

Social Science Information

<http://ssi.sagepub.com>

Simon, Stigler et les théories de la rationalité limitée

Philippe Mongin

Social Science Information 1986; 25; 555

DOI: 10.1177/053901886025003002

The online version of this article can be found at:

<http://ssi.sagepub.com>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:

[Maison des Sciences de l'Homme](#)

Additional services and information for *Social Science Information* can be found at:

Email Alerts: <http://ssi.sagepub.com/cgi/alerts>

Subscriptions: <http://ssi.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Philippe Mongin

Simon, Stigler et les théories de la rationalité limitée

1. Rationalité limitée et optimisation

La notion de rationalité limitée entretient manifestement des rapports complexes avec le programme de recherche “optimisateur” de l’économie traditionnelle. Il est significatif, à cet égard, que l’inventeur et le défenseur de la notion, H. Simon, ait proposé successivement plusieurs analyses de ces rapports. Lorsqu’on les confronte, on voit apparaître une question embarrassante pour la théorie de la rationalité limitée: Simon a-t-il seulement voulu indiquer que l’on devait élargir l’ensemble des contraintes traditionnellement posées sur la décision individuelle, en tenant compte à la fois des contraintes *externes* (technologie, revenu, etc.) et des contraintes *internes* (coûts d’information et de calcul), ou bien la notion de rationalité limitée implique-t-elle, à ses yeux, cette conséquence plus forte – la récusation de toute idée d’optimisation?

Cette question recouvre un enjeu bien déterminé : l’évaluation des théories néo-classiques, “optimisatrices”, de la rationalité limitée, qui se sont développées, depuis 1961, à la suite d’un article important de G. Stigler, “The Economics of Information”. Quand Simon expose pour la première fois ses idées, en 1955, la théorie micro-économique ordinaire semble incapable de relever ses objections et de lui opposer son propre concept de rationalité limitée. En un certain sens, il n’en va plus de même dix ans après, lorsque la théorie de la “recherche” (*search*), inaugurée par Stigler, offre les premiers modèles optimisateurs de la décision en

Je remercie P. A. Chiappori des remarques qu’il a bien voulu faire sur la première version de ce texte.

Information sur les Sciences Sociales (SAGE, Londres, Beverly Hills, Newbury Park et New Delhi), 25, 3 (1986), pp. 555–606

information limitée et coûteuse. Ces modèles prennent au pied de la lettre les recommandations faites naguère par Simon, puisqu'ils *élargissent* le système des contraintes qui sont de nature à influencer la décision individuelle. L'agent représentatif, désormais, tient compte de l'indisponibilité de certaines informations et du coût que représenterait leur obtention. Mais, naturellement, ces modèles sont optimisateurs: l'agent intègre les contraintes nouvelles qui affectent sa décision de la même façon, exactement, qu'il intègre les contraintes externes dans la théorie traditionnelle – *il maximise son utilité relativement à ces contraintes aussi*. Pour beaucoup d'économistes contemporains, cette parade élémentaire constitue la réponse définitive qu'appellent les objections de Simon. On voit donc l'enjeu précis de la question précédente: Simon n'a-t-il écrit ses *Models of Bounded Rationality* que pour stimuler la recherche néo-classique contemporaine?

Simon a pris position contre cette récupération mécanique de ses idées par la théorie traditionnelle ; mais il nous semble aussi vrai qu'il a paru quelquefois l'approuver lointainement. C'est là un côté exégétique, à vrai dire secondaire, de la question soulevée par la rationalité limitée. Peu importe, finalement, que Simon ait pu hésiter dans le jugement qu'il porte sur l'interprétation optimisatrice de la rationalité limitée: l'essentiel est qu'il y ait effectivement matière à débat. Il faut examiner l'argumentation, nullement insignifiante, que proposent les économistes contemporains: l'optimisation étendue aux contraintes *internes* de l'agent représenterait la seule manière de formaliser l'idée, essentiellement fuyante, des limites de la rationalité ; elle serait peut-être même la méthode la plus appropriée pour donner un contenu empirique à des hypothèses qui, dans la présentation relativement informelle de Simon, n'en possèdent pas encore.

L'examen auquel nous procéderons va s'appuyer sur l'énoncé *a priori*, et formalisé jusqu'à un certain point, de la notion de rationalité limitée. Nous montrerons que la théorie optimisatrice peut très facilement se mouler dans une telle définition générique — bien mieux, qu'elle réussit, avec une facilité surprenante, à lui prêter vie. L'apport essentiel de l'optimisation, dans ce contexte, tient peut-être dans un théorème remarquable des modèles de *search* : lorsque l'on dispose d'une information limitée, il est optimal — sous des conditions très générales – de prendre sa décision en faisant intervenir un *seuil de satisfaction* révisable au fur et à mesure que les données sont connues. Or la présentation ordinaire de la rationalité limitée – sinon, peut-être, sa définition générique qui, selon nous, doit éviter de prendre parti à ce stade – fait intervenir, comme une pièce essentielle, le rôle des seuils de satisfaction

endogènes. Et c'est, à notre avis, l'une des faiblesses les moins discutables des modèles proposés par Simon qu'ils n'explicitent pas réellement le mécanisme d'endogénéisation des seuils : la théorie de la "recherche" parvient, en revanche, à donner la règle d'évolution des seuils au cours du déroulement de la recherche.

Mais s'il est vrai que la modélisation traditionnelle s'accommode facilement de la notion abstraite de rationalité limitée, il n'est pas sûr qu'elle en donne une interprétation méthodologiquement et logiquement satisfaisante. Nous montrerons que la puissance de la théorie dans la formalisation se paie d'un double inconvénient. En premier lieu, la dérivation des seuils optimaux apparaît le plus souvent triviale, puisqu'elle résulte de l'injection, dans le modèle traditionnel, de variables libres supplémentaires : d'une part, des coûts de calcul dont la forme exacte n'est pas indiquée, d'autre part, une distribution de probabilité subjective qui est elle aussi librement ajustable. Cette objection, de nature fondamentalement *méthodologique*, gagne à être reformulée dans le langage poppérien des "degrés de testabilité" ou dans celui, lakatosien, des "programmes de recherche dégénérescents". En second lieu, les modèles du *search* ressortissent à une théorie de l'optimisation au second ordre : l'agent qu'ils décrivent tient compte, *d'une façon qui est elle-même optimale*, des difficultés qu'il y a à optimiser. On soupçonne aussitôt, à la lecture d'un tel énoncé, que ces modèles appellent une objection de régression infinie. Il est surprenant que les polémiques sur l'optimisation, qui ont donné lieu à une surabondance d'articles, n'aient pour ainsi dire jamais évoqué cette objection importante.

Nous nous efforcerons de resituer les positions (explicites ou non) de Simon sur notre problème. Il s'agira de rappeler comment il a pu définir la rationalité limitée, et quelles sont finalement les raisons de son désaccord avec la méthode du *search*. Le propos de cet article n'est pas essentiellement exégétique, mais nous prendrons sans cesse appui sur les indications parfois ambiguës, mais toujours précieuses, que donnent les textes de Simon.

C'est d'ailleurs l'une de ses distinctions récentes, celle de la "rationalité procédurale" et de la "rationalité substantielle" (1976, 1978), qui nous semble fournir le meilleur éclairage sur notre problème. Nous montrerons que l'idée de *procédure*, ou encore de *délibération*, concentre finalement l'enjeu du débat entre les économistes néo-classiques et ceux qui se nomment, peut-être improprement, théoriciens de la rationalité limitée. Cette dernière notion est purement taxinomique ; il serait vain de nier qu'elle soit compatible avec l'approche optimisatrice. En revanche, et c'est là, sans doute, la leçon la plus instructive d'une polémique contre

le *search*, l'approche optimisatrice ne parvient pas à restituer l'idée fondamentale d'une rationalité délibératrice. Ce trait permet d'unifier à la fois certaines positions disparates de Simon sur l'optimisation et les critiques d'ordres divers, logique ou méthodologique, qu'appelle, selon nous, la théorie optimisatrice de la rationalité limitée. Toutes ces indications pointent dans la même direction, celle d'un postulat jamais explicite, mais inhérent aux modèles du *search* comme à ceux de la micro-économie naïve : *le caractère non délibératif de la rationalité*.

2. Les limites de la rationalité limitée

2.1. Les positions variables de Simon sur l'optimisation

Simon ne paraît pas avoir toujours analysé de la même manière les rapports qu'entretiennent la rationalité limitée et le programme de la rationalité optimisatrice ou "globale", suivant l'expression qu'il emploie volontiers. Sans doute ne revient-il jamais sur l'objection centrale que, dès 1955, il avance contre ce programme : toutes les données psychologiques dont nous disposons révèlent que l'homme est limité dans son aptitude à recueillir et à traiter l'information ; or ses limites restent en-deçà de la capacité que lui prête le modèle optimisateur le plus simple (1955, pp. 243, 246). L'importance de cette objection tient à ce qu'elle atteint l'essentiel de l'approche économique de la décision, et non pas telle ou telle de ses variantes spécifiques. En un sens, les versions sophistiquées du point de vue "classique" — comme Simon l'appelle aussi — ne font qu'aggraver ses défauts inhérents : pour maximiser une *utilité espérée*, l'agent doit procéder à des calculs plus complexes, et disposer d'une information initiale plus coûteuse, que pour maximiser son utilité en situation de certitude parfaite ; quant à la *théorie des jeux*, elle présuppose, sur le double plan du coût de calcul et du coût de l'information initiale, un agent qui est sans doute plus dispendieux encore que celui de la théorie de l'utilité espérée.¹

Le point de désaccord est globalement bien marqué. Mais Simon a varié dans l'interprétation qu'il lui donne. Dans son premier article, il avance, encore prudemment, que les économistes n'ont pas apporté la preuve empirique de la validité de leur mécanisme favori (1955, p. 246). Par ailleurs, il n'exclut pas qu'un principe de rationalité globale puisse être complémentaire de son principe de rationalité limitée, en ce sens que le premier servirait de *norme de référence* au second. Ceci peut

s'entendre apparemment de deux manières au moins : ou bien, la rationalité globale conserve une valeur prescriptive pour les agents – à défaut de rendre compte de leurs choix empiriques ; ou bien, elle sert de référence théorique au modélisateur, par exemple comme valeur-limite du modèle simonien, lorsque l'on supprime les contraintes posées, en rationalité limitée, sur les aptitudes cognitives des agents.

Par la suite, Simon portera un jugement beaucoup plus sévère sur l'intérêt du modèle optimisateur. Son article de 1959 sur les théories de la décision souligne assez vigoureusement l'inadéquation empirique de celle des économistes ; la théorie de l'utilité espérée, alors en passe de devenir l'outil préférentiel de la micro-économie, est tout particulièrement visée (1959, in 1966, p. 459). A ce stade, il ne s'agit plus de reprocher au principe de rationalité globale d'avoir été introduit *a priori*, sans justifications positives, mais bien de contredire expressément, dans certains domaines tout au moins,² les données fournies par la psychologie empirique. Quant à la thèse unificatrice d'après laquelle l'optimisation conserverait une valeur de référence soit prescriptive, soit théorique, elle est encore défendue dans l'article de 1959, mais elle s'érode, sans disparaître pour autant tout à fait, dans les travaux des années soixante et soixante-dix.

D'après son discours de Stockholm, il semble que Simon regarde comme vraisemblable jusqu'à un certain point la convergence dynamique du modèle de rationalité limitée à *seuil de satisfaction endogène* vers un modèle optimisateur : même s'il était avéré, ce résultat n'aurait, à ses yeux, que très peu d'intérêt (1979a, p. 503). Par ailleurs, le discours révèle une conception originale de la théorie normative de la décision : on pourrait considérer comme également recommandables deux manières d'assister la prise de décision par les agents – construire des modèles simplifiés compatibles avec un algorithme optimisateur, ou bien des modèles complexes qui appellent un algorithme de décision en rationalité limitée, par exemple la comparaison des résultats avec un seuil de satisfaction (1979a, p. 498). Cette idée, fort attirante, d'une équivalence *pragmatique* entre deux méthodes aux postulations théoriques contradictoires, répond, pour le Simon de 1979, à la question de la valeur prescriptive des modèles optimisateurs.

L'évolution de ses jugements reflète, pour une part, le jeu de causes extrinsèques à sa doctrine. La théorie économique s'est indiscutablement modifiée, depuis 1955, d'une façon décourageante pour les adversaires du modèle optimisateur : celui-ci n'a cessé de renforcer son empire, comme en témoigne, parmi bien d'autres signes, le développement d'une théorie dite des "anticipations rationnelles".³ Il se peut donc que Simon

ait renoncé à changer de l'intérieur une situation disciplinaire qui paraît irrémédiablement figée. Ce sentiment a pu se renforcer de l'impression, avérée de multiples façons, que les économistes ne sont pas prêts à tenir compte des données empiriques accumulées sur les processus de décision, voire sur les résultats individuellement observables de ces processus. C'est peut-être l'une des conséquences pratiques les plus néfastes de la méthodologie de l'"irréalisme", aujourd'hui tellement influente parmi les économistes,⁴ qu'elle délimite *a priori* le champ de l'empirie suivant cette ligne de partage : les modèles économiques n'ont pas à être testés sur leurs prédictions situées les plus en amont — celles qui concernent le *processus* de décision, et même ses résultats *individuellement observables* — mais seulement sur des prédictions situées plus en aval — essentiellement les comportements agrégés d'offre et de demande, c'est-à-dire les résultats *collectivement observables*.⁵

Toutefois, on ne peut ramener les positions successives de Simon à la seule évolution de la discipline économique : elles reflètent aussi certaines ambiguïtés inhérentes à son propre concept de rationalité limitée. Ce sont celles qu'il faut s'efforcer maintenant de faire ressortir.

2.2 *La rationalité limitée: un concept taxinomique*

D'emblée, la question s'est posée de savoir si *limited rationality* (1955), puis *bounded rationality* (1972a, 1979b), renvoyaient à l'une ou l'autre de ces deux idées : celle d'une *rationalité affaiblie* par rapport à la rationalité achevée dont ferait preuve le maximisateur d'utilité espérée ou l'agent de la théorie des jeux ; ou bien l'idée d'une *rationalité achevée dans son ordre*, sur le même pied, donc, que la rationalité optimisatrice, mais qui, contrairement à celle-ci, tiendrait compte des limitations effectives dans les capacités cognitives de l'agent. Cette ambiguïté ressort plus évidemment des textes anciens, comme l'article de 1955. Il vaut la peine de citer un passage où elle nous semble apparaître très clairement :

It is precisely because of [the] limitations on its knowledge and capabilities that the less global models of rationality described here are significant and useful. The question of how it is to behave "rationally", given these limitations, is distinct from the question of how its capabilities could be increased to permit action that would be more "rational" judged from the mountain-top of a more complete model (1955, p. 254).

L'alternative que nous venons de poser – rationalité affaiblie, d'une part,

rationalité achevée dans l'ordre particulier des limitations cognitives, d'autre part – n'est sans doute pas très facile à préciser conceptuellement. Mais elle permet d'indiquer l'intervalle de variation extrêmement large sur lequel fluctuent les expressions de *limited* ou *bounded rationality*. Cette souplesse d'emploi va se justifier à la lumière d'un trait généralement mal perçu de la démarche de Simon. Celui-ci ne se contente pas de défendre certains principes de modélisation particuliers : il veut aussi recommander l'usage d'une notion *générique* de rationalité limitée. De là certaines hésitations apparentes dans les jugements qu'il porte sur l'optimisation. Quelquefois, il paraît la rejeter complètement : c'est qu'il la juge du point de vue d'un modèle spécifique, non optimisateur, de la rationalité limitée. Quelquefois aussi, il lui fait place dans la théorie de la décision : c'est qu'il l'apprécie alors d'un point de vue générique – point de vue qui apparaît éventuellement compatible avec l'approche traditionnelle.

Un article de 1972 éclaire parfaitement ce dernier mouvement de la démarche simonienne. Il contient une définition explicite de la rationalité limitée : "on peut désigner comme théories de la rationalité limitée (*bounded rationality*) celles qui incorporent des contraintes sur la capacité de l'acteur à traiter de l'information" (1972a, p. 162). Cette définition a le mérite de rattacher la distinction des types global/limité à un principe taxinomique extrêmement clair. Il se trouve qu'elle s'applique fort bien à certains modèles optimisateurs : il suffit de rendre plus complexe le modèle classique en surajoutant aux contraintes *externes*, comme le revenu ou la technologie, des contraintes *internes* qui décrivent l'absence d'information ou la difficulté à s'en procurer. Tel est, justement, le programme que se donne la théorie néo-classique du *search*. Simon le reconnaît très volontiers, puisque, dans le même texte, il affirme que l'on peut construire des modèles de rationalité limitée en s'aidant des idées de Stigler (p. 163, note 3). Est-ce à dire que Simon cède du terrain à l'adversaire ? Nous ne le croyons pas : dans l'article de 1972, il cherche moins à défendre son modèle particulier de rationalité limitée qu'à caractériser *génériquement* cette notion.

Il n'en va plus de même dans son discours de Stockholm, où il condamne très fermement le projet de l'école du *search* :

Stigler a versé le vin neuf de la théorie de la recherche dans les vieilles outres de la maximisation de l'utilité, puisque le coût de la recherche est [dans son modèle] égalisé avec son rendement marginal . . . Mais la maximation de l'utilité n'était en rien essentielle au schéma de la recherche – heureusement d'ailleurs, car, sinon, il aurait fallu que le décideur fût capable d'estimer les coûts et rendements marginaux de la recherche dans un contexte de décision qui est déjà trop complexe pour que s'y applique une forme de rationalité globale (1979a, pp. 502–503; cf. aussi 1978b).

Cette prise de position – sur laquelle il faudra revenir en détail – n’appartient pas au même niveau logique que la définition de 1972. Simon parle maintenant du point de vue de son modèle *spécifique* de la rationalité limitée: les tentatives de Stigler sont vouées à l’échec, même si, formellement parlant, elles tombent sous la même définition que le modèle simonien.

Ainsi, l’expression *bounded rationality* aurait, dans l’esprit de Simon, un rôle essentiellement taxinomique et signalétique : c’est par cette conjecture que nous proposons d’unifier les positions apparemment disparates qu’il prend sur l’optimisation et les modèles du *search*. En d’autres termes, l’expression *bounded rationality* ne décrirait pas complètement l’objectif que se propose Simon. Sous cette forme, l’hypothèse interprétative devient en quelque sorte testable : Simon a-t-il explicité d’autres notions que celle de rationalité limitée lorsqu’il veut caractériser *spécifiquement* sa démarche ? Or il se trouve qu’à deux reprises au moins, Simon a choisi d’abandonner la dichotomie rationalité globale/rationalité limitée au profit d’un autre couple théorique: en 1976 et 1978, il caractérise son approche de la décision comme “procédurale” (*procedural*) par opposition à l’approche “substantielle” (*substantive*) des économistes.

2.3. De la rationalité limitée à la rationalité procédurale

Contrairement à celle de 1972, la dichotomie de 1976–1978 ne vient pas s’insérer dans une définition au sens strict du mot. Elle peut se formuler de deux manières au moins, qui ne sont sans doute pas logiquement équivalentes :

(1) Les modèles de rationalité procédurale s’intéressent à la rationalité du *processus* du choix : ceux de la rationalité substantielle privilégient la rationalité des *résultats* du choix.

(2) Est procédurale la conception qui met l’accent sur l’aspect délibératif de la décision, c’est-à-dire qui insiste sur le fait que les conditions du choix, du côté des fins ou des fonctions objectifs, aussi bien que du côté des moyens ou des contraintes du choix, ne sont pas données au décideur, mais font l’objet d’une recherche. Est “substantielle” la conception qui regarde les conditions du choix comme fixées, et identifie donc la décision à la simple application d’un critère d’évaluation *donné* à un ensemble des possibles également *donné* (cf. 1976, pp. 130–132 et 1978, pp. 2, 8–9).

On peut s'étonner de l'occurrence à la fois rare et tardive, dans l'œuvre de Simon, des expressions "rationalité procédurale" et "rationalité substantielle". En fait, la distinction des processus et des résultats est loin d'être absente des premiers articles de Simon.⁶ Mais elle n'a pas alors livré encore toutes ses potentialités. Dans le *Dictionary of the Social Sciences* (1964), Simon peut écrire : "La rationalité se rapporte tantôt aux processus de choix qui recourent à la faculté intellectuelle, tantôt aux choix eux-mêmes. La première connotation caractérise l'usage ancien en psychologie, logique ou éthique ; la seconde connotation correspond à l'usage prédominant en économie et en sociologie." En 1976, il reprend la distinction des processus et des résultats, mais pour la faire coïncider avec celle de deux programmes de recherche, et non plus de deux orientations disciplinaires. Même si chaque discipline comporte une orientation préférentielle vers l'un des deux pôles, elle peut très bien faire coexister des recherches du type *substantive* et des recherches du type *procedural*. Ainsi, l'économie n'apparaît pas entièrement fermée aux préoccupations de rationalité procédurale (1976, p. 137). Cette présentation nouvelle de la distinction est sans doute plus exacte que la précédente. Elle en change aussi complètement la portée : des programmes de recherche peuvent entrer en concurrence, ce qui n'est pas le cas de spécialités disciplinaires. Pour le Simon de 1976, la rationalité substantielle et la rationalité procédurale ont le même objet — la décision individuelle — qu'elles appréhendent très différemment; désormais, la question se pose de savoir quelle est la vision la plus féconde.

Le bilan des deux programmes, esquissé dans les articles de 1976 et 1978, trahit une préoccupation d'équilibre : Simon se garde de cautionner exclusivement le point de vue délibératif ou procédural. Il semble parfois même tirer la distinction du procédural et du substantiel vers la taxinomie pure, exactement comme il avait choisi d'interpréter, en 1972, la distinction global/limité. Néanmoins, sa conclusion est univoque : le programme substantiel a été mené aussi loin qu'il était possible, on peut douter qu'il débouche encore sur de nombreux résultats, alors que son rival vient d'entrer dans sa phase la plus féconde.

Il nous semble donc que l'évolution des idées de Simon l'a conduit, dans une large mesure, à *déclasser l'un des deux couples conceptuels au profit de l'autre*. Certes, il ne renonce pas à l'idée d'élaborer une théorie de la rationalité limitée, mais il perçoit mieux, à partir d'un certain moment, le trait qui doit la spécifier : l'aspect procédural, négligé dans l'approche classique. Il serait intéressant de confronter cette hypothèse interprétative avec l'histoire des idées de Simon dans le domaine auquel

il s'intéresse préférentiellement à partir de 1958–1960: l'intelligence artificielle. Ce rapprochement dépasserait le cadre de cet article. Mais une autre cause a dû jouer, dont nous pourrions démontrer plus facilement le mécanisme: l'apparition d'une théorie néo-classique du *search* a contraint Simon à préciser le point exact de son désaccord avec l'optimisation et, sans doute, à comprendre ce qu'il y avait d'exagérément indéterminé dans le projet de faire une théorie de la rationalité limitée.

La section 3 va tenter de clarifier la position que nous prêtons à Simon en construisant une notion *générique* de rationalité limitée.

3. Une notion générique de la rationalité limitée

3.1. La rationalité limitée et la typologie des incertitudes

Pour définir précisément ce qu'est le modèle générique de la rationalité limitée, nous recourons à un formalisme rudimentaire qui intervient, pour ainsi dire, dans toutes les théories de la décision.⁷ Celles-ci reposent traditionnellement sur la donnée de trois ensembles:

- A*, celui des actions ;
- S*, celui des conséquences ;
- E*, celui des états du monde.

Un postulat discutable, mais très généralement employé, veut que l'agent n'attribue pas directement une valeur à ses actions, mais seulement aux conséquences: l'évaluation "remonte", en quelque sorte, des conséquences aux actions.⁸ On se donnera donc une fonction d'évaluation $V : S \rightarrow \mathbb{R}$. Quant à l'information de l'agent, on peut la modéliser de bien des façons, mais notre propos, ici, doit être de rester le plus général possible. Avant de tenter une réponse, il sera utile de considérer et de comparer deux formalisations particulières, celle de Savage dans son ouvrage classique, *The Foundations of Statistics* (1954), et celle de Simon en 1955.

Savage considère que l'agent est "incertain" des conséquences de son action en un sens bien déterminé de ce terme : pour chaque action $a \in A$ considérée, la conséquence qui va se produire dépend aussi de l'état du monde $e \in E$ réalisé ; or l'agent l'ignore. Cette incertitude est parfaitement circonscrite : si l'agent connaissait e , il connaîtrait la conséquence qui découle du choix de a . Pour formaliser cette idée, Savage

pose, pour chaque $a \in A$, l'existence d'une *application* bien définie :

$$\alpha : \epsilon \rightarrow S$$

et c'est une telle application, plutôt que l'élément correspondant a de A , qu'il appelle d'ailleurs un "acte" (1954, p. 14).

Dans son article de 1955 (dont nous essaierons, dans la mesure du possible, de reprendre les notations), Simon décrit ainsi l'incertitude de l'agent: à chaque action $a \in A$, celui-ci associe un sous-ensemble $S_a \subseteq S$ (cf. 1955, pp. 248–250). Une particularité remarquable de l'article de 1955 est qu'il ne fait pas intervenir expressément l'ensemble des états du monde E . Simon ne peut donc pas dire ce qui, dans l'incertitude de l'agent, dépend de son ignorance de l'état du monde ou bien de tout autre facteur. Ce refus de faire intervenir l'ensemble E est à première vue surprenant ; mais il peut s'interpréter comme une préoccupation de généralité. Considérons le formalisme de Savage : à chaque $\alpha \in A$, il associe un sous-ensemble $\alpha(E) \subseteq S$, qui est l'image de l'application α elle-même associée à a . Savage entre donc dans la formulation de Simon à titre de cas particulier : comme lui, il définit l'incertitude par l'existence d'une correspondance $\Psi : a \rightarrow S_a, A \rightarrow P(S)$ — ce symbole désignera l'ensemble des parties de S — mais il se trouve que, pour Savage, le caractère non fonctionnel de la relation tient *uniquement* à l'ignorance, par l'agent, de l'état du monde réalisé ; Simon, quant à lui, préfère ne donner aucune précision sur l'origine de l'incertitude.

La comparaison de ces deux définitions est très éclairante pour l'analyse de la rationalité limitée. Celle-ci se caractérise, on l'a dit, par l'existence de "contraintes sur la capacité de l'acteur à traiter de l'information". Or voici le critère le plus naturel, sans doute, qui permette de tester la présence de ces contraintes : supposons que l'agent connaisse l'état du monde e réalisé ; alors, il est encore en situation d'incertitude, c'est-à-dire il continue d'associer aux différents $a \in A$ des S_a qui, en général, ne se réduisent pas à un seul point. L'idée de cette définition de la rationalité limitée est fort simple : si, par hypothèse, on annule l'incertitude *externe* qui affecte la décision, il subsiste de l'incertitude *interne*. Le joueur d'échecs constitue peut-être l'illustration canonique de cette définition.⁹ En bref, le sujet de décision en rationalité limitée est, par construction même, un *sujet non savagien*.¹⁰

Pour formaliser les idées précédentes, le mieux est d'introduire comme terme primitif la correspondance :

D1

$$\chi : E \times A \rightarrow P(S)$$

qui décrit l'information maximale dont dispose l'agent, lorsqu'il identifie à la fois l'action choisie et l'état du monde réalisé. La correspondance seule envisagée par Simon, $\Psi : A \rightarrow p(S)$, s'introduit alors comme un symbole dérivé, par la définition :

$$(D2) \quad (\forall a \in A) \quad \Psi(a) \stackrel{\text{déf}}{=} \bigcup_{e \in E} \chi(e, a)$$

Le cas particulier de Savage (*incertitude purement externe*) s'écrit:

$$(Cp1) \quad (\forall a \in A) \quad (\exists ! \alpha \in S^E) \quad (\forall e \in E) \\ \{ \alpha(e) \} = \chi(e, a)$$

Cette formule et la définition de Ψ ont pour implication la propriété déjà signalée:

$$(\forall a \in A) \quad \Psi(a) = \alpha(E)$$

Pour des raisons de symétrie, on peut vouloir définir aussi le cas polaire d'une *incertitude purement interne*. Plaçons-nous dans ce cas, et supposons que l'agent passe d'un état d'ignorance complète sur l'état du monde réalisé à la connaissance de cet état du monde \bar{e} : alors, l'agent doit être encore incertain des conséquences attachées à une action quelconque a , et cette incertitude doit être exactement la même que précédemment. Donc:

$$(Cp2) \quad (\forall a \in A) \quad (\forall e \in E) \left\{ \begin{array}{l} \chi(e, a) \text{ en général non réduit à un point} \\ \text{et} \\ \chi(e, a) = \psi(a) \end{array} \right.$$

Quant au cas général (incertitude à la fois externe et interne), il se définit par cette condition moins forte : lorsque l'agent passe d'une ignorance complète sur l'état du monde réalisé à la connaissance de cet état du monde, alors l'agent est encore incertain des conséquences attachées à une action quelconque a . Donc:

$$(Cg.RL) \quad (\forall a \in A) \quad (\forall e \in E) \quad \chi(e, a) \text{ en général non réduit à un point.}$$

Cette condition est évidemment équivalente à la négation du cas particulier (Cp1). Elle nous servira de premier critère de la rationalité limitée:

celle-ci se caractérise par la présence d'une incertitude qui n'est pas purement externe.

Jusqu'à présent, nous avons supposé implicitement que l'agent connaissait l'ensemble des *actions* qui lui sont disponibles, de sorte que l'incertitude portait uniquement sur la détermination des *conséquences*. Or l'existence de contraintes internes peut *a priori* revêtir la forme d'une incertitude sur A aussi bien que sur S . Pour la décrire formellement, le mieux est de suivre un usage courant¹¹ de la théorie de la décision et d'identifier à une *partition* $\pi[A]$ de A l'information dont l'agent dispose sur cet ensemble. Il faut alors redéfinir les applications et en remplaçant A par $\pi[A]$ et $a \in A$ par $p \in \pi[A]$. Intuitivement, l'élément p représentera le sous-ensemble de A dont l'agent perçoit tous les éléments comme identiques.

Un tel formalisme paraît spécialement approprié lorsqu'une action du décideur représente en fait une stratégie en plusieurs étapes, comme dans l'exemple classique du jeu d'échecs : au premier coup, jouer le pion $d4$; au second coup, si l'adversaire a joué $d5$, jouer le pion $c4$, au troisième coup... Formellement parlant, une action est alors un chemin parcouru sur un arbre à branchements multiples, en allant de la racine de cet arbre jusqu'à l'un de ses sommets terminaux. Et l'élément $p \in \pi[A]$ est un sous-ensemble d'actions qui sont engendrées par un même branchement et coïncident donc deux à deux sur une partie initiale des chemins parcourus.

Il est significatif, là encore, que cette description des a et des p en termes d'arbre permette de retrouver le formalisme de Savage comme un cas particulier. La théorie savagienne de l'incertitude analyse effectivement les "actes" comme des chemins (cf. 1954, p. 15). Mais elle ignore les ensembles d'information $\pi(A)$, puisqu'elle suppose que l'ensemble des "actes" est parfaitement connu de l'agent : celui-ci est, par hypothèse, capable de parcourir mentalement tous les chemins jusqu'à leur sommet terminal. Lorsqu'on fait intervenir des "contraintes sur la capacité de l'acteur à traiter l'information", il convient, naturellement, de renoncer à cette hypothèse : on remplacera les a par des p . On retrouve donc, à propos des actions, l'idée fondamentale qui nous était déjà apparue relativement aux conséquences : la théorie de la rationalité est par nature *non savagienne*. Ou, de façon strictement équivalente : on parlera de rationalité limitée *chaque fois que l'incertitude de l'agent n'est pas uniquement externe*, soit qu'elle porte sur l'ensemble des actions, soit qu'elle porte sur l'ensemble des conséquences.

La discussion de cette section 3.1 sur les types d'incertitude nous a-t-elle livré une définition complète de la rationalité limitée? On peut en

douter si l'on se reporte à l'exposé de Simon dans sa conférence de Stockholm. Sans vouloir, dans ce texte, proposer une *définition* au sens strict, Simon y énonce la *caractérisation* suivante: la rationalité limitée met en oeuvre les deux concepts fondamentaux de *recherche* et de *satisfaction*.¹² Quel lien y a-t-il entre notre tentative précédente de définition et cette caractérisation moins abstraite? Faut-il modifier celle-là pour tenir compte de celle-ci?

3.2. "Search" et "satisficing"

Quand Simon parle de *search*, il crée une amphibologie qu'il supprime aussitôt. Il ne s'agit pas de recommander la théorie stiglérienne du choix, mais de rappeler cette idée simple: en rationalité limitée, l'agent procède à une recherche, soit des conséquences, soit des actions elles-mêmes: "If the alternatives for choice are not given initially to the decision maker, then he must search for them. Hence a theory of bounded rationality must incorporate a theory of search" (1979a, p. 502). Pris à ce niveau d'abstraction, le principe de recherche nous semble découler analytiquement de la discussion précédente et de l'idée même de rationalité: en situation d'information interne incomplète, il est normal (= rationnel) que l'agent cherche à améliorer sa connaissance des actions et des conséquences avant d'arrêter son choix, donc que l'on assiste à un processus de recherche plutôt qu'à une décision instantanée. Quelle est la notion de rationalité ici sous-jacente? Il ne faut, à l'évidence, qu'une détermination très pauvre – on aimerait presque la rapprocher de la thèse cartésienne célèbre suivant laquelle il convient d'éviter précipitation et prévention. Ainsi, *le propre de la rationalité limitée est qu'elle ignore l'instantanéité*. Au contraire, la rationalité globale exclut le principe de recherche (au sens abstrait), puisqu'elle dispose d'emblée de tous les moyens des choix.

Le principe du *satisficing* qu'évoque ensuite Simon est, on va le voir, d'une tout autre nature épistémologique. Rappelons qu'il intervient dans le contexte d'une recherche séquentielle sur un ensemble et se définit ainsi: chaque fois que l'agent découvre un nouvel élément de l'ensemble, il redéfinit son niveau d'aspiration et le confronte à tous les éléments déjà connus (y compris, naturellement, celui qu'il vient de découvrir); si l'un de ces éléments satisfait au niveau d'aspiration ainsi révisé, il est retenu, et la recherche s'arrête; si ce n'est pas le cas, la recherche se poursuit. Par exemple, dans le cas où l'ensemble exploré est celui des *actions*, un modèle de *satisficing* doit comporter:

- un mécanisme d'engendrement des actions successivement examinées, et
- un mécanisme d'endogénéisation partielle du niveau d'aspiration.

Formellement parlant, si A est fini ou dénombrable, le premier mécanisme est la donnée des éléments de A sous forme de suite: $a_1, a_2 \dots a_n \dots$; et le second consiste en une infinité dénombrable de fonctions numériques: $k_n(I_n, V(a_1), V(a_2) \dots V(a_n))$, $n \geq 1$, où les $V(a_i)$, $1 \leq i \leq n$, sont les valeurs d'utilité correspondant aux n actions déjà examinées, et I_n désigne les autres informations disponibles à l'étape n . L'agent arrêtera sa recherche à l'instant n si et seulement si

$$(R) \quad (\exists i) \quad (1 \leq i \leq n) \quad V(a_i) \geq k_n(I_n, V(a_1), V(a_2) \dots V(a_n))$$

Dans la pratique de la modélisation, on se donne souvent un *ordre canonique sur A* , ce qui revient à supposer résolue la question de l'engendrement des éléments. Par ailleurs, nombreuses sont les illustrations qui font intervenir seulement le cas, extrêmement particulier, d'un *niveau d'aspiration constant* tout au cours de la recherche.¹³

On voit que le principe du *satisficing* suppose, dans sa formulation même, celui de la recherche séquentielle, c'est-à-dire le *search*.¹⁴ Néanmoins, il ne s'y ramène pas. En effet, (R) n'est pas la seule règle d'arrêt qui soit *a priori* concevable: l'agent pourrait se fixer à l'avance le nombre d'étapes qu'il va parcourir, voire même se donner un nombre d'étapes révisable au vu de ses observations successives; il pourrait combiner (R) à certains critères de régularité imposés sur la suite $a_1 \dots a_n$, etc. Ces différentes manières de clore une recherche séquentielle ne sont pas *a priori* irrationnelles. C'est dire que le principe du *satisficing* ne s'impose pas *a priori*. Par rapport aux considérations qui le précèdent, il introduit une discontinuité épistémologique. Jusqu'au principe du *search* inclusivement, nous n'avons fait qu'explorer les conséquences analytiques d'une certaine définition de la rationalité limitée. Avec les seuils de satisfaction apparaissent maintenant des considérations qui, sans doute possibles, sont empiriques. Simon et les théoriciens de la rationalité limitée se sont ici directement inspirés des données de la psychologie, notamment de certains travaux de Kurt Lewin (cf. Simon 1959/1966, pp. 10 sq; 1979a, p. 503). Notre propos, dans cette section 3, est de déboucher sur une notion générique de rationalité limitée: celle-ci est-elle compatible avec une hypothèse aussi déterminée que celle des seuils de satisfaction?

Pour prendre la mesure du problème, il peut être utile de comparer la formulation “*search plus satisficing*” à celle que Simon et ses collaborateurs ont retenue dans leurs travaux d’intelligence artificielle. Un article de 1972, par exemple – écrit avec Newell et Shaw – décrit ainsi les constituants d’un programme joueur d’échecs: des “générateurs de coups”, qui constituent le mécanisme d’engendrement sur A ; des “évaluateurs”, qui correspondent à la fonction $V(\cdot)$; enfin des “règles d’arrêt”, qui portent, respectivement, sur la recherche des conséquences à action donnée, et sur la recherche des actions envisageables.¹⁵ Cette description frappe par son extrême généralité: le *satisficing* s’y insère naturellement, mais comme un cas particulier. Ne faudrait-il pas préférer le Simon informaticien au spécialiste de théorie de la décision, et choisir, pour une définition générique de la rationalité, l’idée abstraite d’une *règle d’arrêt* plutôt que le principe trop déterminé des seuils de satisfaction? Une réponse positive s’impose sans doute sur le fond, et les protestations que certains psychologues empiriques ont élevées contre un usage trop exclusif des seuils de satisfaction viennent la conforter après coup. Mais il faut tenir compte de l’existence d’une caractérisation coutumière de la rationalité limitée. Nous nous proposons donc de ne pas trancher dans le vif et d’opérer alternativement avec le couple “*search plus satisficing*” et avec une définition générique complètement *a priori* de la rationalité limitée.

3.3. Proposition d’une définition générique

Celle-ci consistera dans la donnée d’êtres mathématiques définis ou définissables en termes ensemblistes, sur lesquels un *minimum de contraintes* sont imposées *a priori*. En voici la liste.

(1) Les ensembles A , E et S , la fonction $V : S \rightarrow \mathbb{R}$, la partition $\pi(A)$ qui décrit l’ensemble d’informations initial de l’agent sur A , la correspondance $\chi : E \times \pi(A) \rightarrow P(S)$, soumise à la condition.

(Cg' .RL) ($\forall p \in \pi(A)$) ($\forall e \in E$) $\chi(e, p)$ en général non réduit à un point.

Toutes ces données permettent de fixer le point de départ de la recherche.

(2) Un mécanisme d’engendrement et une règle d’arrêt pour la recherche menée sur les *actions*, l’état du monde $e = \bar{e}$ étant fixé par ailleurs.

(3) Un mécanisme d’engendrement et une règle d’arrêt pour la recherche menée sur les *conséquences* de chaque action, l’état du monde $e = \bar{e}$ étant fixé par ailleurs.

Il faut naturellement que les processus (2) et (3) soient compatibles entre eux – cette contrainte a pour effet de rendre leur formalisation relativement ardue. Nous nous contenterons de poser deux exigences, raisonnables *a priori*, que ces processus doivent respecter. Pour chaque état donné de l'information sur les actions, état que l'on décrit par une partition $\pi_m(A)$, le mécanisme d'engendrement du (2) doit produire une *information de plus en plus fine sur les conséquences*. On demandera donc qu'il conduise à une suite de correspondances $\chi_0^m, \chi_1^m, \dots, \chi_n^m, \dots : E \times \pi_m(A) \rightarrow P(S)$ vérifiant

$$(C1) \quad \begin{aligned} \chi_0^m : (\bar{e}, p) &\rightarrow S_p^{m,0}(\bar{e}) \\ \chi_1^m : (\bar{e}, p) &\rightarrow S_p^{m,1}(\bar{e}) \subseteq S_p^{m,0}(\bar{e}) \\ &\dots \\ \chi_n^m : (\bar{e}, p) &\rightarrow S_p^{m,n}(\bar{e}) \subseteq S_p^{m,n-1}(\bar{e}) \\ &\dots \end{aligned}$$

Autrement dit, l'agent "raffine" sa correspondance d'informations initiale par étapes successives.¹⁶ De même, pour chaque état donné de l'information sur les conséquences, le mécanisme d'engendrement du (3) devra produire une *information de plus en plus fine sur les actions*. On demandera donc que, $\pi_m(A)$ et χ_n^m étant données, on ait

$$(C2) \quad \pi_{m+1}(A) \text{ au moins aussi fine que } \pi_m$$

L'absence de toute référence aux seuils de satisfaction présente des avantages de simplicité que l'on perçoit sans doute mieux une fois que l'on s'est interrogé sur le problème suivant: supposons que l'on cherche à formuler la définition générique des règles d'arrêt en termes de seuil; on ne rencontrera pas de difficultés essentielles pour le processus (2), mais il n'en ira pas de même pour (3). Comment faut-il, en effet, construire la fonction d'évaluation implicite dans l'arrêt d'une recherche *sur les conséquences*? Faut-il dire que l'agent évalue les conséquences qu'il découvre peu à peu à l'aide de la fonction V , et s'arrête lorsqu'il parvient à un sous-ensemble de conséquences $S_p^{m,n}(\bar{e})$ qui satisfont *toutes* au seuil d'aspiration? Cette proposition, fort simple, qui est d'ailleurs retenue par Simon en 1955 (p. 248), semble déjà trop particulière: l'agent pourrait se contenter d'un sous-ensemble $S_p^{m,n}(\bar{e})$ *globalement satisfaisant* (en un sens statistique à définir). De façon générale, la relation mathématique qu'entretiennent la fonction d'évaluation *terminale* V et la fonction d'évaluation *induite* W (qui est définie sur les $S_p^{m,n}(\bar{e})$) peut

prendre des formes extrêmement diverses.¹⁷ Une formulation entièrement générale de cette relation n'est certainement pas facile à écrire.

On reprochera peut-être à notre définition le parti pris d'abstraction qui l'inspire. Contre cette objection, nous ne pourrions qu'invoquer les principes philosophiques de la distinction du genre et de l'espèce: c'est la *differentia specifica* qui fait apparaître l'information utile, mais celle-ci ne peut ressortir que d'une définition préalable du genre. Il importe également de souligner ce point, qui est de nouveau conforme aux principes théoriques de la classification: la définition du genre "rationalité limitée" est sans doute peu informative, mais elle n'est pas vide de tout contenu empirique. Par construction, cette définition exclut le cas d'une information interne complète. Ceci a une conséquence empirique décisive: *un processus de décision en rationalité limitée ne peut déboucher sur l'optimum que par accident*. En effet, l'absence d'information interne complète implique le caractère séquentiel de la recherche sur les ensembles A et S , et celui-ci implique à son tour que, pour arrêter son choix final, l'agent n'attendra généralement pas de connaître la totalité de ces deux ensembles. Or seule la connaissance *en toutes circonstances* de la totalité de A et de S garantirait que l'agent s'arrête *en toutes circonstances* sur l'optimum. Dans un processus de rationalité limitée, celui-ci ne sera donc réalisé qu'exceptionnellement. On a parfois souligné le caractère sous-optimal (en ce sens déterminé) des processus néo-classiques de *search* (Marciano, 1982, p. 19), ou des processus de recherche avec seuil de satisfaction endogène (Mongin, 1984b, pp. 40–41). Mais il faut bien voir que cette propriété tient simplement au caractère séquentiel de la recherche, et non pas aux modalités particulières retenues pour les règles d'arrêt.

Une limitation plus sérieuse de la définition précédente tient au fait qu'elle repose sur cette expérience *contre-factuelle*: supposons que l'incertitude externe disparaisse, alors il subsiste de l'incertitude interne. En termes équivalents, on pourrait nous reprocher de *supposer déjà connue la structuration de l'ensemble E des états du monde*. Notre définition ne peut rien dire sur la façon dont interagissent l'incertitude interne et l'incertitude externe: or la théorie de la rationalité limitée doit aussi se proposer de saisir cette interaction. Seules des considérations pragmatiques de simplicité nous ont conduit à ignorer cet aspect du concept.

4. Une notion spécifique de rationalité limitée: celle du *search*

4.1. Du consommateur de Stigler au joueur d'échecs

Simon, on l'a vu, reconnaît facilement que les modèles néo-classiques de *search* s'insèrent dans sa définition de la rationalité limitée. Ce faisant, il suppose résolue une difficulté que l'on rencontre dès que l'on tente de mettre en rapport une définition quelconque de la rationalité limitée avec les travaux issus de Stigler. Rappelons que, dans son article célèbre de 1961, celui-ci étudiait le cas d'un consommateur décidé à acheter une unité d'un bien homogène, mais incertain des prix auxquels ce bien est vendu dans les différents magasins de son quartier. Ce consommateur entreprendra donc une recherche séquentielle du meilleur prix. Le problème est de formuler sa règle d'arrêt optimale. Quel rapport y a-t-il, demandera-t-on peut-être, entre cette recherche de type *externe* et la recherche d'un agent en rationalité limitée? L'expression simonienne que nous avons plusieurs fois citée – *constraints on the information-processing capacities of the actor* – pourrait à la rigueur s'appliquer directement à l'exemple choisi par Stigler. Néanmoins, la définition générique de la section 3 est plus contraignante. Elle repose sur une distinction de l'incertitude interne et de l'incertitude externe. Or, les hypothèses explicites du modèle de Stigler l'apparentent, sans aucun doute possible, à un cas d'incertitude *externe*: si le consommateur ignore les conséquences précises qui s'attachent à une certaine action (= acheter dans le n^{ème} magasin), c'est uniquement parce qu'il ignore l'état du monde correspondant (= le prix qui est pratiqué dans le n^{ème} magasin). Suivant les termes exprès de Stigler, l'agent est donc en situation savagienne, et la question se pose aussitôt de savoir comment on peut élaborer une conception de la rationalité limitée à partir d'un exemple aussi manifestement inadéquat.

Cette première difficulté va se dissiper très facilement, car les hypothèses explicites du modèle de 1961 sont moins importantes que le principe général d'analyse qu'il met en oeuvre: en tout état de cause, il s'agit de décrire l'ajustement marginal que l'agent effectue entre *la réalisation d'un certain objectif primaire* (dans le cas particulier: acheter au plus bas prix) et *l'acquisition d'une information supposée coûteuse* (pour connaître ce prix, l'agent doit se rendre dans un magasin, et, pour cela, encourir un certain coût). Suivant une formulation légèrement différente, mais qui s'avère logiquement équivalente, *l'agent déplace son*

problème de décision vers le second ordre: il égalise le coût marginal de la recherche d'information et son rendement marginal brut.¹⁸ Sous l'une ou l'autre forme, le principe imaginé par Stigler est assez général pour s'appliquer au cas d'une recherche interne aussi bien qu'à celui d'une recherche externe. Il suffira, à vrai dire, de partir d'un modèle bien choisi de recherche du consommateur, d'en préserver le formalisme, et d'en modifier la sémantique. *On passera alors du consommateur stiglierien au joueur d'échecs simonien par simple redéfinition des variables.*

Voici quelles sont les variables du modèle initial de Stigler: la suite $c_1, c_2, \dots, c_k, \dots$ des coûts d'observation à chaque étape, et la loi de répartition empirique des prix $F(p)$, chacune étant supposée connue du consommateur. Par la suite, on a généralisé le modèle dans deux directions. D'une part, au lieu de raisonner sur des coûts d'observation monétaires et d'ignorer la fonction d'utilité de l'agent, on a fait intervenir une fonction d'évaluation $V(p)$ et des coûts d'observation $C_1, C_2, \dots, C_k, \dots$ exprimés dans la même unité que $V(p)$ (p. ex. Kohn et Shavell, 1974). D'autre part, on a remplacé l'hypothèse, très forte, d'une loi empirique connue de façon certaine, par celle-ci, qui est plus générale: l'agent forme une loi de probabilité *subjective* sur p , $F^0(p)$, qu'il révisé ensuite après avoir recueilli chaque observation, d'où la suite $F^1(p), \dots, F^k(p), \dots$ (cf. Kohn et Shavell, 1974; Rothschild, 1974). C'est ce modèle plus général, et non pas celui de Stigler, dont nous allons transposer le formalisme pour rejoindre une situation de rationalité limitée.

La méthode de transposition est fort simple: il faut annuler l'incertitude *externe*. On posera donc que, lorsque le consommateur connaît un prix p , il sait aussi dans quel magasin m_p ce prix p est pratiqué.¹⁹ En revanche, il ne dispose pas *mentalement* de tous les couples (p, m_p) , mais doit les rechercher. Cette recherche peut apparaître à l'un ou l'autre de ces deux points de vue: une recherche sur l'ensemble des actions disponibles, une recherche sur les conséquences d'une action donnée. Ou bien, l'on regarde chaque couple (p, m_p) comme action possible (= acheter une unité du produit *au prix p et dans le magasin m_p*) dont les conséquences sont parfaitement connues, et la recherche porte alors sur les $a_1, \dots, a_m, \dots \in A$. Ou bien, l'on considère l'action du consommateur comme fixée (= aller dans un magasin et acheter une unité du produit *sans autre précision*), et la recherche porte sur les conséquences $S_a^0, \dots, S_a^n, \dots$ de cette action (= la nature du magasin et le prix acquitté). Dans l'un et l'autre cas, on a maintenant affaire à une incertitude et à un processus de recherche, qui sont du type *interne*.

Quant à la loi de probabilité $F^0(p)$, elle sert dorénavant à décrire et mesurer l'incertitude subjective de l'agent ou bien sur l'ensemble A

(première version), ou bien sur l'ensemble $S = S_a^0$ des conséquences qui, initialement, sont susceptibles d'être associées à l'action fixée a (deuxième version). Les lois *a posteriori* $F^1(p), \dots, F^n(p), \dots$ correspondent aux différentes étapes de révision de cette loi *a priori*. Dans la seconde interprétation, les $S_a^0, \dots, S_a^n, \dots$ s'interprètent comme les *supports* des lois de probabilité successivement envisagées par le consommateur. La réduction dans la taille du support S_a^n se traduit par une réduction de la variance pour la loi associée $F^n(p)$.

On pourrait maintenant vouloir effectuer un pas de plus et produire grâce à une nouvelle réinterprétation des variables *un modèle néo-classique du joueur d'échecs*. Cette opération n'a guère d'intérêt en elle-même, et l'on a compris le propos qu'elle servirait à illustrer: utilisés d'abord pour résoudre des problèmes économiques bien déterminés (choix du consommateur ou encore du demandeur d'emploi), les formalismes du *search* sont assez souples pour servir à élaborer *une théorie néo-classique de la rationalité limitée*.

Voici comment ils s'insèrent finalement dans la définition générique du 3.3 (*supra* p. 570).

(1) Comme pour tout autre modèle de rationalité limitée, on se donnera des ensembles A, E et S , une fonction d'évaluation V et éventuellement une partition initiale $\pi(A)$.

(2') et (3') Les règles d'engendrement des processus seront de même nature dans le cas des actions et des conséquences: il s'agit de lois de probabilité subjectives, révisées normalement après chaque observation, $F^1(p), \dots, F^n(p)$. De même, les règles d'arrêt obéiront à un principe uniforme: on retiendra celles qui assurent l'utilité maximale (au sens de V), étant donné la suite des lois subjectives révisées $F^1(p), \dots, F^n(p)$, et la suite des coûts d'observation C_1, \dots, C_n , (mesurés dans la même unité que V).

Le trait saillant des modèles de *search* tient peut-être dans le double rôle qu'ils font jouer aux lois de probabilité subjectives: elles tiennent lieu de *mécanisme d'engendrement* et elles interviennent aussi, par le biais des égalités marginales, dans la *règle d'arrêt* de chaque processus de recherche.

Pour décrire plus concrètement l'apport de la théorie néo-classique du *search* à une élaboration de la rationalité limitée, il faut invoquer un résultat important, dû, sous cette forme générale, à Kohn et Shavell (1974): si l'on accepte certaines hypothèses apparemment peu contraignantes sur les processus de recherche du type (2') - (3'), il est optimal de les arrêter suivant une règle de *satisficing*. Ce théorème, que nous

allons maintenant rappeler, constitue, à nos yeux, le véritable intérêt des modèles post-stiglériens pour une théorie de la rationalité limitée.

4.2. Une théorie néo-classique du "satisficing"

Les modèles post-stiglériens se divisent grossièrement en deux classes, suivant qu'ils traitent plutôt de la recherche du consommateur ou plutôt de celle du demandeur d'emploi. Dans le premier cas, on admet généralement qu'une hypothèse dite de *perfect recall* s'applique bien: le consommateur peut décider d'accepter toute proposition qui lui a déjà été faite dans le passé, ou, plus concrètement, de retourner dans un magasin qu'il a déjà visité. Dans le second cas, on choisit le plus souvent d'exclure le *perfect recall*: le demandeur d'emploi examine successivement des propositions qu'il doit accepter ou refuser de manière définitive. Cette distinction ne suffirait pas à fonder une taxinomie des modèles,²⁰ mais elle est certainement la plus utile pour l'application particulière que nous avons ici en vue.

Une recherche *interne* suppose en effet un retour en arrière toujours possible sur des actions ou des conséquences qui ont déjà été examinées. On pourrait vouloir modéliser le rôle de l'oubli – en considérant, par exemple, la possibilité du retour en arrière comme une variable aléatoire.²¹ Mais s'il faut choisir une hypothèse de première approximation parmi les deux hypothèses extrêmes, celle du *perfect recall* est manifestement la plus appropriée. Or, il se trouve justement qu'elle correspond au cas le plus facile à traiter mathématiquement et le mieux étudié jusqu'à présent (cf. Rosenfield et Shapiro, 1981). On pourra, en particulier, invoquer une classe de résultats qui supposent un modèle de Stigler généralisé et font tous appel à une hypothèse de *perfect recall*.

En guise de règle d'arrêt, Stigler avait proposé celle qui consiste à *fixer ex ante le nombre optimal de recherches à effectuer*. Il ne parvient pas à calculer ce nombre \hat{k} , mais il en démontre facilement l'existence et l'unicité, sous une hypothèse (prévisible) de coûts d'observation non-décroissants. \hat{k} est déterminé par la condition suivante, qui est l'analogie exact, pour une variable discrète ($k \in \mathbb{N}$), des égalités marginales de la théorie traditionnelle: si $k < \hat{k}$, l'agent a encore intérêt à recueillir une observation, si $k > \hat{k}$, il a eu tort de recueillir la dernière observation. L'inconvénient de cette règle saute aux yeux: supposons que l'agent rencontre le prix le plus bas auquel il prête une valeur de probabilité (ou de densité) positive, soit p_{min} ; il devra alors s'arrêter avant d'avoir effectué \hat{k} recherches. Il en ira de même, plus

généralement, si, en $k < \hat{k}$, il rencontre un prix inférieur à p_{\min} augmenté du coût de la recherche suivante.

Cette remarque a induit différentes reformulations²² qui, rétrospectivement, se comprennent mieux à la lumière de ce résultat général: toute règle faisant intervenir *ex ante* le nombre d'observations est inférieure (en utilité attendue) à une règle qui ferait intervenir pour seul critère un seuil de satisfaction r_t révisé au fur et à mesure des observations.

Voici l'énoncé précis du théorème (Kohn et Shavell, 1974, pp. 101–102). Par hypothèse, l'agent explore séquentiellement un certain ensemble X . Chaque observation est un tirage dans la loi de X ; les tirages sont indépendants. L'agent dispose d'une fonction d'utilité $V : X \rightarrow \mathbb{R}$, à laquelle on demandera seulement d'être bornée. A un instant t de sa recherche, l'information dont il dispose est décrite par une loi subjective F^t sur les suites d'observations futures $(x_{t+1}, x_{t+2}, \dots)$. Par ailleurs, l'agent a la possibilité de choisir l'une quelconque des observations précédentes ou contemporaines, $x_1 \dots x_t$: c'est l'hypothèse de *perfect recall*. Chaque observation a un coût $c(t)$ exprimé dans la même unité que V . Une règle d'arrêt optimale consistera alors à :

comparer le meilleur niveau d'utilité disponible en t , soit

$$(R') \quad y_t = \max \{V(x_1), \dots, V(x_{t-1}), V(x_t)\}, \text{ et un seuil} \\ r_t = K_t(y_t, C(t+1), F^t)$$

Il est instructif de comparer (R') et la forme générale d'une règle de *satisficing*, donnée par (R) (*supra*, p. 569). On voit que dans (R') , le seuil r_t est une fonction des utilités passées ou contemporaines $V(x_1), \dots, V(x_t)$, à travers la seule variable $y_t = \max \{V(x_1), \dots, V(x_t)\}$. Cette formulation n'est peut-être pas véritablement restrictive. En l'absence, néanmoins, d'une démonstration contraire, il faut considérer que (R') n'est qu'un cas particulier de (R) et donc relativiser l'importance du résultat démontré. Du point de vue de l'agent, (R') paraît bien comporter une perte d'information relativement à (R) .

Cette remarque permet de comprendre ce qu'il faut entendre exactement par "une règle d'arrêt optimale" dans l'énoncé précédent. Kohn et Shavell se sont, *a priori*, restreints à une certaine classe de stratégies de recherche: celles qui consistent, ou bien à s'arrêter, ou bien à continuer, *au vu d'une information résumée dans* $(y_t, C(t+1), F^t)$. C'est pour cette classe qu'ils démontrent l'optimalité de la règle consistant à fixer un seuil r_t ²³.

D'autres remarques préliminaires s'imposent pour que l'on puisse interpréter le théorème.

(1) Celui-ci est compatible avec toute espèce de loi subjective initiale et de mécanisme de révision. La littérature contemporaine sur le *search* incorpore quelquefois explicitement un tel mécanisme (p. ex. Rothschild, 1974). Mais la démonstration du théorème ne requiert aucune hypothèse particulière sur la nature de F^t . Le résultat affirme donc seulement que l'agent constate la valeur de x_t , révisé sa loi subjective F^{t-1} en F^t , révisé son seuil r_{t-1} en r_t (au vu de l'information x_t et de F^t), et s'arrête si et seulement si ce seuil est dépassé.

(2) Le théorème s'applique à toute fonction d'utilité bornée et il généralise donc le cas de Stigler, dans lequel l'utilité et les coûts étaient directement exprimés en monnaie (cas de neutralité par rapport au risque).

(3) Pas plus que le résultat de Stigler, celui de Kohn et Shavell n'est constructif. L'existence et l'unicité du seuil sont établies par un raisonnement de continuité sur une équation intégrale qui n'a pas nécessairement de solution explicite.²⁴

4.3. Résumé: l'apport de la théorie du "search"

Celle-ci se conforme également bien à une définition générique de la rationalité limitée (3.1) et à la caractérisation ordinaire par *search* et *satisficing* (3.2). Lorsqu'il traite des niveaux d'aspiration, Simon souligne toujours qu'ils s'ajustent aux performances réalisées; néanmoins, son formalisme de 1955 décrit seulement le cas particulier d'un niveau fixé, et l'on peut lui reprocher de n'avoir, par la suite, donné aucune indication déterminée sur la forme des processus de révision. Les modèles de *search* offrent une réponse partielle à cette question: l'agent révisé son seuil r_t au vu de la dernière réalisation et d'une loi de probabilité qui est elle-même révisée suivant un mécanisme approprié; une fois précisé le mécanisme de révision des probabilités,²⁵ la suite des seuils apparaît complètement endogénéisée (sinon calculable). On voit donc que le théorème de Kohn et Shavell peut donner lieu à une analyse entièrement "classique" des relations entre rationalité limitée et optimisation: ou bien les notions de recherche et de niveau d'aspiration en resteront au niveau générique ou intuitif, c'est-à-dire ne donneront *presque* aucune information scientifiquement utile, ou bien elles se spécifieront dans le cadre de la théorie optimisatrice.

Avant de présenter les objections les plus sérieuses qu'appelle ce point de vue, deux réserves s'imposent à l'attention. Nous avons souligné le fait qu'en l'absence d'une démonstration contraire, le résultat de Kohn et Shavell ne paraissait pas suffisamment général. Aucun théoricien du

search – à notre connaissance – n’a remarqué le trait commun à tous ces modèles: chacun présuppose une restriction *a priori* de l’ensemble des stratégies de recherche. Stigler se limitait aux stratégies qui consistent à fixer à l’avance un nombre d’observations et démontrait l’existence d’un \hat{k} optimal *pour cette classe de stratégies*. Kohn et Shavell reproduisent la même démarche; ils ont choisi un ensemble beaucoup plus vaste, mais qui, néanmoins, ne s’identifie pas *a priori* avec celui des stratégies possibles. Par ailleurs, il faut garder en mémoire le caractère non constructif des résultats. Ce trait est banal dans la théorie optimisatrice de la décision; mais il imposera manifestement d’en préciser l’intérêt *normatif*: quel sens y a-t-il à prescrire des stratégies dont on ne sait pas calculer la forme exacte?

5. Discussion critique des modèles du *search*

5.1. Les objections de Simon contre les modèles du “*search*”

Il nous semble que l’on peut regrouper ces objections sous trois rubriques: les modèles du *search* reposent sur une structure des coûts et des utilités qui contredit certains faits d’expérience; ils réclament un agent capable de mener à bien des calculs d’une complexité excessive; enfin, ils lui prêtent une information initiale démesurée. Ces reproches nous semblent pouvoir être gradués dans cet ordre. Le plus radical va porter sur l’hypothèse, commune à tous les modèles post-stiglériens, suivant laquelle l’agent connaît initialement ses coûts de recherche.

L’exemple de la recherche opérationnelle et celui du jeu d’échecs peuvent donner lieu à une objection locale, mais sans doute intéressante, contre la modélisation du *search*. Simon reconnaît que, dans certaines configurations de recherche, “nous nous trouvons contraints de faire un arbitrage (*tradeoff*) entre la qualité de la solution et le coût de calcul” (1978a, p. 496). C’est exactement la situation hypothétique dans laquelle se place la théorie du *search*. Mais le point important est que cet arbitrage peut fort bien rester “seulement implicite” (*ibid.*, p. 496).

Voici comment Simon fait ressortir ce point:

D’un côté, lorsqu’il est possible de trouver des algorithmes pour la solution optimale qui ne demandent qu’un temps de calcul raisonnable, la qualité de la solution disparaît de l’équation, et nous nous préoccupons seulement du coût du calcul. D’un autre côté, et à l’inverse, quand l’espace des possibilités à explorer est très vaste, et que nous sommes

incapables d'y découvrir une structure permettant une recherche efficace, les coûts n'interviennent pour ainsi dire pas dans l'équation: nous sommes alors contraints de rechercher n'importe quel algorithme qui produise des solutions acceptables au prix d'un effort de calcul acceptable (*ibid.*, p. 496).

En d'autres termes, *l'arbitrage stiglierien entre la qualité de la solution et le coût de calcul ne se produira pas en pratique.*

On peut amplifier quelque peu ces deux remarques. Dans le premier cas, l'agent doit choisir entre des algorithmes qui lui fournissent tous la solution optimale au degré d'approximation désiré: il ne peut donc les différencier que par les coûts de calcul – par exemple par les temps d'utilisation des machines – qu'ils lui imposent respectivement. Il est instructif d'identifier les algorithmes aux magasins visités par le consommateur de Stigler et de se ramener au modèle du *search*. Si l'on classe les algorithmes par coût absolu croissant, on obtient des coûts marginaux non négatifs, mais qui peuvent décrire une courbe quelconque, croissante ou décroissante suivant les intervalles (le coût marginal du premier algorithme est son coût absolu, celui des suivants est la différence de deux coûts absolus successifs). Quant aux gains marginaux, ils sont positifs pour le premier élément (c'est la valeur commune de la "qualité de la solution") et nuls ensuite. On se trouve manifestement dans une configuration dégénérée du modèle du *search*, où le choix se porte sur le premier élément examiné (cf. Figure 2a).

L'autre situation qu'envisage Simon fait apparemment intervenir une "qualité de solution" variable avec l'algorithme considéré, tandis que les coûts de calcul sont à la fois très élevés et rapidement croissants avec l'amélioration obtenue dans la "qualité de la solution": le jeu d'échecs pourrait ici servir à nouveau d'exemple. Si l'on classe les algorithmes par "qualité de solution" croissante, on obtient une courbe de gain marginal positive et de forme quelconque *a priori*, cependant que la courbe de coût marginal croît rapidement, à partir d'un niveau initial élevé. Là encore, le modèle de Stigler dégénère en une variante triviale: le *search* s'arrête sur l'un des premiers éléments examinés, peut-être même à nouveau sur le premier (cf. Figure 2b).

Il est très significatif que les deux cas de figure privilégiés par Simon correspondent à deux possibilités exclues par hypothèse implicite dans la théorie néo-classique du *search*: les prix sont tous identiques; le coût des premières recherches est tel que le consommateur achète toujours dans le premier magasin rencontré. Pour Simon, ces cas sont les plus fréquents en algorithmique, et la configuration aberrante est, précisément, celle que décrit Stigler: l'égalisation *progressive* des coûts et des gains

FIGURE 1
Le cas "général" de Stigler

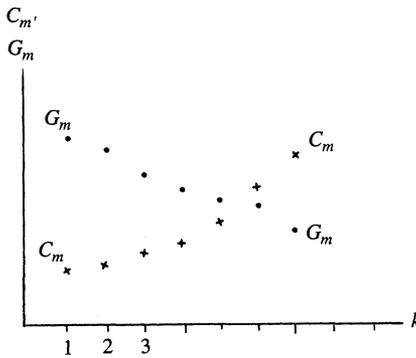
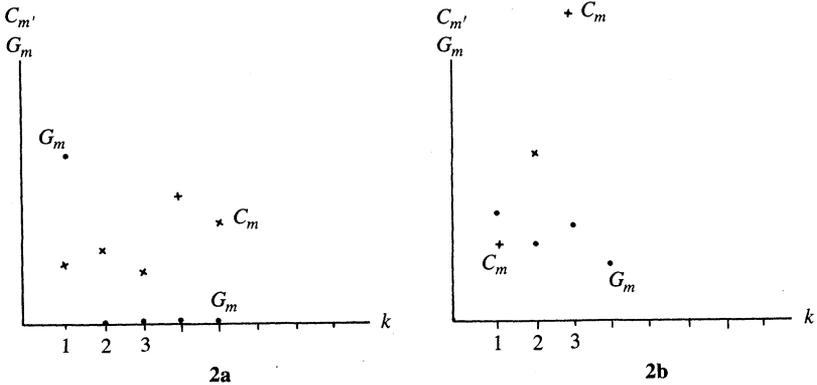


FIGURE 2
Le cas "général" selon Simon



Légende: k désigne le nombre d'observations, C_m et G_m désignent, respectivement, le coût marginal et le gain marginal de la recherche.

incrémentaux. Ainsi, la modélisation du *search* apparaît comme inadéquate à un secteur de la réalité tout au moins; et comme il s'agit, avec la recherche d'algorithmes, d'un cas patent d'incertitude interne et donc de rationalité limitée, le contre-exemple est véritablement embarrassant.

L'objection décisive, qui atteint la théorie dans toute son extension, concerne les exigences très lourdes qu'elle fait peser à la fois sur l'information de l'agent et sur sa capacité à la traiter. On sait que Simon a toujours objecté aux modèles optimisateurs la présence des *constraints on the information-processing capacities*. Dans leur visée explicite,

les travaux du *search* prennent ce fait pour point de départ. Mais, dit Simon, ils ne font que déplacer la difficulté et même l'aggraver. En déplaçant la décision de l'agent vers une optimisation du second ordre, ils accroissent à la fois l'information initiale qu'on lui prête et la complexité des calculs qu'il doit effectuer (1979a, p. 503, cité ici p. 561).

Une manière de préciser cette objection consisterait à comparer les axiomatiques sous-jacentes à la théorie de l'utilité espérée et à celle de la théorie ordinaire ("statique") de la maximisation de l'utilité.²⁶ Ce faisant, on traiterait jusqu'à un certain point du problème soulevé par Simon, puisque l'agent du *search* est (parmi d'autres caractéristiques) un maximisateur d'utilité espérée. Les deux axiomatiques posent l'existence d'un préordre complet sur les choix possibles, mais la notion de "choix" est entendue différemment dans les deux cas – ce sont des choix certains, généralement représentés par des vecteurs $x \in IR^n$, pour la théorie ordinaire, et des "loteries", ou combinaisons probabilisées des x précédents, pour la théorie de l'utilité espérée. On voit déjà, sur ce premier axiome, que celle-ci impose un surcoût d'information et d'aptitude au calcul par rapport à celle-là: les objets du choix qui sont préordonnés suivant la théorie ordinaire forment un *sous-ensemble strict* de l'ensemble correspondant pour la théorie de l'utilité espérée.

Recourir aux axiomatiques du choix présente un avantage subsidiaire: on découvre que la distinction, qui est à première vue naturelle, entre l'*information initiale* de l'agent et son *aptitude à calculer* les solutions compte tenu de cette information, est en réalité incertaine et variable suivant les modèles. Cette distinction est sans doute parfaitement claire dans le cas d'un modèle optimisateur: l'agent y dispose de certaines données – qui se ramènent à un ensemble des possibles et à une fonction d'évaluation; par ailleurs, on le suppose capable, au vu de ces données, de déterminer son optimum contraint. Mais si l'on recourt à un autre formalisme – qui peut d'ailleurs être logiquement équivalent au modèle optimisateur considéré – la distinction se brouille: que "connait" l'agent, que "calcule"-t-il exactement, d'après l'axiome suivant lequel "il existe un préordre complet sur l'ensemble des loteries"?

Cette remarque explique peut-être que Simon ne cherche pas à distinguer systématiquement les types de *difficultés cognitives* rencontrées par le décideur humain. Il emploie pour les désigner un grand nombre d'expressions qui paraissent interchangeables. Néanmoins, pour un modèle *donné*, et en particulier un modèle optimisateur, il peut y avoir du sens à distinguer ce qui est "information initiale" et ce qui est "aptitude à la traiter". C'est pourquoi on peut, sans risque de confusion, distinguer les objections de Simon contre les modèles optimisateurs

suivant qu'elles se rattachent à l'une ou l'autre de ces deux rubriques. Ce faisant, on dégagera un trait caractéristique, semble-t-il, de la critique simonienne: les objections les plus sérieuses *ne portent pas sur la capacité de l'agent à traiter optimalement l'information dont il dispose, mais plutôt sur l'information initiale démesurée que lui prêtent ces modèles.*

On a rappelé dans la section 2 que Simon n'était pas systématiquement hostile à l'optimisation. Il n'est pas déraisonnable d'imaginer que l'agent maximise sa fonction objectif lorsque la situation de choix est simple, c'est-à-dire que les ensembles A et S sont de petite taille et que la fonction d'évaluation V ne prend qu'un nombre limité de valeurs. On peut donc, face à un modèle optimisateur, se livrer à une expérience mentale: supposons que l'agent dispose effectivement des données nécessaires à son programme d'optimisation; le calcul de la solution fait-il encore problème dans ces conditions? Pour la théorie élémentaire de l'entreprise tout au moins, la réponse est facile à donner: "Si les courbes effectives de demande et de coût sont données, le calcul effectif de l'optimum est trivial" (1976, p. 137). C'est dire que la difficulté de la théorie élémentaire tient, presque exclusivement, dans l'information initiale qu'elle suppose. Qu'en est-il pour les modèles du *search*?

Simon dirait sans doute qu'ils imposent à l'agent des exigences de calcul qui ne sont nullement "triviales". Il est significatif que les principaux résultats de la théorie prennent la forme de théorèmes d'existence et que les auteurs des modèles ne soient généralement pas en mesure de calculer les seuils r_i . Ce fait – qui n'est pas exceptionnel, loin de là, dans la recherche néo-classique – soulève une question méthodologique embarrassante: est-il raisonnable de supposer, chez l'agent qu'on décrit, une aptitude à déterminer instantanément et sans coût des solutions que l'on ne sait pas soi-même calculer? Quel degré de cohérence faut-il supposer *entre les capacités que le modélisateur prête aux agents et l'état actuel de la science*? On devine que la réponse de Simon à ces questions ne serait guère favorable aux modèles du *search*.

L'objection décisive, néanmoins, concerne l'information initiale plutôt que les exigences de calcul au sens étroit. Elle est contenue dans une remarque que nous extrayons de l'article de 1955. Simon se demande ce qu'il adviendrait de son modèle si l'on supposait – hypothèse qu'il s'est gardé de faire jusqu'à ce point – que l'agent mesure directement ses coûts de recherche en termes d'utilité. Mais il ajoute aussitôt: "the central argument of the present paper is that the behaving organism does *not* in general know these costs" (1955, p. 254). Ainsi, la fallace inhérente aux modèles du *search* est qu'ils supposent donnée une

information du second ordre qui n'est pas disponible sous cette forme à l'agent: soit que la notion de "coût de recherche" apparaisse, du point de vue de l'agent, comme purement et simplement dénuée de sens; soit qu'il parvienne, dans une certaine mesure, à lui associer une quantité, mais que cette quantité soit incommensurable aux données du premier ordre (les utilités primaires).

Peut-on, ici encore, appliquer un principe de cohérence intellectuelle entre ce que fait le modélisateur et ce qu'il décrit chez l'agent du modèle? Les recherches en algorithmique visent à formaliser la notion de complexité.²⁷ Néanmoins, elles ne semblent pas devoir nécessairement conduire à l'élaboration d'un concept *quantifié* de coût de recherche. Dans un article qui n'est pas très ancien, Simon rappelait que l'on en est souvent réduit, pour les algorithmes de programmation linéaire, à cette méthode primitive de comparaison des coûts: on les fait opérer sur des problèmes choisis et l'on relève les temps de calcul qui sont nécessaires à l'ordinateur dans chaque cas (1978a, p. 496).

Les objections de Simon contre le *search*, que nous venons de rappeler ou de reconstituer, soulèvent toutes un problème d'interprétation. On peut y voir tout d'abord des objections *empiriques* du type le plus immédiat: c'est un fait que les individus ne peuvent faire de calculs compliqués, etc. On peut aussi les présenter comme des objections de type *méthodologique*. Une approche qui rejette certains paramètres fondamentaux de la décision dans l'information initiale de l'agent, c'est-à-dire dans l'exogène, se condamne à ne rien expliquer véritablement: elle est triviale. Enfin, on peut faire apparaître, à propos de l'hypothèse stiglérienne d'un coût de recherche connu de l'agent, la menace d'une régression à l'infini: on débouche alors sur une objection *logique*, qui, semble-t-il, n'a jamais été envisagée en termes exprès par Simon, mais peut constituer un développement naturel de sa position.

5.2. De l'objection empirique à l'objection méthodologique

Invoquer directement la théorie empirique de la décision contre les approches optimisatrices revient à faire une postulation d'empirisme. Par exemple, l'axiomatique de l'utilité espérée, dans sa version von Neuman-Morgenstern, comporte un principe, dit des "loteries en plusieurs étapes", suivant lequel le décideur sait réduire une loterie complexe en ses constituants ultimes.²⁸ Ce principe demande donc à l'agent de savoir manier la notion de probabilité composée – à tous les niveaux de complexité envisageables. Or les travaux de Kahneman

et Tversky (1974) ont mis en évidence bien des anomalies dans l'usage que les décideurs empiriques font des probabilités: on constate notamment qu'ils sous-estiment la probabilité des événements composés par disjonction. Ce dernier cas correspond justement à celui de l'axiomatique von Neuman-Morgenstern. Il n'est donc pas douteux que le principe des "loteries composées" tombe – ainsi, d'ailleurs, que la plupart des axiomes de l'utilité espérée – sous des objections empiriques très fortes. Mais le vrai problème est de savoir *si ces objections sont recevables telles quelles*: on ne peut l'affirmer sans postuler une interprétation radicalement empiriste de la théorie de l'utilité espérée.

C'est dire que la critique empirique de l'optimisation débouche presque immédiatement sur la critique méthodologique. Il y a deux manières, nous semble-t-il, de mener celle-ci à bien. La première consiste à prendre parti sur la question difficile de l'*irréalisme* des prémisses en économie et dans la théorie de la décision. La seconde, moins extérieure sans doute, choisit comme point de départ la théorie optimisatrice de la décision telle qu'elle se donne, et non pas telle qu'elle doit être. Il s'agira d'accepter les termes exprès de son programme de recherche et de lui appliquer les critères ordinaires de *progressivité* (en cohérence interne, d'une part, et en contenu testable ou en contenu testé, d'autre part) que la philosophie des sciences élabore abstraitement.

Il est bien évident que la critique simonienne de l'optimisation se fonde, au moins implicitement, sur une récusation absolue de la thèse friedmanienne: "a theory cannot be tested by the 'realism' of its 'assumptions'" (Friedman, 1953, p. 23). On ne doit pas s'étonner que Simon lui-même soit intervenu dans les débats suscités par cette thèse. Il l'a fait avec une brièveté et une discrétion remarquables: celles-ci témoignent, à la fois, de sa méfiance bien connue à l'égard des controverses *purement* méthodologiques, et du peu de considération qu'il a pour la thèse de Friedman (Simon, 1963, pp. 229–231). Il est dommage, néanmoins, que Simon n'ait pas cherché à relier plus nettement les positions fondamentales de sa philosophie des sciences et sa critique de l'"irréalisme". On sait que l'épistémologie simonienne incline vers un empirisme sophistiqué: elle a réhabilité et, dans une assez large mesure, mené à bien le projet néo-positiviste qui consiste à *réduire* la place des termes théoriques en science.²⁹ On peut se demander si cette position en philosophie pure commande logiquement la récusation des thèses de Friedman: la réponse ne va pas de soi *a priori*. On peut aussi, ce qui est sans doute plus important, se demander si la "réduction" simonienne des termes théoriques amène, dans le cas particulier de la théorie de la décision, à cautionner les objections naïvement

empiristes comme celle que nous avons citée en exemple. La réponse, ici encore, est loin d'aller de soi.

Nous n'avons pas voulu ici aborder le détail de ces questions, car il nous semble possible d'éviter le détour assez complexe qu'elles nous imposeraient de faire. En effet, les modèles de *search* se prêtent à une critique méthodologique beaucoup plus directe et facile à mener à bien. Nous ne ferons qu'en indiquer le principe général. Celui-ci consiste à admettre *ex hypothesi* la restriction implicite dans la théorie optimisatrice de la décision: on ne s'intéressera pas à l'analyse du *processus* de décision, mais seulement de ses *résultats*, individuels ou collectifs. S'agissant, par exemple, de l'utilité espérée, le problème du test se spécifie alors de la manière suivante: on considérera des ensembles d'objets sur lesquels les préférences empiriques des agents sont manifestes, et l'on se demandera si les ensembles ainsi ordonnés sont compatibles avec la théorie. Par exemple, on interrogera directement un agent pour savoir dans quel ordre de préférence il classe certaines "loteries". Ou bien, avec Friedman et Savage (1948; 1952), on cherchera à expliquer le fait reconnu que certains agents achètent des polices d'assurance, alors que, par ailleurs, ils vont jouer au casino. Ces différents tests ont en commun une notion réductrice du choix humain: celui-ci est ramené à son point d'impact sur le monde extérieur (la "loterie", l'acquisition d'une police d'assurance, etc.). On disqualifie, en revanche, comme non pertinents, les tests qui font intervenir la structuration interne du choix. Du coup, tous les contre-exemples fondés sur l'information initiale de l'agent, ou sa capacité à calculer les solutions, perdent leur fonction réfutatrice. Il en va ainsi, notamment, de l'objection examinée sur les "loteries en plusieurs étapes".

Une fois admise cette restriction *a priori* des "falsificateurs potentiels", il reste à voir si la théorie se développe de manière "progressive", suivant le terme employé par Lakatos (1970). Il en ira ainsi lorsque la théorie progresse en contenu testable et, dans une certaine mesure aussi, en contenu testé; cette condition est nécessaire, mais non pas suffisante, car il faut encore que le gain de contenu ne soit pas obtenu de manière *arbitraire*. On réclamera donc aussi que la théorie ait progressé d'une manière qui soit compatible avec son heuristique fondamentale. Dans la pratique, la question de la progressivité de l'approche optimisatrice – ou de toute autre théorie – peut se traiter en étudiant la façon dont elle se réincorpore des anomalies, ou bien surmonte des expériences réfutatrices.

C'est dans cette perspective méthodologique que nous reviendrons brièvement sur l'émergence des modèles de *search*: le problème est alors de savoir dans quelle mesure ces modèles ont fait *progresser* l'approche

optimisatrice de la décision. Les termes de référence (par rapport auxquels on appréciera la “progressivité” ou la “dégénérescence” éventuelle de l’approche) sont, d’une part, la théorie élémentaire, dite “statique”, du consommateur, d’autre part la théorie de l’utilité espérée.

Quand Stigler écrit son article fondamental, “The Economics of Information” (1961), il a précisément en vue certaines anomalies embarrassantes pour la théorie élémentaire:

- on constate une dispersion élevée des prix de vente, non seulement pour les biens de qualité différente, mais aussi pour des biens relativement homogènes;

- les transactions effectuées sur une marchandise homogène et divisible font souvent apparaître un aspect de *higgling*, c’est-à-dire une discrimination positive ou négative du vendeur à l’égard de ses acheteurs.

Le modèle ébauché dans l’article met en place les hypothèses nécessaires à la réduction de ces anomalies: les acheteurs prennent leur décision au vu d’une loi de probabilité (généralement subjective) sur les prix de vente, et de certains coûts de recherche. Le caractère progressif ou non de l’explication va maintenant dépendre de ces deux critères: est-elle conforme à l’heuristique générale de la théorie optimisatrice? fait-elle progresser celle-ci en contenu testable et en contenu testé?

Il n’y a guère de doute que Stigler a transposé, autant que faire se pouvait, le principe général de l’optimisation élémentaire. Son modèle revient simplement à rajouter des contraintes au programme du consommateur ordinaire. Il ne hiérarchise pas la contrainte cognitive, exprimée dans le coût de recherche, et les contraintes technologiques ou financières seules prises en compte dans la micro-économie naïve. Plus profondément, la démarche de Stigler ne se contente pas de reproduire un mouvement qui préexisterait dans un état antérieur de la science: la fidélité à l’heuristique signifie *aussi* que la variante la plus tardive révèle complètement des aspects de cette heuristique qui étaient seulement virtuels ou implicites. Or voici l’une des constantes fondamentales de la théorie optimisatrice que Stigler nous permet de comprendre plus nettement: *à tout phénomène de rareté, la théorie croit pouvoir associer une estimation numérique*, qu’il s’agisse d’un prix de marché ou d’un “coût d’opportunité” non directement observable. Ainsi, l’agent stiglérien manque de certaines informations: le modèle comportera *donc* des coûts de recherche correspondant à chaque information manquante.³⁰

Bref, la fidélité de Stigler à l’heuristique traditionnelle ne fait guère problème. Il reste à considérer les gains éventuels de l’approche en contenu testable et testé. Par construction même, les *explicanda* qui

correspondent aux deux anomalies (dispersion empirique des prix, phénomènes de *higgling*) sont réinsérés dans l'approche optimisatrice. Il faut alors vérifier que l'*explicans* n'est pas "ad hoc", c'est-à-dire qu'il sert aussi à prédire d'autres *explicanda*, si possible éloignés des anomalies initiales. Stigler est parfaitement conscient de cette exigence – classique – de testabilité indépendante. Il argumente ainsi en faveur de ses hypothèses: "La publicité... permet aux acheteurs éventuels d'identifier leurs vendeurs" (1961, p. 220). "La publicité sur les prix dispense les acheteurs éventuels d'un effort de recherche considérable" (*ibid.*, p. 224). Bref, les deux hypothèses du modèle (l'acheteur forme une loi de probabilité subjective sur les prix, il est soumis à des coûts de recherche positifs) apparaissent cohérentes avec l'existence de la publicité, et, plus généralement, d'institutions diffusant l'information disponible sur les prix de vente. Le consommateur accepte de financer de telles institutions – éventuellement sous la forme d'un surcoût grevant le produit acheté – dès lors qu'elles lui permettent de supporter, par ailleurs, des coûts de recherche plus faibles.

La dérivation de ce nouvel *explicandum* rencontre néanmoins deux difficultés. Stigler traite fort honnêtement l'une d'entre elles: c'est généralement le vendeur, et non pas l'acheteur, qui finance une opération de publicité (*ibid.*, p. 221). On pourrait aussi objecter que la publicité ne consiste pas seulement, ni même sans doute principalement, dans l'affichage des prix. La différenciation du produit semble être l'une des visées constantes des messages publicitaires, ce qui va directement à l'encontre du but envisagé par Stigler – informer le consommateur sur le prix d'un bien *homogène*. Quoiqu'il en soit de ce point, la vraie faiblesse du modèle n'est pas dans le choix, finalement assez ingénieux, d'un *explicandum* de contrôle. Elle réside dans l'absence de toute prévision quantitative: du modèle de *search*, on ne peut tirer la forme exacte de la courbe de demande à l'entreprise qui vend au prix p ; on ne peut, naturellement, faire de prévision numérique sur les montants de publicité engagés en fonction des différents biens. La raison en est évidente: Stigler a généralisé le modèle élémentaire en introduisant deux ensembles de paramètres nouveaux, les coûts de recherche c_k et les lois subjectives $F^k(p)$. Or les données dont il dispose ne lui permettent d'assigner que des contraintes très vagues sur ces paramètres nouveaux.

La donnée principale tient dans cette constatation: l'hypothèse d'une "recherche" par le consommateur doit rester compatible avec le maintien (empiriquement constaté) d'une dispersion non nulle des prix; il faut donc faire appel à des c_k et $F^k(p)$ tels qu'il ne soit pas avantageux pour le consommateur et le vendeur d'éliminer toute dispersion (cf. Stigler, 1961, pp. 219–220). De cette constatation, on peut vraisemblablement inférer la

forme croissante des coûts et celle, décroissante, du gain marginal (pour justifier le fait que la recherche du consommateur s'arrête bien avant l'identification du meilleur prix). De là, une double restriction, sur les c_k et sur les $F^k(p)$. Mais elle est évidemment très faible. On ne saurait s'en étonner véritablement, puisqu'elle découle de l'*explicandum* principal (l'existence d'une dispersion empirique des prix), et non pas d'une inférence supplémentaire. La considération d'un *explicandum* de contrôle (la publicité) n'a permis d'ajouter aucune restriction nouvelle sur les deux ensembles de paramètres.

On nous objectera sans doute que nous réclamons trop des modèles de *search*. Le problème n'est pas d'apprécier leur contenu testable et testé dans l'absolu, mais seulement de les comparer à leurs ancêtres immédiats; or les modèles de la théorie micro-économique élémentaire ne débouchent pas non plus immédiatement sur des prévisions numériques de comportement, puisque la fonction d'utilité "statique" n'est pas directement observable. Cette objection doit nous conduire à reformuler la critique précédente dans une perspective comparative et historique. On sait que les modèles élémentaires permettent, au moins sous certaines conditions, de dériver la forme décroissante de la courbe de demande à l'entreprise.³¹ Or, l'un des résultats les plus curieux de la théorie du *search* est que, dans les mêmes conditions exactement, elle ne permet pas de dériver une forme croissante ou décroissante de la demande à l'entreprise. On s'en convaincra par un exemple très simple, que nous empruntons à Kohn et Shavell (1974). Supposons qu'après avoir effectué la $k^{\text{ème}}$ observation, le consommateur pense que deux lois de probabilités seulement sont possibles:

$$\begin{array}{l} 1) \quad \text{Proba } (p_{k+1} = 3) = 1 \\ 2) \quad \left\{ \begin{array}{l} \text{Proba } (p_{k+1} = 1) = 0,99 \\ \text{Proba } (p_k = 2) = 0,01 \end{array} \right. \end{array}$$

Si le consommateur observe $p_{k+1} = 2$ et que le coût c_{k+2} de l'observation suivante ne soit pas trop élevé, il doit poursuivre sa recherche. Si, en revanche, il observe $p_{k+1} = 3$, il doit s'arrêter, quelle que soit la valeur de c_{k+2} . Il en résulte que le consommateur adresse au $(k+1)^{\text{ème}}$ magasin une demande qui n'est pas décroissante avec le prix!

Le modèle "statique" n'est pas la seule référence qui s'impose: nous devons aussi considérer la théorie de l'utilité espérée, puisque l'on peut considérer tout modèle de *search* comme l'un de ses développements

possibles. On sait que l'axiomatique de l'utilité espérée peut donner lieu non seulement à des tests empiriques très divers, mais encore à un procédé de construction particulièrement simple de la fonction d'utilité "von Neuman-Morgenstern". Pour chaque agent qui répondrait, sans violer les axiomes, à un questionnaire approprié, on arrive à *estimer une fonction d'utilité cardinale*, définie à une transformation linéaire près (p. ex. Friedman et Savage, 1948; 1952). Ainsi, un modèle qui prête aux agents une fonction d'utilité "von Neuman-Morgenstern" peut conduire non seulement à des résultats qualitatifs, mais encore à certaines prévisions quantitatives – si, du moins, on prend la peine d'estimer comme il convient les fonctions. L'insertion dans un tel modèle d'un coût de recherche non observable ferait disparaître cette possibilité et signifierait donc, sans équivoque, une perte de contenu informatif. Afin de parer à cette objection, la théorie devrait s'imposer de *révéler directement les coûts de recherche*, suivant le principe qui a permis naguère d'opérationnaliser la théorie de l'utilité espérée. Certaines études de marketing ont paru s'engager dans cette direction, mais l'approche néo-classique s'en est, jusqu'à présent, tenue éloignée. Elle continue à traiter les coûts de recherche comme des termes théoriques, soumis à des contraintes qualitatives très vagues, qui, dans le meilleur des cas, ne peuvent conduire qu'à des prédictions elles-mêmes qualitatives.

On pourrait maintenant franchir un nouveau pas et s'interroger sur la "progressivité" de l'approche optimisatrice lorsqu'elle transpose le modèle du *search* au problème spécifique de la rationalité limitée. Il est très probable que, lorsqu'elle porte sur une recherche *interne*, comme celle du joueur d'échecs, la théorie perd certaines possibilités de détermination indépendante des coûts de recherche qui existaient dans le cas du consommateur. Cette difficulté supplémentaire demanderait à être précisée soigneusement. Mais on a sans doute compris, maintenant, quel est le principe général de notre critique méthodologique. Dans sa *Logik der Forschung* (§ 36), Popper distingue deux manières dont une théorie peut gagner en contenu testable. Comme elle a une forme conditionnelle "si A, alors B", elle pourra progresser ou bien par extension de son antécédent A, ou bien par restriction de son conséquent B. Popper appelle le premier cas un gain en *universalité*, le second un gain en *précision* (1972, pp. 121–122). La discussion de l'apport de Stigler donne à penser que la théorie optimisatrice de la décision a progressé en universalité, mais perdu en précision. Il nous semble que l'étude détaillée d'un modèle de *search* "interne" conduirait à reproduire cette conclusion. *Pour rendre compte de phénomènes qualitatifs nouveaux, la théorie perd certaines des restrictions mesurables qu'elle imposait à ses prévisions.*

En résumé, la critique dont nous venons d'indiquer le principe consiste en une application rudimentaire de la méthodologie des programmes de recherche.³² Nous n'avons pas voulu isoler notre objet principal – les modèles de *search* – et le contexte scientifique de son apparition: au contraire, nous l'avons situé dans la continuité d'une évolution (dont le point de départ serait la micro-économie "statique" élémentaire). Nous n'avons pas cherché à discuter dans l'absolu, mais uniquement de façon relative – c'est-à-dire en termes de "progressivité" – les contenus testables et testés des modèles du *search*. Cette méthode se recommande par de nombreuses considérations qui ne peuvent être rappelées ici, mais on voit qu'elle présente cet avantage immédiat: on peut traiter de manière précise la question de la testabilité sans avoir, pour autant, à définir une unité de mesure pour cette notion; ce sont les états antérieurs de la théorie qui servent de point de référence. Nous pensons que l'objection empirique sophistiquée, ou l'objection méthodologique, qu'appellent les modèles du *search*, gagne à être reformulée dans la conceptualisation poppéro-lakatosienne des programmes de recherche.

5.3. Esquisse d'une objection logique

Lorsqu'il incrimine l'information initiale requise dans le modèle de Stigler, Simon semble tout près de donner une forme logique à ses objections. La théorie, dit-il en substance (voir note 18), déplace l'optimisation vers le second ordre en transformant la décision primaire d'allocation – acheter à un certain prix – en une décision d'allocation de l'information – s'arrêter de chercher; la faiblesse dirimante de cette théorie vient de ce que les données nécessaires à l'optimisation du second ordre ne sont pas réellement disponibles. Il suffirait d'ajouter à cette analyse: "donc l'agent doit les rechercher à leur tour, ce qui déplace l'optimisation vers le troisième ordre", pour qu'apparaisse la difficulté logique inhérente, nous semble-t-il à tous ces modèles: *ils entraînent l'agent dans une régression à l'infini*.

C'est un trait à nos yeux fort surprenant de la conception simonienne de la rationalité que *jamais* elle ne débouche explicitement sur cette forme de la critique. Mieux, il semble que Herbert Simon ait voulu condamner par avance toute formulation *a priori* de l'échec de l'optimisation:

For most problems that Man encounters in the real world, no procedure that he can carry out with his information-processing equipment will enable him to discover the optimal solution, even when the notion of 'optimum' is well defined. *There is no logical reason*

why this need be so; it is simply a rather obvious empirical fact about the world we live in – a fact about the relation between the enormous complexity of that world and the modest information-processing capabilities with which Man is endowed (1976, p. 135, italiques de l'auteur).

Il faudra revenir sur cette thèse, que nous ne pouvons accepter sous la forme où elle est exprimée ici: nous nous proposons en effet, dans cette section, d'examiner une "raison logique" qui est de nature à empêcher le décideur de "découvrir la solution optimale". Ce faisant, nous suivons la voie ouverte par un économiste fortement inspiré des thèses simoniennes, Sidney Winter. Celui-ci n'invoque pas la *régression infinie* des modèles optimisateurs de la décision, mais se demande – question étroitement liée à la précédente, comme on va le voir – si l'on peut parler, sans contradiction logique, d'une "*optimisation tenant compte de ses propres coûts*" (1975, p. 83). Au demeurant, Winter ne fait que signaler l'existence d'une difficulté sans prétendre la traiter complètement, faute d'un "cadre formel" adéquat.³³

Un modèle de *search* comporte toujours deux types de données spécifiques, la suite des coûts de recherche et une loi de probabilité sur les éléments du choix, qui représentent un surcroît d'information par rapport à ce qu'un modèle de la micro-économie élémentaire réclamerait de l'agent. Comme ces données ne sont disponibles nulle part, il est raisonnable de penser qu'elles sont à leur tour recherchées, et que la théorie classique de la décision doit être en mesure de décrire leur acquisition dans des termes optimisateurs. Il faudrait donc montrer comment l'agent choisit optimalement une loi subjective parmi toutes celles qui sont initialement à sa disposition, et comment il arrête, de nouveau optimalement, une structure de coûts de recherche parmi toutes celles qui se présentent à lui. Le *search* du second ordre, c'est-à-dire l'optimisation du troisième ordre, consistera à doter les deux ensembles pertinents – celui des lois de probabilité et celui des coûts – d'une mesure de probabilité, et à décrire un ajustement marginal qui endogénéise le choix d'une certaine $F^0(p)$ et d'une certaine suite $C_1, C_2, \dots, C_k, \dots$.

Dans un travail antérieur (Mongin, 1985), nous nous sommes penché sur une question plus simple, mais formellement similaire. Considérons une optimisation élémentaire, c'est-à-dire qui ne tienne compte ni des coûts de calcul de la solution, à paramètres donnés, ni des coûts de recherche des paramètres: elle définit le premier niveau logique de la décision. Lorsque ces "coûts simoniens" sont réintégré au calcul de l'optimum, on dispose d'un modèle nouveau, qui se place à un niveau logique immédiatement supérieur. On peut regarder les modèles du

search comme des cas particuliers de cette optimisation du second ordre. Or l'une des leçons prévisibles, mais importantes, des modèles du *search* est qu'ils décrivent un agent moins bien loti que celui de la micro-économie élémentaire: l'agent stiglierien n'achète pas, en règle générale, au plus bas prix qui se rencontre sur le marché. Nous avons suggéré que cette perte d'optimalité dépend du caractère séquentiel de la recherche plutôt que de la manière, optimisatrice ou non, d'arrêter la recherche (cf. *supra*, 3.3). Quoiqu'il en soit, elle affecte tous les modèles de *search*, et révèle qu'il est à tout le moins *possible* que l'optimum second ne coïncide pas avec l'optimum premier. Cette *éventualité* suffit à rendre la théorie optimisatrice aussi fragile qu'un jeu de cartes. Supposons en effet que l'on veuille tenir compte des "coûts simoniens" apparus dans le calcul de l'optimum second. On construira un modèle du troisième niveau qui, peut-être, déclassera l'optimum second. La remontée des niveaux logiques successifs peut fort bien interdire à l'agent de "stabiliser" une solution optimale.

Dans le cas bienheureux où il n'en va pas ainsi, on dira que la régression infinie *converge en décision*. Formellement parlant, on définira la convergence en décision par l'existence d'un niveau fini N tel que, pour tous n et n' vérifiant $n' > n \geq N$, la décision optimale du niveau n se trouve confortée par les décisions optimales des niveaux $n+1$, $n+2$, ..., n' . D'un point de vue pragmatique, il vaudrait mieux pour la théorie optimisatrice que N soit petit; mais du point de vue logique, cette propriété est évidemment dénuée d'importance. L'existence de N est l'une des conditions requises pour qu'une théorie de la décision opérant à des niveaux *finis* simule l'effet d'une théorie plus riche qui – à supposer que cela ait un sens humain – parviendrait à tenir compte d'une *infinité dénombrable* de niveaux logiques. Nous allons voir que cette condition est nécessaire, mais n'est pas suffisante.

Supposons en effet qu'il existe un niveau "stable" N au sens de la convergence en décision. Il peut très bien se faire que les optimisations de niveau $N+1, \dots, N+j, \dots$, qui, par hypothèse, confortent l'optimisation en N , imposent à l'agent des coûts simoniens *croissants*, $K_{N+1} \leq K_{N+2} \leq \dots \leq K_{N+j} \leq \dots$. Cela pourrait être le cas, notamment, si l'algorithme d'optimisation implique une énumération exhaustive de l'ensemble des possibilités, et si la taille des ensembles à énumérer croît avec le niveau logique. Dans une telle situation, la théorie optimisatrice n'aurait remporté qu'une victoire à la Pyrrhus: la stabilité en décision du niveau N se paierait d'une perte d'utilité égale à $\sum_{j \geq 1} K_{N+j} = +\infty!$ La théorie optimisatrice n'éviterait une première absurdité – celle qui

consiste, dans le cours de la régression, à déclasser sans cesse les décisions du niveau inférieur – qu'en tombant dans une deuxième absurdité – consistant à laisser décroître sans cesse l'utilité de l'agent. Pour tenir compte de cette difficulté, Mongin (1985) introduit la notion de *convergence en utilité*, qui définit une nouvelle contrainte sur la théorie optimisatrice; celle-ci ne sera finalement sauvée que si elle parvient à réunir les deux critères de convergence. La définition formelle de la convergence en utilité est plus lourde à manier que celle de la convergence en décision, et nous devons ici nous contenter d'une indication approximative. *Grosso modo*, une régression infinie résultant de l'optimisation converge en utilité s'il existe au niveau fini N' tel que, pour tous n et n' vérifiant $n' > n \geq N$, la différence entre les coûts simoniens comptabilisés en n' et les coûts simoniens comptabilisés en n soit inférieure à un montant ε prédéterminé.

On peut remarquer que le niveau $N'' = \sup \{N, N'\}$ fournit une interprétation possible de ce que Winter appelle "l'optimisation qui tient compte de ses propres coûts". En effet, pour l'observateur, sinon pour l'agent, N'' ne se distingue pas de $N'' + 1$, $N'' + 2$, ... puisque l'optimisation en $N'' + 1$ reconduit les valeurs de N'' , celle de $N'' + 2$ reconduit les valeurs de $N'' + 1$, puis celle de N'' , et ainsi de suite. En outre, du point de vue de l'agent, il est presque équivalent de demeurer en N'' ou de remonter à un niveau quelconque $N'' + j$, $j \geq 1$, puisque la différence d'utilité est "négligeable" – elle est majorée par ε prédéterminé. Il en résulte que le niveau N'' représente tous les niveaux supérieurs pour la théorie de l'optimisation. Celle-ci a de bonnes raisons de ne pas distinguer entre N'' et $N'' + j$. Tout se passe comme si le niveau N'' avait une propriété de réflexivité – comme s'il tenait compte lui-même de considérations qui, normalement, doivent s'intégrer à des niveaux logiques autres, $N'' + 1$, $N'' + 2$, etc. Cette remarque ne fait, à vrai dire, qu'illustrer les liens, bien connus des logiciens, entre les phénomènes d'autoréférence et ceux de régression à l'infini. Très grossièrement, lorsqu'une régression à l'infini n'est pas "vicieuse", on peut la traduire, d'une manière également non "vicieuse", dans le langage de l'autoréférence – et réciproquement.

En l'absence d'hypothèses supplémentaires, les deux notions de convergence qui viennent d'être présentées sont logiquement indépendantes. On ne peut exclure que l'optimisation ne parvienne pas à les vérifier simultanément, de sorte qu'il y aurait lieu de prévoir, à côté de la résolution favorable du problème, une issue intermédiaire – l'un des critères serait satisfait et non pas l'autre. Il résulte finalement de cette discussion informelle que les théoriciens classiques disposent *a priori* de trois méthodes pour affronter l'objection logique: ou bien ils la nient purement et simplement en arrêtant

la régression à un niveau qui ne présente aucune des propriétés désirables; ou bien, il se satisfait de l'une des deux formes de convergence; ou bien, ils parviennent à triompher de l'objection en démontrant la double convergence.

La première méthode (α) consiste donc à arrêter d'office la remontée des niveaux logiques à une valeur ν *a priori* quelconque. Une telle attitude peut se justifier ainsi: la théorie optimisatrice *ne doit pas* s'appliquer aux niveaux $n > \nu$; il n'y a pas de sens à la faire servir au-delà de cette valeur. S'il ne veut pas tomber dans la pétition de principe pure et simple, le théoricien optimisateur doit justifier cet interdit par des raisons de type théorique ou empirique. Certains auteurs – économistes, mais aussi, peut-être spécialistes d'intelligence artificielle – diraient volontiers que l'existence d'un niveau optimisateur maximal ν est un fait de nature: l'homme ne sait pas calculer au-delà. Un tel recours à l'empirie comporte un danger évident: si l'on consulte les données psychologiques, elles risquent fort d'être défavorables non seulement à l'optimisation au niveau $\nu + 1$, mais à l'optimisation au niveau 1! Une autre démarche, plus subtile, fait intervenir des justifications théoriques extrinsèques à la théorie optimisatrice de la décision – par exemple des déterminants biologiques ou quasi-biologiques. On sait qu'il existe une "théorie" de la "sélection naturelle économique"³⁴: celle-ci prétend dériver, à partir des contraintes de la concurrence parfaite, les dispositions qu'auraient les entreprises représentatives à optimiser *au sens élémentaire*. Il serait ingénieux de transposer l'idée à des ordres logiques supérieurs: on pourrait imaginer une théorie sélectionniste qui prédirait la survie à long terme des seuls agents *capables d'optimiser au niveau ν* . Quoiqu'il en soit de ce projet, ou de toute autre justification théorique extrinsèque du choix de ν , il est clair que l'on ne peut rationnellement s'en remettre qu'à des théories constituées et des résultats effectifs. Le théoricien optimisateur ne doit pas de contenter d'invoquer la simple *possibilité* d'une explication biologique de l'optimisation au niveau ν , il doit l'articuler effectivement ou bien s'en passer, si, d'aventure, il n'y parvient pas.³⁵ C'est pourquoi, dans l'état actuel du savoir, il devra se ramener à l'une ou l'autre des propositions qui suivent.

Puisque la régression infinie est inévitable, on peut vouloir en récuser le caractère "vicieux" en affirmant soit (β) la réalisation de l'un des critères de convergence pris isolément, soit (γ) la résolution complète du problème par la réalisation simultanée des deux critères. La mathématique sous-jacente à ces variantes est loin d'être élucidée à ce jour. Il est évident que si la suite des "coûts simoniens" ne comporte plus que des termes nuls à partir d'un certain rang, la double convergence est assurée. De même, c'est une condition nécessaire évidente de la double

convergence (ou de la convergence en utilité) que la suite des “coûts simoniens” soit décroissante. Ce point avait été d’ailleurs suggéré par Winter (1975, pp. 83–84). Pour énoncer des conditions plus subtiles, soit nécessaires, soit suffisantes, il faudrait développer un formalisme complet et expliciter les hypothèses faites sur les coûts propres aux différentes règles. Nous avons entrepris de le faire dans un travail qui, au stade actuel, débouche sur des conclusions nuancées: la solution (γ) n’est ni toujours applicable, ni toujours exclue – tout dépend de la structure de l’environnement et des propriétés particulières de la suite des “coûts simoniens”. Le principe de cette réponse ne saurait évidemment surprendre: si l’on parvient à invoquer (β) ou (γ) contre l’objection logique, ce ne peut être qu’en s’appuyant sur des constatations d’ordre *empirique*. Dans cette perspective favorable, le théoricien optimisateur reconnaîtrait donc que la régression infinie est inévitable, mais il affirmerait aussitôt que les conditions de l’optimisation humaine sont de nature à la rendre bénigne. Si telle est bien la structure de l’argumentation orthodoxe, on s’expliquerait que Simon ait refusé de donner une forme purement *logique* au débat sur la théorie optimisatrice. Il est probable que la plausibilité des réponses (β) ou (γ) se jugera sur la même classe des données que celle que mobilisaient les deux premières strates d’objections, empiriques et méthodologiques.

A supposer que l’on parvienne à identifier une configuration de “coûts simoniens” qui serait plus vraisemblable qu’une autre, une difficulté subsisterait: l’observateur seul, et non l’agent, sait que les “coûts simoniens” – pour prendre l’exemple le plus simple – s’annulent à partir de N . L’agent ne dispose évidemment pas de la suite infinie des “coûts simoniens”, il la découvre peu à peu: *après* avoir optimisé de manière “élémentaire”, il perçoit les coûts négligés et décide d’en tenir compte au second ordre; *après* avoir optimisé au second ordre... L’agent peut s’arrêter à un moment quelconque de cette vertigineuse escalade, mais il n’est jamais sûr d’avoir atteint le sommet: même si, par extraordinaire, il découvrirait que les coûts sont nuls au niveau N , il ne pourrait rien savoir des niveaux $N + 1$, $N + 2$,... tant qu’il ne les aurait pas encore parcourus. Ainsi, les réponses (β) et (γ) peuvent sauver l’optimisation du point de vue d’un observateur omniscient, mais elles ne garantissent pas que l’optimisation soit possible du point de vue de l’agent décrit dans le modèle.

Conclusion

Au cours de cette longue enquête, nous avons peut-être donné l’impression de vouloir placer tous les arguments de notre côté plutôt que de

choisir celui qui nous paraît décisif. Mais le fait est qu'on ne peut dissocier complètement les trois raisons qui obligent, nous semble-t-il, à juger sévèrement les tentatives du *search* et, plus généralement, celles de l'optimisation au second ordre. Ces tentatives soulèvent un problème (a) empirique – la plupart des données “naïves” sont défavorables à l'optimisation – (b) méthodologique – les modèles du *search* font dégénérer le programme d'une théorie optimisatrice de la décision – et (c) logique – ils entraînent l'agent dans une régression infinie qui, pour lui sinon pour l'observateur, est toujours “vicieuse”. La proximité des rapports qu'entretiennent ces trois objections – elles paraissent toujours sur le point de se transformer l'une dans l'autre – s'explique aussitôt, lorsque l'on considère leur point d'impact dans les modèles incriminés: celui-ci est bien le même pour les trois objections. Le problème unique et lancinant que pose l'optimisation au second ordre tient à ce qu'elle se *donne* les paramètres qui – dit très généralement – quantifient la difficulté à optimiser. Dans le cas de Stigler et de sa nombreuse postérité, que signifient ces “coûts de recherche” dont l'agent disposerait, comme par miracle, afin de réviser sa décision du premier ordre? On peut moduler cette question de différentes manières: en affirmant que ces coûts n'existent pas en tant que tels, puisqu'ils ne sont observables ni sur le marché, ni dans la psychologie de l'agent; en montrant que l'introduction de ces coûts affaiblit le contenu testable de l'optimisation; ou encore en rendant sensible la régression infinie dans laquelle la théorie classique se précipite inconsciemment. Il s'agit là de variations sur un thème commun.

Il est remarquable que les conceptions de la *rationalité procédurale* ne cherchent pas à quantifier la rareté des ressources cognitives de l'agent: dans les descriptions que les meilleurs philosophes nous ont laissées des processus délibératifs, dans les travaux les moins programmatiques de l'intelligence artificielle – comme ces études du joueur d'échecs écrites par Newell, Shaw et Simon – la notion de “coût de recherche” n'apparaît nulle part. Le programme de recherche recommandé par Simon évite, de ce fait, bien des difficultés. On peut lui reprocher de s'interdire par avance de traiter le paradoxe d'une délibération inefficace à force d'efficacité: celui d'un processus tellement long et coûteux qu'il consumerait toutes les ressources que, précisément, la décision doit allouer. Mais on a vu que l'optimisation ne parvient pas à résoudre véritablement ce paradoxe: faut-il reprocher aux conceptions procédurales de ne pas permettre de la soulever?

A la réflexion, on ne doit pas s'étonner qu'un programme centré sur les problèmes de la *délibération* évite la notion de “coût de recherche”. Il serait tout aussi juste de dire qu'en l'introduisant, la théorie classique

exclut toute préoccupation délibérative. En se donnant les “coûts de recherche”, la théorie transforme le problème de décision en *calcul*. Le propre du raisonnement calculatoire est qu’il opère sur des symboles qui sont physiquement disponibles: il transforme du donné. C’est pourquoi Engels (et d’autres avant et après lui) ont cru pouvoir analyser l’activité mathématique suivant l’analogie de la production. Le raisonnement délibératif n’opère pas exclusivement sur des données, il consiste pour partie en une invention d’objets. Chacun a pu faire l’expérience de la préparation d’un coup aux échecs: la manoeuvre à laquelle on se résout ne préexistait pas mentalement à sa découverte; elle s’est constituée dans le cours de la délibération.

La concentration de difficultés qui entourent l’optimisation et le programme “substantiel” et nous inciteraient à conclure sur leur échec, ne doit pas induire, par contrecoup, une vision trop flatteuse du programme “procédural”. Celui-ci recouvre un ensemble de thèses ou d’opinions passablement hétérogènes et quelquefois très peu articulées. L’analyse spécifique de la *bouleusis*³⁶ a sollicité depuis longtemps les philosophes et les écrivains, plus récemment des auteurs qui se rattacheraient à la nébuleuse de l’intelligence artificielle, mais ce n’est guère que dans les écrits de Simon qu’elle apparaît comme un programme de recherche, au sens que Popper ou Lakatos donneraient à ce mot. Encore n’est-ce qu’un programme embryonnaire. On ne peut soupçonner que très vaguement les obstacles qu’il rencontrera. L’un d’entre eux, néanmoins, nous semble déjà perceptible. Le principe délibératif doit s’exprimer dans la définition adéquate d’un ordre de parcours sur l’ensemble des actions ou celui des conséquences. Pour l’observateur, cette difficulté – qui reflète le phénomène que nous avons désigné comme invention d’objets – s’exprimera de la manière suivante: quel *mécanisme d’engendrement* (au sens du 3.3) faut-il définir pour ces ensembles? La théorie du *search* supposerait dans ce cas une loi de probabilité, solution triviale qui désigne un problème plutôt qu’elle ne le résout, mais qui peut devenir inévitable en l’absence d’une plus ample information. Des analyses procédurales comme celles du joueur d’échecs ont sans doute privilégié une situation qui est trop particulière: le problème du joueur d’échecs a une structure combinatoire préalable, grâce à laquelle divers mécanismes d’engendrement des coups ou des conséquences peuvent se définir assez naturellement. Mais qu’en est-il des ensembles qui n’ont pas de structure combinatoire *a priori*, par exemple de l’ensemble des règles de décision possibles pour un chef d’entreprise? La réussite du programme supposerait toutefois que l’on parvienne à modéliser les processus de décision des chefs d’entreprise avec la même rigueur formelle que l’on a su appliquer aux délibérations du joueur d’échecs.

Philippe Mongin (né en 1950) est Chargé de Recherches au Laboratoire d'Economie Politique de l'Ecole Normale Supérieure. Ses domaines de recherche sont théorie de la décision et l'épistémologie des sciences sociales. *Articles récemment publiés*: "Modèle rationnel ou modèle économique de la rationalité?", *Revue Economique* 35 (jan.), 1984; "Un modèle de politique monétaire avec aléas stratégiques", *Revue Economique* 35 (sept.) (1984) (avec P. A. Chiappori); "La controverse sur l'entreprise (1940-1950) et la formation de l'irréalisme méthodologique", *Oeconomia* 5 (1986); "Are 'all-and-some' statements falsifiable after all? The example of utility theory", *Economics and Philosophy* (1986) (à paraître). *Adresse de l'auteur*: Laboratoire d'Economie Politique, ENS, 45 rue d'Ulm, 75005 Paris, France.

Notes

1. Simon (1955), p. 245. Ces objections contre le traitement économique de l'incertitude sont reprises par exemple en 1959 (in 1966, p. 13).

2. En 1959, Simon reste conciliant: il fait ressortir l'insuffisance des théories "classiques" lorsqu'on aborde les problèmes de la dynamique, ou certains secteurs qu'il désigne nommément (1959, pp. 3, 25). Il se pourrait qu'il s'agisse là d'une précaution rhétorique. Lorsqu'il mentionne le "choix dans l'incertain" parmi les "secteurs" où les théories classiques sont défaillantes, Simon porte en réalité contre elles une condamnation d'ensemble.

3. Dans sa visée explicite, cette théorie se propose d'étendre l'hypothèse de choix optimisateur au traitement de l'information: l'agent formerait des anticipations d'une manière qui maximise ses profits ou son utilité.

4. Le représentant le plus éminent de ce courant, M. Friedman, a proclamé que l'on peut fonder une science empirique sur des prémisses "irréalistes"; certains zéloteurs revendiquent même la possibilité de faire reposer la science économique sur des prémisses *fausses*. Cette méthodologie étrange, mais fort répandue aujourd'hui, est discutée in Mongin (1984a).

5. On peut se demander si la méthodologie de l'"irréalisme" telle que Friedman l'a énoncée (1953) implique logiquement une conséquence aussi déterminée. A notre avis, ce n'est pas le cas. Il n'en reste pas moins que l'influence *pratique* de cette méthodologie se mesure à la force de cette idée: le test porte sur le comportement au marché des agents représentatifs, c'est-à-dire sur les relations agrégées d'offre et de demande.

6. Elle apparaît déjà fondamentale dans la revue critique des théories de la décision (1959, in 1966, p. 16).

7. Cf. Marschak et Radner (1972), Radner (1972). On va voir qu'il s'applique également, comme à deux cas particuliers, à Savage (1954) et Simon (1955).

8. Ce postulat fondamental rattache la théorie de la décision à une attitude philosophique générale que l'on peut appeler *conséquentialisme*. Jusqu'à présent, on a surtout traité du conséquentialisme dans le contexte restreint des questions morales (p. ex. Williams (1973)). Mais cette attitude n'est pas moins problématique lorsqu'elle apparaît dans le contexte plus large d'une théorie de la rationalité individuelle. Il est significatif que certains travaux d'intelligence artificielle paraissent la remettre en cause (nous pensons à la distinction intéressante des "coups" et des "stratégies" pour le joueur d'échecs - cf. Simon (1972a), p. 169).

9. Il est vrai que, à un moment donné de la partie, un joueur d'échecs ignore le coup que va choisir son adversaire: de ce point de vue, il est soumis à une incertitude externe. Mais *si le joueur suppose que son adversaire est rationnel*, cette incertitude externe disparaît: le théorème de Zermelo lui garantit que, pour chacun de ses coups, il existe une meilleure réponse de son adversaire. L'incertitude du joueur d'échecs est alors due au fait qu'il ne peut calculer cette meilleure réponse, ni énumérer la totalité des coups dont il dispose. Elle est purement interne.

10. En un sens différent, et sans doute plus radical, que celui retenu par Munier (1984).

11. Cf. la notion d' "ensemble d'information" chez Marschak et Radner (1972, *passim*).

12. Cf. Simon (1979), p. 502: "Two concepts are central to the characterization: *search and satisficing*." Nous avons suivi cette caractérisation coutumière dans une présentation antérieure de la rationalité limitée (Mongin, 1984a; 1984b).

13. Dans son article de 1955, Simon ne traite pas explicitement de *l'engendrement* des $a \in A$: sans doute suppose-t-il donné un ordre canonique sur cet ensemble. C'est, semble-t-il, grâce à des travaux d'intelligence artificielle qu'il en vient à poser le problème de l'engendrement des éléments de A ou de S (cf. Newell, Shaw et Simon, 1958). Par ailleurs, en 1955, Simon choisit de raisonner sur un niveau d'aspiration constant (p. 252). Néanmoins, il écrit alors, et ne cesse d'affirmer par la suite (1959, 1979a) que les agents révisent leurs seuils de satisfaction: le modèle de 1955 ne peut être qu'une simplification provisoire.

14. Sur les rapports logiques entre le principe des seuils de satisfaction et celui de la recherche séquentielle, cf. Mongin (1984b), pp. 31–32.

15. Simon (1972a), p. 168. Dans le jeu d'échecs, une *action* est un coup (p. ex. jouer le pion d4) et une *conséquence* est un état du jeu obtenu par continuation à partir de ce coup (p. ex. gambit du pion dame accepté). La distinction des deux processus de recherche y est donc particulièrement facile à saisir, ce qui n'est malheureusement pas toujours le cas dans d'autres contextes. Il faut aussi bien voir que, grâce à la structure combinatoire du jeu, il existe un *mécanisme d'engendrement naturel des conséquences* lorsque l'on fixe l'action $a = \bar{a}$: le sous-ensemble de conséquences associé à l'étape n , S_a^n , se déduit de S_a^{n-1} par l'analyse des continuations, c'est-à-dire par des moyens purement combinatoires. De ce fait, la tâche de spécialiste d'intelligence artificielle apparaît grandement simplifiée: il pourra, en première analyse tout au moins, concentrer ses efforts sur la définition d'une *règle d'arrêt* appropriée en supposant résolu le problème de l'engendrement. Nous reviendrons sur la portée de cette simplification.

16. Cf. Simon (1955), p. 248. L'analyse des continuations, aux échecs, fournit l'exemple le plus simple de ce premier processus de recherche.

17. Newell, Shaw et Simon (1962) ont – en substance – soulevé cette question, sans prétendre lui donner une réponse générale; cf. aussi l'ouvrage de Newell et Simon (1972).

18. Cf. Simon (1972a), p. 410: "the decision question has been switched to the question of how much of the actor's resources should be allocated to search"; et Baumol et Quandt (1964), p. 23: "the marginal cost of additional information gathering or more refined calculation [should] be equal to its marginal (expected) gross yield".

19. Nous nous sommes facilité la tâche en supposant qu'un seul magasin, m_p , est associé à un prix p donné.

20. Parmi les autres critères pertinents, on peut faire figurer: la durée de la recherche (indéfinie, finie ou aléatoire); la présence ou non d'un taux d'actualisation; la nature du mécanisme de révision pour les $F^n(p)$.

21. Cette hypothèse de *uncertain recall* apparaît dans un modèle de Karni et Schwartz (1977).

22. P. ex. Gastwirth (1976), et Marciano, (1982), qui fait une bonne revue des transformations apportées au modèle de Stigler.

23. Cf. Kohn et Shavell (1974), pp. 96–97. Les auteurs ne discutent malheureusement pas la portée de la restriction qu'ils ont imposée *a priori* sur l'ensemble des stratégies disponibles.

24. Plaçons-nous dans le cas particulier de Stigler, où le seul problème du consommateur est d'acheter au plus bas prix. On démontre l'existence d'un seuil r_t au-dessous duquel le consommateur doit acheter en étudiant l'équation intégrale:

$$p_{\min}^t \int_{r_t}^r {}^t(r_t - p) dF^t(p) = c(t + 1)$$

Cette équation s'interprète facilement en termes marginalistes: si le consommateur rencontre un prix $p = r_t$, son *coût marginal* attendu (membre de droite) est exactement compensé par le gain espéré d'une réduction dans le prix r , c'est-à-dire par son *gain marginal attendu* (membre de gauche).

25. Eventuellement, mais non nécessairement, bayésien. Rappelons, encore une fois, que la théorie du *search* est compatible avec un mécanisme de révision *quelconque* des probabilités.

26. Pour l'axiomatique de la théorie statique, cf. Debreu (1959); pour celle de l'utilité espérée, différentes axiomatiques sont envisageables, principalement celle de von Neuman-Morgenstern (reprise Harsanyi (1977)) et celle de Savage (1954).

27. Sur le développement de cette discipline, cf. Simon (1978a, 1978b).

28. D'après cet axiome, si une loterie a la forme $L = \{1F \text{ avec proba } p, L' \text{ avec proba } 1-p\}$, L' étant une loterie $= \{2F \text{ avec proba } q, 0, 4F \text{ avec proba } 1-q\}$, l'agent est indifférent entre L et $L'' = \{1F \text{ avec proba } p, 2F \text{ avec proba } (1-p)q\}, 0, 4F \text{ avec proba } (1-p)(1-q)\}$. La théorie classique de la décision se préoccupe si peu des coûts de calcul que, dans la présentation de Harsanyi (1977, p. 24), le principe des "loteries composées" *ne figure même pas parmi les axiomes*: il est présenté comme une simple "convention d'écriture".

29. Cf. les articles réunis dans la dernière partie de *Models of Discovery* (1977).

30. C'est ici qu'il faut souligner la parenté entre Stigler et un autre économiste chicagôite, Becker. Pour celui-ci, l'approche "économique" consiste justement à assigner des coûts à chaque phénomène de rareté (1976, p. 6). L'expression de *economic approach* est sans doute trop vague pour désigner une heuristique aussi précise. C'est plutôt, nous semble-t-il, une constante des *approches optimisatrices* de la décision que Becker identifie – avec profondeur – lorsqu'il énonce cette définition. On trouvera une discussion utile de la méthodologie beckerienne in Rosenberg (1979).

31. Nous nous plaçons dans le cas le plus favorable, qui est également celui du *search*: la théorie s'intéresse à la demande adressée à l'entreprise par un consommateur *isolé* et pour un bien *normal*. Dans ce cas, la micro-économie élémentaire peut conclure à la forme décroissante de la courbe de demande. Il n'en va plus de même lorsqu'on en vient à la demande *totale* adressée à l'entreprise: la répartition des revenus et des préférences influence alors la courbe de demande, qui peut fort bien n'être pas décroissante, même pour un bien normal.

32. Les exposés canoniques sur cette méthodologie se trouvent dans Popper (1963) et Lakatos (1970). On trouvera des développements utiles dans Boyer (1983), qui précise notamment les différences entre Popper et Lakatos.

33. Nous avons retenu la formulation de Winter en 1975, parce qu'elle nous semble plus claire que celle d'un article antérieur (1964), où il mettait déjà en évidence les possibilités de régression infinie contenues dans la théorie optimisatrice. Elster (1979, p. 57) commente Winter en rapprochant ingénieusement la critique logique de l'optimisation et la théorie cartésienne de l'action.

34. Chiappori (1984) fait une utile mise au point sur cette construction dont l'idée remonte à Alchian (1950) et Friedman (1953).

35. Un coup d'oeil à l'histoire de la théorie économique montre que celle-ci cherche très souvent à étayer ses postulats fondamentaux en s'aidant de résultats que les autres sciences pourraient – logiquement – ou devraient – dans un avenir proche – lui fournir. La psychologie a joué pendant un temps ce rôle de pseudo-fondation (voir p. ex. Rosenberg, 1981). Alchian et Friedman lui substituent une discipline créée de toutes pièces, aux confins de la biologie et de l'économie, la "théorie" de la "sélection économique naturelle". On peut envisager bien des modulations sur ce thème singulier.

36. Aristote distingue soigneusement l'acte du choix, ou *proairesis*, qui consiste à désigner l'objet préféré, et la délibération antérieure, ou *bouleusis*, qui consiste à passer en revue les moyens, à prévoir les conséquences et à rendre celles-ci commensurables aux fins. Cette distinction annonce très nettement celle des deux types de rationalité, *substantive* et *procedural*, chez Simon. En même temps qu'il approfondit la distinction des deux concepts, Aristote montre qu'ils coïncident extensionnellement: "L'objet de la délibération et l'objet du choix sont identiques" (*Ethique à Nicomaque*, III, 5, 1113a). Considérée isolément, cette thèse ferait d'Aristote l'ancêtre du programme procédural en théorie de la rationalité. En fait, il faut la rapprocher d'autres passages des deux *Ethiques* où il semble prendre parti en faveur de l'optimisation. Par ailleurs, Aristote s'éloigne certainement du programme procédural lorsqu'il défend la thèse célèbre du caractère donné, et intangible au cours de la délibération, des fins subjectives ("Nous délibérons non pas sur les fins elles-mêmes, mais sur les moyens d'atteindre les fins" (1112b)).

Références

A. Travaux de H. A. Simon

On trouvera sous cette rubrique, rangés par ordre chronologique, tous les articles et ouvrages de H. A. Simon auxquels le texte fait référence, y compris des travaux écrits en collaboration. Ces derniers sont aussi inclus dans la rubrique "B. Autres travaux".

- 1955 "A behavioral model of rational choice", *Quarterly Journal of Economics* 69: 99–118; repris in: Simon (1957), pp. 241–260 (édition citée).
- 1957 *Models of Man*. New York, Wiley.
- 1958 (en collaboration avec A. Newell et J. C. Shaw) "Chess-playing programs and the problem of complexity", *Journal of Research and Development* 2: 320–335.
- 1959 "Theories of decision making in economics and behavioral science", *American Economic Review* 49: 253–283; repris in: Royal Economic Society and American Economic Association, *Surveys of Economic Theory*, (1966). Londres, MacMillan, t. III, chap. I (édition citée).

- 1962 (en collaboration avec A. Newell et J. C. Shaw) "The process of creative thinking", pp. 63–119 in: H. E. Gruber, G. Terrell and M. Wertheimer (eds.), *Contemporary Approaches to Creative Thinking*. Lieber-Atherton; repris in: Simon (1979b), pp. 140–174.
- 1963 "Problems of methodology – discussion", *American Economic Review* 53: 229–231.
- 1964 "Rationality", pp. 573–574 in: J. Gould et W. L. Kolb (eds.), *A Dictionary of the Social Sciences*, Glencoe, IL, The Free Press.
- 1972a "Theories of bounded rationality", pp. 161–186 in: R. Radner et C. B. McGuire (eds.) (1972).
- 1972b (en collaboration avec A. Newell) *Human Problem Solving*. Englewood Cliffs, NJ, Prentice Hall.
- 1976 "From substantive to procedural rationality", pp. 129–148 in: S. J. Latsis (ed.), *Method and Appraisal in Economics*. Cambridge, Cambridge University Press.
- 1977 *Models of Discovery*. Dordrecht, Reidel.
- 1978a "On how to decide what to do", *The Bell Journal of Economics* 9: 494–507.
- 1978b "Rationality as process and as product of thought", *American Economic Review* 68: 1–16.
- 1979a "Rational decision making in business organizations" (texte du discours que Simon a prononcé à Stockholm en 1978 pour la réception du prix Nobel), *American Economic Review* 69: 494–513.
- 1979b *Models of Thought*. New Haven, CT, Yale University Press.
- 1983 *Models of Bounded Rationality* (2 tomes). Cambridge, MIT Press.

B. Autres Travaux

Alchian, A.

- 1950 "Uncertainty, evolution and economic theory", *Journal of Political Economy* 58: 211–222.

Aristote

- 1972 *Ethique à Nicomaque* (trad. J. Tricot). Paris, Vrin.

Baumol, W. J.; Quandt, R. E.

- 1964 "Rules of thumb and optimally imperfect decision", *American Economic Review* 54: 23–46.

- Becker, G. S.
1976 *The Economic Approach to Human Behavior*. Chicago, IL, University of Chicago Press.
- Boyer, A.
1983 "L'explication en histoire. Essai sur des théories anglo-saxonnes de l'Histoire". Paris, Université de Paris I – Panthéon Sorbonne. (Thèse pour le doctorat de troisième cycle.)
- Chiappori, P. E.
1984 "Sélection naturelle et rationalité absolue des entreprises", *Revue Economique* 35: 87–108.
- Debreu, G.
1959 *Theory of Value*. New York, Wiley.
- Elster, J.
1979 *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge, Cambridge University Press.
- Friedman, M.
1953 "The methodology of positive economics", pp. 3–46 in: *Essays in Positive Economics*. Chicago, IL, University of Chicago Press.
- Friedman, M.; Savage, L. J.
1948 "The utility analysis of choices involving risk", *Journal of Political Economy* 56: 279–304.
1952 "The expected-utility hypothesis and the measurability of utility", *Journal of Political Economy* 60: 463–474.
- Gastwirth, J. L.
1976 "On probabilistic models of consumer search for information", *Quarterly Journal of Economics* 90: 38–50.
- Harsanyi, J. C.
1977 *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge, Cambridge University Press.
- Kahneman, D.; Tversky, A.
1974 "Judgment under uncertainty: heuristic and biases", *Science* 185: 1124–1131.
- Karni, E.; Schwartz, A.
1977 "Search theory: the case of search with uncertain recall", *Journal of Economic Theory* 16: 38–52.
- Kohn, M. G.; Shavell, S.
1974 "The theory of search", *Journal of Economic Theory* 9: 92–123.

- Lakatos, I.
1970 "Falsification and the methodology of scientific research programmes", in: I. Lakatos et A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge, Cambridge University Press.
- Marciano, M.
1982 "Acquisition optimale d'information par le consommateur sur le prix d'un bien homogène". Aix-en-Provence, Faculté d'économie appliquée, GRASCE. (Note de recherche 82.01.)
- Marschak, J.; Radner, R.
1972 *Economic Theory of Teams*. New Haven, CT, Yale University Press.
- Mongin, P.
1984a "Réexamen du déterminisme situationnel. Critique de quelques exemples monétaristes", Aix-en-Provence, Faculté d'économie appliquée. (Thèse d'Etat.)
1984b "Modèle rationnel ou modèle économique de la rationalité?", *Revue Economique* 35: 9-64.
1985 "Est-il logiquement impossible d'optimiser?", Paris, Ecole Normale Supérieure, Laboratoire d'économie politique. (Document de recherche 104.)
- Munier, B.
1984 "Quelques critiques de la rationalité économique dans l'incertain", *Revue Economique* 35: 65-86.
- Newell, A.; Shaw, J. C.; Simon, H. A.
1958 "Chess-playing programs and the problem of complexity", *Journal of Research and Development* 2: 320-335.
1962 "The process of creative thinking", pp. 63-119 in: H. E. Gruber, G. Terrell et M. Wertheimer (eds.), *Contemporary Approaches to Creative Thinking*. Lieber-Atherton; repris in: Simon (1979b), pp. 140-174.
- Newell, A.; Simon, H. A.
1972 *Human Problem Solving*. Englewood Cliffs, NJ, Prentice Hall.
- Popper, K.
1934 *Logik der Forschung*. Trad. anglaise: *Logic of Scientific Discovery*. Londres, Hutchinson (1ère édition 1959, 6ème édition révisée 1972).
1954 *Postscript*, t. III, aujourd'hui publié sous le titre *Quantum Theory and the Schism in Physics*. Londres, Hutchinson, 1982.
- Radner, R.
1972 "Normative theory of individual decision: an introduction", chap. 1, pp. 1-18 in: R. Radner et C. B. McGuire (eds.) (1972).
- Radner, R.; McGuire, C. B. (eds.).
1972 *Decision and Organization. A Volume in Honour of Jacob Marschak*. Amsterdam, North-Holland.

- Rosenberg, A.
1979 "Can economic theory explain everything?", *Philosophy of the Social Sciences* 9: 509–529.
1981 "A sceptical history of microeconomics", pp. 47–61 in: J. C. Pitt (ed.), *Philosophy in Economics*. Dordrecht, D. Reidel.
- Rosenfield, D. B.; Shapiro, R. D.
1981 "Optimal adaptive price search", *Journal of Economic Theory* 25: 1–20.
- Rothschild, M.
1974 "Searching for the lowest price when the distribution of prices is unknown", *Journal of Political Economy* 82: 689–711.
- Savage, L.
1954 *The Foundations of Statistics*. New York, Dover Publications (2ème édition 1972) (édition citée).
- Stigler, G. J.
1961 "The economics of information", *Journal of Political Economy* 69: 213–225.
- Williams, B.
1973 "A critique of utilitarianism", pp. 77–150 in: J. J. C. Smart et B. Williams, *Utilitarianism For and Against*. Cambridge, Cambridge University Press.
- Winter, S. G.
1964 "Economic natural selection and the theory of the firm", *Yale Economic Essays* 4: 225–272.
1975 "Optimization and evolution in the theory of the firm", pp. 73–118 in: R. H. Day et T. Groves (eds.), *Adaptive Economic Models*. New York, Academic Press.