

# *La complexité sociale et la critique hayekienne du 'constructivisme'*

Dans le cadre de l'atelier organisé par Kevin Mulligan  
Université de Genève 22-23 avril 2016

**Jean Petitot** (CAMS, EHESS, Paris)

## **INTRODUCTION**

Je suis très heureux de participer à cet atelier et d'avoir ainsi une nouvelle occasion de discuter avec Kevin. Depuis une trentaine d'année que nous nous connaissons, nous avons beaucoup discuté de phénoménologie, de sciences cognitives et d'épistémologie mais peu de philosophie politique. Cet atelier est donc une belle opportunité.

Cela dit, je ne suis pas un spécialiste de sciences politiques. J'ai surtout travaillé d'un côté en philosophie des mathématiques et de la physique et, d'un autre côté, en théorie des *systèmes complexes* (en particulier sur les modèles mathématiques en neurosciences cognitives). C'est dans ce dernier domaine que certaines de mes réflexions ont croisé les conceptions de Friedrich von Hayek qui a été un pionnier de la théorie des systèmes complexes non seulement en sciences sociales (économie, droit, politique) mais aussi, c'est moins connu, en sciences cognitives (son grand livre *The Sensory Order*, paru en 1952, a été en fait commencé dès 1920!).

Dans les années 2000, à l'époque où je dirigeais le CREA, un Centre de l'Ecole Polytechnique créé par mon collègue et ami Jean-Pierre Dupuy, spécialiste des modèles *d'auto-organisation* "du physique au politique", j'ai organisé avec Philippe Nemo de l'ESCP (European School of Management) un séminaire de quatre ans (2001-2005) sur *l'Histoire du libéralisme en Europe* dont les actes ont été publiés en 2006 aux PUF (1427 pages). Comme Kevin a abordé dans ses cours la critique du "constructivisme" développée par Hayek, nous nous sommes dit que je pourrais intervenir autour de ce thème.

Les autres intervenants dans l'atelier sont beaucoup plus compétents que moi en sciences sociales. Je me focaliserai ici uniquement sur des problèmes liés à la complexité.

Sous le nom de "constructivisme", Hayek a toujours critiqué la croyance que le volontarisme politique serait à même de planifier et d'organiser rationnellement de façon "top down" les dynamiques économiques et sociales en évitant les écueils de la glaciation bureaucratique. Selon lui, les ressources cognitives humaines (surtout celles des politiques)

sont beaucoup trop limitées pour qu'un tel volontarisme puisse être opératoire. Les dynamiques économiques et sociales relèvent en effet d'ordres spontanés, complexes et auto-organisés dont le contrôle rationnel est impossible.

Le cadre de cette conception est l'*individualisme méthodologique complexe* (IMC) et je commence donc par quelques remarques à ce sujet.

## I. L'INDIVIDUALISME METHODOLOGIQUE COMPLEXE

### I.1. Les paradigmes de l'ordre social

Hayek (1899-1992, Nobel d'économie 1974) s'inscrit dans un paradigme caractéristique de la modernité, celui d'un ordre social ouvert fondé sur un pluralisme et une liberté individuelle garantis par un Etat de droit. La thèse centrale en est que le pluralisme et la liberté ne sont pas des facteurs de désordre et de division sociale, mais au contraire une forme supérieure d'organisation sociale.

Ce paradigme s'oppose à trois autres paradigmes :

1. Le paradigme de *l'ordre hiérarchique*, celui de l'absolutisme, théorisé dès la Renaissance par Machiavel (1469-1527), puis par Bodin (1529-1596) et Hobbes (1588-1679) et réalisé, par exemple en Espagne, par Charles-Quint et Philippe II et, en France, par Richelieu, Louis XIV et Napoléon. C'est en réaction contre l'absolutisme que se développèrent les revendications libérales de tolérance et de droits de l'homme, de Grotius (1583-1645), Pierre Bayle (1647-1706) et Locke (1632-1704) à Kant (1724-1804), Humboldt (1767-1835) et Benjamin Constant (1767-1830). C'est de là que sont issues les révolutions hollandaise, anglaise, américaine et française (avant la Terreur). C'est de là qu'ont pris leur essor les sciences et les techniques, la révolution industrielle, la croissance économique.
2. Le paradigme de *l'ordre artificiel*, celui du constructivisme rationnel qui refuse la société ouverte au nom des idéaux d'égalité et de justice et vise la planification volontariste de la société. C'est surtout celui-là que critique Hayek.
3. Le paradigme de *l'ordre inné* qui refuse la société ouverte au nom d'un holisme organiciste et combat la modernité pour avoir "atomisé" la société par l'individualisme et détruit les "communautés naturelles" (famille, corporations, etc.).

Le paradigme de l'ordre spontané conçoit l'ordre social comme *ni* permanent et universel, *ni* artificiel et rationnellement construit, mais pluraliste et auto-organisé, non hiérarchique et polycentré. Des exemples évidents de tels ordres sont les langues naturelles : elles ne sont pas innées car sinon elles ne manifesteraient pas une telle diversité anthropologique, mais elles ne sont pas non plus artificielles car aucun individu n'en a jamais décidé. Comme disaient au XVIIIe les maîtres du Scottish Enlightenment David Hume (1711-1776) et Adam Ferguson (1723-1816), ce sont

“des ordres qui résultent des actions des hommes mais non de leurs intentions”.

Dans l'état de droit, l'état n'a pour fonction ni de maintenir un ordre inné ni d'imposer un ordre rationnel mais seulement de garantir les institutions qui sont les conditions de possibilité de l'émergence d'un macro-ordre social spontané, ouvert et évolutif.

Le rapport à la temporalité change complètement. Il n'y a plus de finalité dans l'histoire sociale mais une évolution sans fin impossible à anticiper et à prévoir.

## **I.2. Les diverses formes d'individualisme méthodologique complexe**

La question centrale de l'individualisme méthodologique complexe (IMC) est de savoir si les phénomènes sociaux peuvent ou non être expliqués à partir d'entités collectives macro-sociales et supra-individuelles antérieures aux individus et les transcendant, comme l'état, le marché, les églises, les classes, les nations, etc. L'existence substantielle et l'action causale de telles entités relève du *holisme*. Le holisme sera dominant pendant une bonne partie du 19<sup>ème</sup> siècle et prendra diverses formes.

(i) Chez Saint-Simon (1760-1825, *De la physiologie appliquée à l'amélioration des institutions sociales* : 1813) et chez son disciple Auguste Comte (1798-1857, *Système de politique positive* : 1851) il s'agira d'un “organicisme”, c'est-à-dire d'une approche “physiologique” du social qui s'oppose radicalement à “l'atomisme” et au “mécanicisme” issus des Lumières.

(ii) Avec Durkheim (1859-1917, *De la division du travail social* : 1893, *Règles de la méthode sociologique* : 1895, *Les formes élémentaires de la vie religieuse* : 1912) se développe un holisme non plus métaphoriquement “biologique” mais résolument sociologique. Les totalités sociales existent et subsistent et *déterminent* les comportements des individus empiriques. Certes il existe des interactions “horizontales” empiriques entre individus, mais la vraie causalité sociale est “verticale” et “top down”. Les actions et même les consciences individuelles sont l'expression de leurs déterminations sociales.

L'IMC s'oppose au holisme. Il récuse l'hypostase substantialiste des concepts totalisants. Sur le plan métaphysique, on peut le faire remonter au nominalisme médiéval de Guillaume d'Ockham (1288-1349) et à son conflit avec les réalistes à propos des universaux. Les ensembles sociaux sont des agrégats interactifs (compétitifs-coopératifs) d'individus singuliers et non pas des substances. L'IMC est *méthodologique* dans la mesure où il ne concerne pas l'ontologie mais l'explication, la modélisation, la reconstruction des phénomènes sociaux. Il est *complexe* dans la mesure où il n'est pas du tout un atomisme réductionniste: il repose sur les *interactions* multiples et sur les propriétés *émergentes* qui en découlent. Faisons deux citations à ce propos.

Karl Popper (*Misère de l'historicisme*, 1944, p. 198) :

“L'individualisme méthodologique (est) la doctrine tout à fait inattaquable selon laquelle nous devons réduire tous les phénomènes collectifs aux actions, interactions, buts, espoirs et pensées des individus et aux traditions créées et préservées par les individus.”

Jon Elster (*Karl Marx, une interprétation analytique*, p.19) :

“Par individualisme méthodologique, j'entends la doctrine suivant laquelle tous les phénomènes sociaux — leur structure et leur changement — sont en principe explicables de manières qui impliquent les seuls individus avec leurs qualités, leurs croyances, leurs objectifs et leurs actions.”

Les premiers classiques de l'IM sont bien connus.

- (i) John Locke (1632-1704). L'individu est premier mais toujours en relation originelle avec autrui dans une société contractuelle garantie par le droit (rule of law). Les lumières écossaises : David Hume (1711-1776) et Adam Ferguson (1723-1816) (déjà évoqués).
- (ii) J'aimerais m'arrêter un peu sur Bernard de Mandeville (1670-1733) qui, selon Hayek, “asked the right question” à propos des sociétés humaines. Mandeville était d'une famille de médecins libéraux et progressistes de Rotterdam qui, suite à des démêlées avec les orangistes et les calvinistes, fut banni et s'installa en Angleterre. Son célèbre apologue “*The Grumbling Hive: or, Knaves Turn'd Honest*” (La ruche murmurante, ou les fripons devenus honnêtes gens) (1705) aussi dit “*The Fable of the Bees; or, Private Vices, Publick Benefits*” (1714) eut un énorme retentissement.

La thèse est qu'une ruche ne fonctionne collectivement bien que lorsque chaque abeille individuelle, qui n'a que des ressources représentationnelles extrêmement limitées, fait ce qu'elle a à faire dans le cadre de la division du travail en suivant des règles prescrites (en ce cas innées) sans se préoccuper du bien-être collectif. Ce sont *les interactions et l'interdépendance* qui produisent la prospérité de la ruche et non pas les vertus morales altruistes d'abeilles individuelles qui, surestimant leurs ressources cognitives, se croiraient capables de se représenter la ruche alors que toute représentation de ce type leur est impossible. Bref, la fable porte sur l'utilité sociale de l'égoïsme dans les sociétés qui comportent une myriade d'individus et sont fondées sur la division du travail.

Mandeville explicite de façon provocante l'émancipation du libéralisme naissant par rapport à la morale religieuse et l'éthique chrétienne et la découverte depuis le XVIIe siècle du fait que “l'amour de soi” (comme on disait à l'époque) et “l'amour-propre éclairé” peuvent être socialement positifs. La thèse se trouve déjà chez Pascal et surtout chez Pierre Nicole (1625-1695, proche des jansénistes de Port-Royal, cf. ses *Essais de morale* de 1671) : pour réformer le monde il faut se fonder sur l'amour-propre éclairé et non pas sur la charité. C'est le débat théologique, religieux et moral sous-jacent au libéralisme des Lumières.

Mandeville a inspiré en partie Adam Smith. La prospérité des nations est guidée par un mécanisme de “main invisible” (thèse de l'auto-organisation) et, sur le plan moral, cette

“main invisible” fonctionne comme une “ruse de la raison”. Comme le disait déjà Pierre Nicole, faire concourir les intérêts égoïstes au bien-être social est “le plan secret de Dieu”, “l'ordre caché de Dieu”. “Il n'y a pas besoin que les individus soient vertueux pour que la société soit vertueuse”. Des individus égoïstes

“sont capables de faire, *sans le savoir ni le vouloir*, une (...) chose admirable: plus chacun ne vise que son intérêt, plus les uns se rendent *dépendants* des autres, et plus il en résulte une *réalité supérieure susceptible de transcender la mise de chacun*”.

C'est déjà de la théorie des jeux et de la théorie de l'émergence. Cette “chose admirable” (une ruche humaine) n'est possible qu'à cause de la complexité des interactions. J'insiste sur le fait que la réalité sociale ainsi produite *transcende* les individus.

La *Fable des abeilles* déclencha une polémique terrible (par exemple avec Berkeley) et fut condamnée en 1724 pour ses “diabolic attempts against religion” car elle introduisait un principe *d'inversion* entre les intentions individuelles micro-sociales et les propriétés macro-sociales émergentes. Les individus sont intentionnellement égoïstes et dominés par leurs intérêts particuliers, mais, à cause de leurs interactions, ils engendrent de façon *non* intentionnelle un ordre social global favorable à l'intérêt général. C'est l'inverse de l'adage “l'enfer est pavé de bonnes intentions”.

(iii) Adam Smith (1723-1790) a repris cela avec la “main invisible” (*Theory of Moral Sentiments*, 1759, *The Wealth of Nations*, 1776, IV.2).

“(L'individu) ne pense qu'à son propre gain ; en cela, comme en beaucoup d'autres cas, il est conduit par une main invisible à remplir une fin qui n'entre nullement dans ses intentions ; et ce n'est pas toujours ce qu'il y a de plus mal pour la société, que cette fin n'entre pour rien dans ses intentions. Tout en ne cherchant que son intérêt personnel, il travaille souvent d'une manière bien plus efficace pour l'intérêt de la société, que s'il avait réellement pour but d'y travailler.”

On voit ainsi apparaître explicitement un principe d'auto-organisation *qui n'est pas rationnellement calculé par les agents*. La cohésion sociale, la coopération, la prospérité commune — qui relèvent de l'intérêt général — sont des effets émergents non intentionnels issus de l'agrégation de libres intentions individuelles.

Sur cette première base se développeront de nombreuses variantes de l'IM, plus utilitaristes (John Stuart Mill 1806-1873, l'école néo-classique Léon Walras, 1834-1910 et Wilfrid Pareto, 1848-1923 : *Traité de sociologie générale*, 1916), ou plus organicistes (mais non holistes, Herbert Spencer 1820-1903 : *Introduction à la science sociale*, 1873).

On s'accorde à considérer que l'IM *complexe* a été fondé par l'école autrichienne de Carl Menger, Ludwig von Mises, Friedrich von Hayek. Une fois éliminée la magie du

holisme et une fois compris que l'atomisme d'agents séparés sans interactions est totalement insuffisant, on part des motivations, des intentions et des comportements individuels pour en étudier les interdépendances, les interactions compétitives-coopératives, les synergies et les intégrations.

C'est avec Carl Menger (1840-1921, *Grundsätze des Volkswirtschaftslehre*, 1871; *Untersuchungen über die Methode des Socialwissenschaften*, 1883) que s'introduit explicitement la problématique de la *complexité*. Des ordres auto-organisés comme le langage, la religion, le droit, la monnaie, le marché, l'état, etc. sont des "résultats non attendus et non recherchés". Elles ne sont pas "l'œuvre d'une volonté commune délibérément destinée à les établir." D'où un problème fondamental de conditions de possibilité. Cela est possible à cause de l'échange et de sa complexité.

"Les phénomènes sociaux (...) sont le résultat non intentionnel des efforts humains tendant à atteindre des buts essentiellement individuels."

On en arrive ainsi à Hayek.

- (i) L'ordre est une conséquence de l'auto-coordination des individus.
- (ii) Les structures émergentes s'autonomisent bien qu'elles soient causalement réductibles aux interactions individuelles.
- (iii) Elles n'existent stablement que si les agents respectent des règles de droit.
- (iv) Ces règles résultent d'une évolution culturelle.
- (v) Les structures émergentes sont *inintentionnelles* et imprévisibles, non planifiables.
- (vi) Les penser de façon intentionnelle produit des effets pervers dramatiques (totalitarisme).
- (vii) Il n'y a pas de téléologie historique mais une évolution non finalisée.

### **I.3. Epistémologie de l'émergence**

Epistémologiquement, l'IMC est inséparable d'une théorie de *l'émergence* et donc des débats sur le *réductionnisme* pour les systèmes complexes possédant au moins *deux niveaux* de réalité : un niveau *micro* sous-jacent comprenant un grand nombre d'unités élémentaires en interaction et un niveau *macro* où émergent des structures et processus macro d'auto-organisation.

Le réductionnisme postule que les phénomènes, structures, processus de haut niveau peuvent être réduits, en ce qui concerne leur explication causale, à des phénomènes, structures, processus de niveau sous-jacent. Un exemple paradigmatique est celui de la réduction de la thermodynamique macroscopique à des mouvements atomiques et moléculaires (température = énergie cinétique moyenne, etc.).

Le réductionnisme est alors inséparable de concepts duaux comme ceux d'"émergence", de "survenance" ("supervenience") ou de "fonctionnalisme". Le fonctionnalisme signifie ici que les macro-structures ayant un rôle *fonctionnel* ne peuvent

exister que si elles sont matériellement implémentées dans un substrat matériel sous-jacent, mais que, dans le même temps, en tant que fonctionnelles, elles sont largement *indépendantes* des propriétés “fine grained” de ce substrat.

Dans les systèmes complexes micro-macro (en général avec des niveaux méso intermédiaires), il existe des comportements *collectifs* macro régis par des lois qui ne sont pas celles du niveau micro sous-jacent. C'est le cas des phénomènes critiques en thermodynamique, de la percolation, des équations de réaction-diffusion à la Turing, des processus morphogénétiques à la Thom, des structures dissipatives à la Prigogine, de la turbulence, des automates cellulaires à la Wolfram, des réseaux de neurones, des colonies de fourmis, des marchés, etc. Suivant la conception que l'on se fait des lois, on peut donner différentes interprétations de ces phénomènes émergents (pensez par exemple à la possibilité de réduire ou non l'esprit aux neurones).

- (i) *Eliminativisme et épiphénoménisme*: les lois macro n'étant que des régularités empiriques sans contenu ontologique, les structures émergentes sont épiphénoménales et doivent être éliminées de la description scientifique.
- (ii) *Réalisme holistique* (c'est la position inverse): les lois macro ont un contenu ontologique et une efficacité causale. Les structures émergentes ne peuvent pas être réduites et induisent une causalité descendante (top-down) “immergente”.
- (iii) *Réductionnisme causal et émergentisme objectif*: les lois sont *objectives*, empiriquement fondées et mathématiquement formulables (cf. les lois de la thermodynamique), mais le niveau émergent macro n'a pas pour autant de contenu ontologique et causal *spécifique*. Il est *causalement* réductible à des interactions complexes du niveau micro sous-jacent, mais il possède néanmoins une autonomie théorique en tant que niveau de réalité descriptible (cf. la thermodynamique en physique statistique).

L'IMC relève de ce troisième point de vue.

## II. HAYEK ET LE PROBLEME DE LA COMPLEXITE

Hayek a beaucoup insisté sur les caractères spécifiques de cet ordre spontané qu'est l'ordre socio-économique dans les sociétés libérales modernes. Sa complexité endogène auto-organisationnelle est *irréductible*. Elle résulte du fait que les connaissances, les compétences et les informations sont *distribuées* entre une *myriade* d'agents cognitivement *limités* constituant des systèmes ouverts dont l'équilibre est dynamique.

Les conséquences de ce fait sont nombreuses :

- (i) D'abord la complexité est incompatible à la fois, d'un côté, avec une organisation hiérarchique centralisée et, d'un autre côté, avec une organisation associative contractuelle par concertation. Dans *Law, Legislation and Liberty*, Hayek explique que la morale des petites sociétés closes traditionnelles peut être fondée sur l'amour, la charité, la compassion, la justice, la loyauté, la solidarité, la réciprocité des dons, mais que les grandes sociétés ouvertes

ne le peuvent pas pour des raisons structurelles. Dans les grandes sociétés ouvertes, l'échange se fait dans le cadre de règles de droit abstraites et de systèmes de signes comme les prix. La liberté individuelle, la division du travail, la division du savoir et la productivité en sont constitutives.

Dans ces sociétés, la communication entre agents ne peut plus être assurée par des consensus mais uniquement par des signaux (comme les prix dans un marché) permettant d'échanger. D'où le rôle essentiel (mais souvent incompris) du marché. Le marché est avant tout une forme de circulation de l'information et de *coordination* des actions dans une société multi-agents auto-organisationnelle opaque.

Précisons ce point en nous référant à l'ouvrage classique de Philippe Nemo : *La société de droit selon Hayek*. Dans ce que Hayek a appelé la *catallaxie* (du grec échanger et coopérer), l'auto-organisation du marché (la "main invisible" d'Adam Smith) est la seule façon de créer un lien social viable fondé sur la pluralité des différences individuelles. Elle remplace une impossible communauté de fins (où la différence des objectifs ne peut qu'engendrer la guerre hobbesienne de tous contre tous) par une communauté de moyens. Dans un marché, chacun coopère avec chacun, mais indépendamment de visées communes. Le marché garantit la coopération malgré la divergence des intérêts et la concurrence des fins.

(ii) Ensuite le constructivisme rationaliste en matière de politique, de droit et de morale est inadéquat pour des raisons de principe. En effet, la complexité est *évolutionnaire*. Elle résulte d'un processus de sélection — une sélection historico-culturelle de règles de comportement, de pratiques, d'institutions — effectué par une intelligence collective qu'il est impossible de maîtriser conceptuellement.

(iii) Une troisième conséquence de la complexité est que les règles qui régissent les échanges et la communication sociale sont abstraites et formelles. Les systèmes sociaux complexes auto-organisés sont régis par les règles d'un droit civil abstrait (*nomos*) et non pas par celles d'un droit public positif ayant pour source la volonté constructiviste d'une souveraineté, qu'il s'agisse de celle de l'absolutisme monarchique ou de celle du peuple.

Selon Hayek, le libéralisme démocratique fondé sur le droit, les sciences, les techniques et l'économie de marché constitue un mode supérieur d'action, de socialisation et de communication produit par l'évolution historico-culturelle.

Sur le plan de la psychologie cognitive, Hayek a par ailleurs beaucoup insisté sur le rôle fondamental des connaissances de situations typiques et génériques au moyen de schèmes abstraits simplifiant la complexité et permettant des anticipations plausibles. Ces schèmes sont sélectionnés de façon évolutionniste et dominent l'ordre communicationnel du social sous la forme de conventions, de règles et de normes.

### III. LA CRITIQUE DU RATIONALISME CONSTRUCTIVISTE

#### III.1. La critique du rationalisme constructiviste

La critique hayekienne du constructivisme dénonce le rationalisme “inconditionné” en matière de politique. La cognition des agents étant limitée, il est cognitivement impossible d'effectuer un “calcul” des causes et des effets des actions. L'information nécessaire sur la société serait tellement énorme et complexe qu'elle exigerait des agents *omniscients*. A cause de la complexité même des interactions et du caractère *distribué* des connaissances associées, on ne peut pas aller au-delà d'une simple *coordination* cohérente des agents. Les structures organisationnelles ne sont pas récapitulables dans une intelligence individuelle. Leur synthèse rationnelle est impossible et, lorsqu'elle est imposée par un pouvoir politique, elle ne peut déboucher que sur des effets pervers. C'est à cause de vérités systémiques profondes que la planification produit toujours l'inverse de ce qu'elle vise.

Une conséquence de la nature complexe de l'ordre catallactique est qu'il y existe une *opacité structurelle*. La convergence des jugements et des comportements n'est possible que dans des *micro-communautés* et les grandes sociétés ouvertes modernes ne sont pas de telles communautés. Et c'est précisément parce que “l'ordre étendu” est un *macro-ordre* de coopération *globale* où les règles traditionnelles de solidarité et d'altruisme ne peuvent opérer que *localement* pour des micro sous-communautés, qu'il n'est pas “naturel” pour les individus. Nous rencontrons là ce que j'appelle le *paradoxe de la réflexivité sociale*. C'est comme pour la ruche de Mandeville : pour des raisons de complexité, une réflexivité cognitive de leur structure sociale est inatteignable pour les agents. La société est localement homogène et planifiable mais globalement hétérogène et immaîtrisable (sauf à faire dégénérer sa complexité). C'est la confusion du local et du global, du micro et du macro, de l'homogène et de l'hétérogène, qui est la source du “fatal conceit” dénoncé par Hayek.

#### III.2. L'origine évolutionnaire des règles de conduite

Mais il faut aller un peu plus loin et s'interroger sur les règles de conduite et les maximes d'action. Nous avons vu (avec Nicole et Mandeville) que des individus vertueux ne font pas forcément des sociétés vertueuses et que des “fripons” peuvent concourir malgré eux au bien public par une sorte de “ruse de la raison”. Mais évidemment, il faut aussi des règles de conduite qui échappent à ce double bind mandevillien et soient à la fois “justes” pour les sujets et “utiles” pour les sociétés. Il n'est pas du tout évident que de telles règles existent. C'est ici qu'intervient chez Hayek une défense remarquable du *sens commun*.

Il existe selon Hayek une origine *évolutionniste* des “bonnes” règles de conduite, conventions et normes. Elles sont le résultat d'une *sélection culturelle* — donc d'un apprentissage collectif-historique — fonctionnant comme un processus concurrentiel ayant avantagé les groupes les ayant adoptés. Elles permettent d'agir sans devoir à chaque fois récapituler toutes les expériences permettant d'agir. Le sens commun est un ensemble de

connaissances *tacites* (et non pas explicites) et de schèmes *pratiques* (et non pas théoriques) permettant d'agir et de prévoir dans notre environnement, sans nous laisser submerger par le haut flux d'informations non pertinentes charriées par sa complexité. Les maximes du sens commun ont une valeur cognitive opératoire. Pour Hayek, elles sont des “savoirs incorporés” d'origine en quelque sorte “phylogénétique” (au sens de l'évolution culturelle) et il est par conséquent rationnel de s'y conformer “ontogénétiquement”.

De même qu'en biologie évolutionniste les a posteriori phylogénétiques fonctionnent comme des a priori ontogénétiques, on peut dire que ces règles fonctionnent pour les sujets comme des a priori. En ce sens il y aurait chez Hayek un apriorisme évolutionniste expliquant l'auto-transcendance des règles. Comme les grammaires des langues, elles procèdent d'institutions symboliques ne relevant ni d'une intelligence individuelle omnisciente ni d'un contrat social délibératif mais bien plutôt d'une sélection naturelle.

On voit ainsi s'articuler de façon précise chez Hayek la psychologie cognitive et la sociologie des macro-ordres spontanés complexes. Au niveau cognitif, les actions humaines dans un environnement complexe opaque sont guidées par des propensions, des dispositions et des tendances. Les patterns d'action et les schèmes de perception permettent d'anticiper. Ils permettent de reconnaître des traits abstraits *génériques* des situations et de mobiliser les comportements adaptés correspondants.

### **III.3. La nature de l'évolution culturelle**

Un problème difficile est de savoir quelle est la nature de l'évolution culturelle en tant que sélection naturelle de règles de conduite. L'évolution doit sélectionner des individus qui adoptent des règles maximisant les performances collectives du groupe. Il faut d'ailleurs noter, comme le font Robert Nadeau et Paul Dumouchel, que les maximes d'action ici en jeu sont procédurales. Ni mœurs primitives, ni règles rationnelles réflexivement adoptées, elles sont en général abstraites et négatives. Elles ne concernent pas les fins de l'homme, mais des conditions de possibilité du bien général.

Par quels mécanismes de telles règles abstraites ont-elles pu être sélectionnées ? Une sélection de groupe est difficile à admettre car toute sélection de type darwinien doit porter sur des individus. Dans ce cas, il faudrait penser qu'elles constituent une sorte de “génotype” qui, exprimé “phénotypiquement”, se convertirait en stratégies individuelles (maximes d'action). Une règle individuelle (micro) n'aurait alors rien à voir avec l'intériorisation de régularités sociales (macro). Le problème est extrêmement délicat.

Quoi qu'il en soit, *pour les individus*, ces normes *socialement utiles* opèrent comme des *valeurs*, car ceux-ci ne comprennent pas leur véritable opérativité sociale (impossibilité de réflexivité sociale). L'évolution culturelle est trop longue. Elle encode une expérience collective historique non récapitulable par les individus. L'individu ne pouvant pas comprendre l'efficacité *utilitariste* des normes, il est donc bon qu'il les suive par devoir. On retrouve ainsi un point de vue “pratique” (au sens éthique). Mais les normes étant socialement

utiles on retrouve aussi les théories utilitaristes de l'éthique (Jeremy Bentham et John Stuart Mill). Simplement le “calcul” de l'utilité des règles et actes moraux est *cognitivement impossible* pour les individus.

Bref, selon Hayek, l'évolution culturelle implique que certaines maximes d'action puissent être des impératifs “catégoriques” transcendants pour les individus tout en étant des règles utiles pour les sociétés. Cela compense en partie, par le recours à une réflexivité éthique, l'absence de réflexivité sociale.

Un impératif est un jugement normatif formulant un devoir. Un impératif “catégorique” est purement procédural et relève d'une conception *non conséquentialiste* des actions. C'est un “know how” au sens de Ryle (savoir comment agir) et non pas un “know that” (savoir pourquoi c'est utile). Un “know that” conséquentialiste supposerait d'une façon ou d'une autre de pouvoir effectuer un “calcul des fins”. Or un tel calcul est impossible pour un sujet cognitif de ressources limitées. Mais il est effectuable par *l'intelligence collective distribuée et l'évolution culturelle*. C'est ce qui se passe dans une conception évolutionniste comme celle de Hayek. L'évolution culturelle est “utilitariste” mais sélectionne des maximes qui pour les sujets sont des conventions ou des normes.

#### IV. L'EXEMPLE DES JEUX EVOLUTIONNISTES

J'aimerais souligner qu'il existe des modèles rigoureux tout à fait non triviaux de la thèse hayekienne dans *la théorie des jeux évolutionnistes*. Depuis les travaux fondateurs de Robert Axelrod, W. Poundstone, M. Nowak et K. Sigmund, on a beaucoup travaillé sur certains systèmes complexes adaptatifs “sociaux” pour lesquels on sait analyser les mécanismes causaux sous-jacents.

##### IV. 1. Le dilemme du prisonnier

Il y a deux joueurs  $A$  et  $B$ . Pour chaque comportement possible des joueurs, à savoir  $d$  = défection (trahir) et  $c$  = coopération, la matrice du jeu donne les gains (payoffs) des joueurs. La matrice fait intervenir 4 gains :  $T = (d, c)$  = Tentation,  $D = (c, d)$  = Duperie,  $R = (c, c)$  = Récompense,  $P = (d, d)$  = Punition. Pour que le jeu soit intéressant, il faut que les gains satisfassent les conditions :  $T > R > P > D$  et  $(T + D)/2 < R$ .

Avec une telle matrice de gain, la double stratégie  $(d, d)$  (perdant-perdant) est le seul *équilibre de Nash* du jeu, i.e. la double stratégie telle que chaque joueur fasse moins bien s'il change de stratégie de façon unilatérale. Inutile de dire que l'échec perdant-perdant sévit en toute occasion toujours et partout.

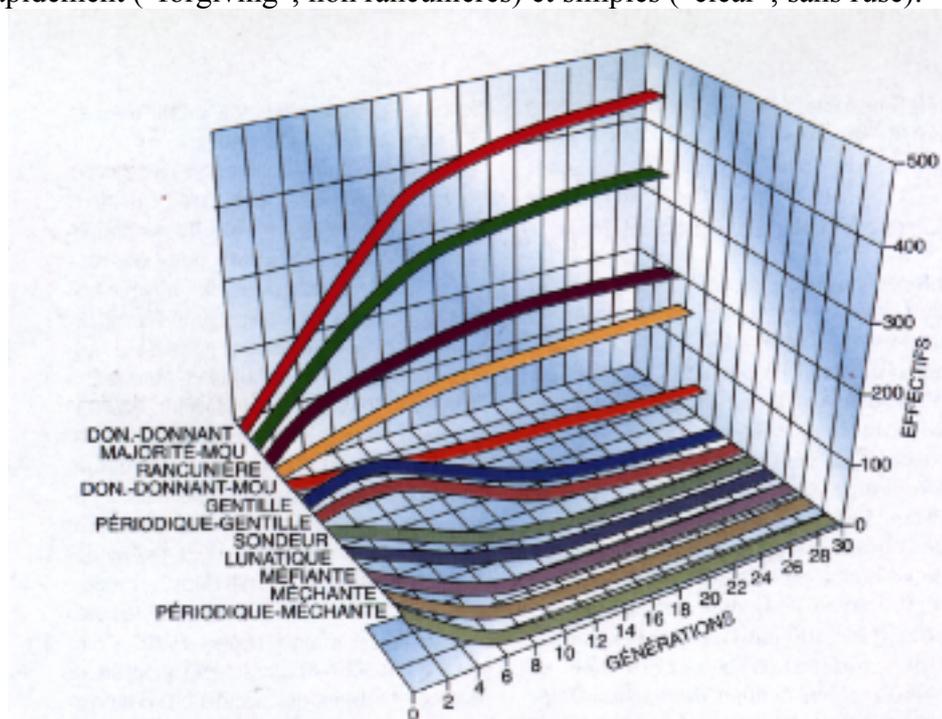
On peut généraliser cet exemple de multiples façons, par exemple en introduisant des asymétries, des inégalités larges, un comportement neutre (refus de jouer), des joueurs multiples, des probabilités, etc. Mais le phénomène que nous venons de décrire se révèle *robuste* pour les parties à un seul coup. Comment expliquer alors à partir de ce constat initial

la façon dont la coopération peut être sélectionnée évolutivement. C'est un problème théorique fondamental.

#### IV. 2. Le dilemme du prisonnier itéré (IPD)

La situation du dilemme change complètement lorsque l'on *itère* le jeu car la défection peut alors être sanctionnée et la coopération récompensée. On peut dans ce cas introduire de véritables *stratégies*. On suppose que le nombre de coups est indéterminé et l'on teste des stratégies du genre : *G* = “gentille” = toujours *c*; *M* = “méchante” = toujours *d*; *TFT* = “donnant-donnant” (tit for tat) = d'abord *c* puis jouer ce que l'autre a joué à la partie précédente; *R* = “rancunière” = *c* mais toujours *d* dès que l'autre a trahi une fois, etc. On confronte ces stratégies sur un grand nombre de parties et l'on étudie leurs scores.

On constate alors que pour un pool de stratégies simples il y a une supériorité très nette de la stratégie *TFT*. De façon générale, les simulations montrent qu'il y a une supériorité manifeste des stratégies coopératives (“nice”), vite réactives aux trahisons (“retaliatory”), pardonnant rapidement (“forgiving”, non rancunières) et simples (“clear”, sans ruse).



*Compétition entre stratégies (d'après J-P. Delahaye). Population de 1200 agents. 12 stratégies représentées par 100 agents chacune. A la stabilisation, il n'y a plus que des stratégies coopératives et TFT domine. (Majorité-mou = jouer ce qu'a joué la majorité et si 50%-50% jouer c; méfiante = TFT avec d initial).*

#### IV. 3. Les jeux évolutionnistes

La théorie des jeux évolutionnistes consiste à considérer des *populations* d'individus utilisant différentes stratégies et à définir les nouvelles générations à partir des scores obtenus

dans une confrontation généralisée. Les stratégies avec bons scores augmentent leurs représentants alors que celles avec mauvais scores disparaissent progressivement.

#### **IV. 4. La stratégie du “tit for tat” : du sens commun aux modèles**

Dans ces modèles, les agents sont considérés comme des “phénotypes” exprimant des stratégies “génotypes” et leurs stratégies “micro” influent sur la dynamique “macro” de leur population. Les simulations de cette dynamique fournissent des résultats extrêmement intéressants. Pour des stratégies simples comme ci-dessus Axelrod a montré :

- (i) Il y a élimination des stratégies anti-coopératives et la coopération s’installe et se stabilise.
- (ii) C’est la stratégie *TFT* qui domine, mais dans les cas où il peut exister des mutants elle est *fragile* car les mutants “gentils” ont le même comportement que les agents *TFT* dans un environnement *TFT* et peuvent donc se substituer progressivement aux *TFT*; mais alors des mutants “méchants” peuvent facilement les déstabiliser et envahir le système.
- (iii) Pour une stratégie, la réactivité aux trahisons est une condition pour être *collectivement stable*, c'est-à-dire ne pas pouvoir être déstabilisée par un mutant.
- (iv) Si l'on introduit des stratégies complexes, alors il peut se produire énormément de phénomènes subtils différents dès qu'il existe au moins trois stratégies en interaction. Par exemple une stratégie non coopérative peut en utiliser une autre pour éliminer les stratégies coopératives et l’éliminer ensuite à son tour. Ou encore le désordre permet à des stratégies non coopératives de survivre et même de gagner, etc.
- (v) La maîtrise computationnelle de ces situations permet alors de montrer qu'il existe d'autres stratégies que *TFT* qui sont maximales performantes dans ces contextes élargis. Autrement dit, on peut raffiner le *TFT* sélectionné par le sens commun.

Bref on constate “expérimentalement” (au sens d'une synthèse computationnelle) une extrême complexité des dynamiques possibles.

Nous rencontrons ici un exemple typique de *modèle du sens commun* :

- (i) Les simulations confirment un sens commun politique, social, pédagogique qui constitue un savoir incorporé et permet d’agir de façon efficace sans avoir à répéter indéfiniment les mêmes expériences déceptives.
- (ii) Mais dans le même temps, elles permettent de dépasser le sens commun, non pas en lui donnant tort, mais en sélectionnant des règles en quelque sorte d’“hyper” sens commun dans le cadre *expérimental* (au sens d'une synthèse computationnelle) d'évolutions culturelles *virtuelles*.

Les jeux évolutionnistes sont intéressants dans la mesure où ils remplacent une intelligence rationnelle individuelle surhumaine (irréalisable) par l’intelligence collective (réalisable) d’une population d’agents en interaction dont la rationalité et les ressources cognitives sont limitées. L’optimisation n’est plus individuelle mais collective et peut être

obtenue *sans hypothèse de rationalité individuelle forte*. Elle relève d'un sens commun "artificiel" mimant le sens commun "naturel".

Ce qui m'intéresse le plus dans ces résultats est un paradoxe apparent : on peut donner raison à l'anti-constructivisme de Hayek au moyen de formalisations mathématiques et de simulations informatiques alors que Hayek fait partie des penseurs qui mettent en avant l'impossibilité pour les individus d'effectuer de tels calculs. Mais il n'y a en fait aucune contradiction. Même si les formalisations mathématiques sont cognitivement accessibles, les simulations informatiques qui calculent leurs conséquences ne sont pas, quant à elles, cognitivement effectuelles. Il s'agit d'un phénomène fondamental mais tout à fait banal en sciences. La loi de Newton est une équation différentielle très simple, mais le calcul de ses solutions est vertigineusement compliqué et encore en grande partie inaccessible malgré les ordinateurs. Cela est lié au fait qu'intégrer une équation différentielle consiste à *itérer* son générateur infinitésimal et que l'itération de règles simples conduit en général à une très grande complexité. C'est ce qui se passe dans l'exemple ci-dessus.

Cette justification *computationnelle* de l'anti-constructivisme hayekien rend le constructivisme politique qu'il dénonçait encore plus obscurantiste.

## V. EUSOCIALITE ET SOCIALITE

Pour conclure, je vais revenir à la "fable des abeilles" pour souligner le fait que la référence à des *insectes sociaux* est particulièrement pertinente. En fait les problèmes que nous avons esquissés plongent leurs racines dans toute l'épaisseur de l'évolution naturelle et culturelle.

1. Quel que soit le niveau évolutif des individus d'une espèce donnée, il y a plusieurs façons de "faire société". La première, qui est dominante chez les mammifères supérieurs, est celle de ce qu'on peut appeler les "communautés" comme groupes familiaux élargis. Les niveaux cognitifs sont très différents suivant les cas, mais la caractéristique de ces communautés avec leurs règles éthologiques de comportement (hiérarchie, altruisme, etc.) est que les liens communautaires sont *accessibles* à l'expérience des sujets. Disons que l'organisation communautaire est "cognitivement commensurable". Pour l'homme, cela correspond au fait que l'expérience communautaire peut être *réflexive* (réflexivité sociale).

A l'autre extrême, l'évolution a développé des modes de socialité qui *transcendent* les individus. C'est le cas des insectes sociaux comme les abeilles, les fourmis ou les termites. Les traits caractéristiques de ces étonnantes organisations sociales sont

(i) que les individus y suivent certaines règles éthologiques, précises, *élémentaires* et différenciées au sens de la *division du travail* ;

(ii) que le nombre d'individus est extrêmement *grand* et qu'il est donc pertinent de distinguer *deux niveaux* radicalement différents, celui "micro" des individus et celui "macro" du social, le niveau macro étant organisé par des structures émergentes résultant d'un nombre considérable d'interactions entre individus micro ;

(iii) que ce nombre énorme d'agents micro se comportant et interagissant en suivant des règles élémentaires est capable de construire par un processus auto-organisationnel d'“ordre spontané” des architectures sociales totalement *incommensurables* aux ressources cognitives des agents. Il faut insister sur ce miracle, cette “chose admirable” pour parler comme Nicole, de l'évolution biologique.

Les “architectures sans architectes” que sont par exemple les termitières sont d'immenses *macro*-structures spatiales collectives (à notre échelle, une dizaine de kms) hautement sophistiquées dont les modèles mathématiques sont tout à fait non triviaux (cf. par exemple les travaux de Jean-Louis Denebourg, Guy Theraulaz, Eric Bonabeau, Bernard Manderick). Leurs architectures sont fonctionnelles pour les populations d'agents micro qu'elles hébergent et, comme de vrais édifices, comprennent des piliers, des murs externes, des galeries, des membrures et des spirales pour la ventilation et le refroidissement, des chambres (chambre royale, nursery), une base terrassée avec des conduits et des vannes, etc. De telles architectures sont totalement irréprésentables au niveau des agents qui pourtant les construisent (non réflexivité sociale).

Pour parler de cet autre mode de socialisation radicalement opposé au mode communautaire, on utilise celui d'organisation “*eusociale*” en mettant en avant son intelligence collective distribuée (“swarm intelligence”). Les organisations eusociales sont déjà des “grandes sociétés” purement naturelles.

2. Le point crucial est qu'il existe ainsi au moins *deux* complexifications évolutives nettement différentes de la socialité. D'une part une *complexification “verticale” de nature cognitive* comportant un enrichissement des facultés cognitives des agents individuels, en particulier représentationnelles, et d'autre part une *complexification “horizontale” eusociale* comportant une division du travail, une “swarm intelligence” et un changement de niveau micro-macro. Soit l'intelligence individuelle se développe mais le groupe reste communautaire, ce qui se passe dans l'évolution conduisant vers les primates et l'*homo sapiens*, soit l'intelligence individuelle reste limitée mais les groupes deviennent eusociaux, ce qui permet à une intelligence collective distribuée d'émerger.

Nous sommes des primates évolués et une partie des caractéristiques humaines se situe sur l'axe “vertical” de l'évolution cognitive relayée par l'évolution culturelle. Mais il n'y a pas d'organisation *eusociale* chez les primates. Or, une caractéristique essentielle de l'évolution historique de l'homme est que, à partir de l'émergence de ce que l'on a appelé les grandes civilisations *urbaines* (qu'il faut ici opposer aux cultures traditionnelles), après la sédentarisation des chasseurs-cueilleurs, l'apparition de l'agriculture, de l'élevage, des techniques puis, surtout, des villes, de l'écriture, de l'arithmétique comptable et de la géométrie des habitats, de la monnaie, des voies de circulation terrestres et maritimes, des échanges commerciaux, etc., une rupture fondamentale s'est produite et une *complexité “horizontale” eusociale* s'est développée de façon vertigineuse, complexité complètement

étrangère à la lignée évolutive de l'*homo sapiens* et qui n'existait jusque-là que chez les insectes sociaux.

L'intelligence humaine est donc non seulement maximale, mais elle l'est en quelque sorte “au carré”, à la fois dans sa dimension cognitive individuelle et dans sa dimension eusociale collective. Et les “grammaires” de juste conduite font précisément le lien entre ces deux dimensions.

C'est pourquoi Mandeville est si pertinent dans sa “fable des abeilles”. En utilisant le genre littéraire de la fable, il met en place une analogie que je reformulerai de la façon suivante : la civilisation moderne est à l'*homo sapiens* ce qu'une ruche est à une abeille. D'une part, les individus (intelligents, rationnels, etc.) y suivent des règles de comportement et des maximes d'action qui, si les règles sont bonnes et bien suivies, transforment les innombrables interactions en ordre *eusocial* spontané. D'autre part, la complexité interne de ce dernier est *incommensurable* avec l'intelligence et la rationalité des agents. C'est à cause de cette incommensurabilité que les effets des interactions entre intentionnalités individuelles sont des effets “*in-intentionnels*”.

C'est je pense cette prise en compte du caractère eusocial des grandes sociétés ouvertes modernes fondées sur les technosciences et l'économie de marché qui justifie la critique du “constructivisme” chez Hayek : dans une collectivité eusociale, aucune intelligence individuelle n'est à même de rendre la complexité interne du social compréhensible et transparente, perceptible et descriptible. Il s'agit là d'une sorte de principe *d'incomplétude* : pour une espèce d'individus eusociaux *A* d'intelligence et de rationalité données, la complexité *eusociale* des macro-sociétés formées d'agents de type *A* transcende les ressources cognitives de *A*.

Si l'on admet l'hypothèse de la *double* complexité alors se pose tout naturellement la question de leur compatibilité et rien ne permet de supposer que la compétence cognitive soit *compatible* avec la complexité eusociale. La thèse de l'incommensurabilité dit même qu'il y a une certaine incompatibilité et c'est précisément cela le “paradoxe de Mandeville” qui affirme une sorte *d'inversion* de valeur entre l'individuel et le social. Le problème kantien de “l'insociable sociabilité” (1784, *Idée d'une histoire universelle d'un point de vue cosmopolitique*) vient sans doute du fait que la civilisation a transformé l'*homo sapiens* en primate eusocial.

## BIBLIOGRAPHIE

- Axelrod R., Cohen M., Rislo, R. 1998. “The Emergence of Social Organization in the Prisoner's Dilemma: How Context Preservation and other Factors Promote Cooperation”, à paraître.
- Binmore K., 1994. *Playing Fair*, Cambridge, MIT Press.

- Delahaye J-P., Mathieu P., 1999. “Des surprises dans le monde de la coopération”, *Les Mathématiques sociales, Pour la Science*, 58-66.
- Dumouchel, P., 1999. *Emotions: Essai sur le corps et le social*, Institut Synthélabo, Paris.
- Dupuy, C., Torre, A., 1999. “The morphogenesis of spatialized cooperation relations”, *European Journal of Economic and Social Systems*, 13, 1, 59-70.
- Dupuy, J-P., 2006. « Friedrich Hayek ou la morale de l'économie », *Histoire du libéralisme en Europe*, (Ph. Nemo et J. Petitot eds), Presses Universitaires de France, Paris, 1151-1213.
- Hayek, F. A. 1944. *The Road to Serfdom*, Routledge & Kegan Paul, London.
- Hayek, F. A. 1952. *The Sensory Order: An inquiry into the Foundations of Theoretical Psychology*, U. of Chicago Press.
- Hayek, F. A. 1982. *Law, Legislation and Liberty*, Routledge & Kegan Paul, London.
- Hayek, F. A. 1988. *The Fatal Conceit: The Errors of Socialism. The Collected Works of F. A. Hayek*, Vol. 1. (W. W. Bartley, ed.), Routledge, London.
- Hofbauer J., Sigmund K., 1988. *The Theory of Evolution and Dynamical Systems*, Cambridge UP.
- Kirman A., 1998. “La pensée évolutionniste dans la théorie économique néoclassique”, *Philosophiques*, XXV, 2, 219-237.
- Laurent, A., 1994. *L'individualisme méthodologique*, Que sais-je ?, n° 2906, Paris, Presses Universitaires de France.
- Nadeau R., 1998. “L'évolutionnisme économique de Friedrich Hayek”, *Philosophiques*, XXV, 2, 257-279.
- Nadeau, R., 2006. « Friedrich Hayek et le génie du libéralisme », *Histoire du libéralisme en Europe*, (Ph. Nemo et J. Petitot eds), Presses Universitaires de France, Paris, 1115-1150.
- Nemo, P., 1988. *La société de droit selon F.A. Hayek*, Paris, Presses Universitaires de France.
- Petitot, J., 2000. “Vers des Lumières hayekiennes : de la critique du rationalisme constructiviste à un nouveau rationalisme critique”, Colloque de Cerisy *Friedrich Hayek et la philosophie économique* (A. Leroux et R. Nadeau eds), *Philosophie économique*, 2, 9-46.
- Petitot, J., 2006. « Libéralisme et Liberté : Hayek avec Kant ou une éthique de la finitude », *Histoire du libéralisme en Europe*, (Ph. Nemo et J. Petitot eds), Presses Universitaires de France, Paris, 1215-1223.
- Petitot, J., 2006. « Modèles formels de la ‘main invisible’ : de Hayek à la théorie des jeux évolutionniste », *Histoire du libéralisme en Europe*, (Ph. Nemo et J. Petitot eds), Presses Universitaires de France, Paris, 1095-1114.
- Poundstone W., 1993. *Prisoners Dilemma*, Oxford UP.

Samuelson L., 1997. *Evolutionary Games and Equilibrium Selection*, Cambridge, MIT Press.

Weibull J., 1996. *Evolutionary Game Theory*, Cambridge, MIT Press.