « A General Theory of Rational Behavior in Game Situations », dans *Econometrica*, vol. 34, n°3, p. 613-634.
- Joulian F. (1998) : « La transmission de la « culture » chez l'animal », dans *Pour la science*, n° 254.
- McLeod (1996) : « Decision, contract, and emotion : some economics for a complex and confusing world », *Canadian Economic Review*, n°4, p. 788-810.
- Munck de J. (1999) : *L'institution sociale de l'esprit*, puf.
- Manski C. (2000) : « Economic Analysis of Social Interactions », Working Paper 7580, *National Bureau Of Economic Research*, 41 p.
- Nash J. (1953) : « Two-Person Cooperative Games », dans *Econometrica*, 21, p. 128-140.
- Neumann von J. et Morgenstern O. (1944) : *Theory of Games and Economic Behavior*, Princeton University Press.
- Orléan A. (1994) ss. la dir. : *L'analyse économique des conventions*, puf.
- Orléan A. (1997) : « Jeux évolutionnistes et normes sociales », dans *Économie Appliquée*, n°3, p. 177-198.
- Ponsard J.-P. et alii (1996) : « Jeux d'équipe avec interface et coordination dans l'entreprise », dans *Revue Économique*, n°5, p.1063-1088.
- Postel N. (1998) : « L'économie des conventions : une approche instrumentale de la rationalité individuelle », dans *Revue Économique*, n°6, p. 1473-1495.
- Quéré L. (1993) : « A t-on vraiment besoin de la notion de convention ? », dans *Réseaux*, 62, p. 21-41.
- Rallet A. (1993) : « La théorie des conventions chez les économistes », dans *Réseaux*, 62, p. 43-59.
- Rousseau J.J. (1755) : *Le discours sur l'origine de l'inégalité parmi les hommes* (repris dans *Du Contrat social*, éditions 10/18, 1973).
- Salais R. (1998) : « À la recherche du fondement conventionnel des institutions », dans *Institutions et conventions, Raisons pratiques*, éditions de l'EHESS, p. 255-291.
- Salanié B. (1994) : *La théorie des contrats*, Economica.
- Searle J. (1995) : *The Construction of Social Reality*, éditions Free Press, New York.
- Sugden R. (1986) : *The Evolution of Rights, Cooperation, and Welfare*, Basil Blackwell.
- Theraulaz G. et alii (1998) : « Les insectes architectes ont-ils leur nid dans la tête ? », dans *La recherche*, octobre, n°313, p. 84-90.

Thévenot L. (1995) : « L'action en plan », dans *Sociologie du travail*, n°3, p. 411-434.

Thierry B. (1998) : « L'intelligence sociale des singes », dans *Pour la science*, n° 254.

Turner J.-C. (1987) : *Rediscovering the social group : a self-categorization theory*, Basil Blackwell, Oxford.

Vinokur A. (1998) : *Décisions économiques*, textes réunis par, Economica.

Young P.H. (1996) : « The Economics of Convention », dans *Journal of Economic Perspectives*, n°2, p.105-122.