

**"IL DEVIENT TOUT AUSSI PASSIONNANT
de rechercher l'organisation des processus de pensée
que de découvrir l'organisation du mouvement des planètes .**

Commentaires et réponses présentées par *H.A. Simon*¹

au Colloque de la Grande Motte, 1-4 février 1984

SCIENCES DE L'INTELLIGENCE, SCIENCES DE L'ARTIFICIEL, avec Herbert A. Simon,
(publié dans les actes édités par les PUL (1986, ISBN 7297 0287 3, épuisé) pp. 577-612

Avertissement : Ce texte, qui ne fut publié qu'en français, reprend l'enregistrement des réponses et commentaires que H.A.Simon proposa en guise de conclusion lors du Colloque de la Grande Motte, 1-4 février 1984 «SCIENCES DE L'INTELLIGENCE, SCIENCES DE L'ARTIFICIEL, avec Herbert A. Simon, aux soixante dix participants de ce séminaire original: Pendant trois jours, ceux ci venaient de présenter leurs lectures de l'œuvre transdisciplinaire de H A Simon et les questionnements qu'elles suscitaient. Nous disposons ainsi d'un document original qui permet de brosser une sorte de perspective cavalière de la pensée et de l'œuvre d'H.A. Simon qui témoigne de son étonnante unité dans la diversité des expériences disciplinaires qu'il a pratiquées. Ce texte fut publié dans les actes édités par les PUL (1986), actes aujourd'hui épuisés. JLM.

Au long de ces trois journées, j'ai entendu une discussion des écrits d'un nommé Herbert Simon, lequel a manifestement écrit, depuis 47 ans, beaucoup de choses sur diverses questions qui relèvent des sciences sociales. Je présume qu'il n'est pas très surprenant qu'il ait tant écrit sur tant de sujets, en 47 années ; peut-être même qu'il ait pu oublier, certaines années, ce qu'il avait écrit en d'autres moments ; et qu'il se soit ainsi parfois contredit. Comme on a tenté d'organiser l'ensemble de ses écrits autour de quelques thèmes centraux, je vais essayer de présenter mes réflexions de ce matin de la même façon. De ce fait, quelques-uns de mes écrits seront laissés de côté : je ne vous en parlerai pas maintenant, même si une partie d'entre vous en favorise certains qui ne s'intègrent pas dans ce cadre général.

Ce matin, mes commentaires seront surtout basés sur vos textes écrits et moins sur les exposés et discussions ; d'une part, parce que j'ai disposé de plus de temps pour assimiler les textes écrits ; d'autre part, parce que je ne suis pas certain d'avoir toujours interprété correctement les idées qui furent développées au cours des débats. En outre, mes commentaires seront surtout orientés en direction de ce qui me semble être les questions ouvertes : celles qui ne me paraissent pas avoir été bien résolues, celles pour lesquelles il pourrait exister quelques désaccords ; celles sur lesquelles j'aimerais me démarquer par rapport aux positions exprimées dans les textes présentés.

¹ Traduit de l'anglais, à partir d'une transcription de l'enregistrement de l'exposé du Professeur H.A. Simon lors de la séance de clôture du colloque de la Grande-Motte (3 Février 1984). Francette Perrot a assuré la transcription ; Jean-Louis Le Moigne s'est chargé de la traduction et des notes en bas de page. Texte revu et corrigé par H.A. Simon, le 4 Juillet 1984. En septembre 2001, André Demailly a assuré la numérisation et la présentation de ce document à partir de l'original dactylographique qu'il avait eu la sagesse de conserver, ce dont les le Programme européen MCX le remercie très sincèrement. .

En calculant le temps dont je disposerais pour faire au moins un bref commentaire de chacun de vos papiers, je me suis aperçu que je n'aurais que 5 minutes, à peine, pour chacun d'eux. Or, certains d'entre eux sont très intéressants et méritent d'y passer plus de temps. C'est pourquoi, je ne serai pas toujours en mesure de citer spécifiquement, dans mes commentaires, chacune de vos contributions. J'espère que vous me pardonneriez de ne pas aborder directement votre thème favori mais, de toute façon, nous pourrions y revenir ensuite lors de la discussion générale.

Par ailleurs, dans ma conception de la production scientifique, le scientifique en tant qu'individu n'est pas une île. Tous les travaux scientifiques prennent place dans un fleuve, d'autres travaux, sont toujours influencés par ceux qui les entourent et ceux qui les ont précédés ; et, s'il s'agit d'un travail qui a abouti avec succès, il va à son tour influencer ce vaste courant. Une publication scientifique n'est pas forcément une synthèse de tout ce qu'on connaît ni même de tout ce qu'on connaît sur une question particulière. Une publication scientifique est un rapport sur un progrès en cours (*«progress report»*) qui contribue à former ce courant, en rejoignant d'autres rapports sur d'autres progrès. Les gens y prennent tantôt de l'eau et tantôt en déversent (je ne suis pas sûr que l'image de l'eau soit la bonne métaphore... peut-être faudrait-il parler de «prendre du vin» ! En tout cas, ils prélèvent et déversent quelque chose dans ce fleuve de la production scientifique. C'est pourquoi j'étais très heureux du ton général de toutes vos discussions qui visaient beaucoup moins ce que Herbert Simon avait dit ou voulait dire que ce que cela avait à voir avec une foule de questions et d'idées. Ces questions et idées ne sont, «en aucune façon, issues de moi-même mais le produit de ce *«Zeitgeist»* que j'ai essayé de vous évoquer hier soir ² - Il s'agit d'un courant de pensée très important, qui s'est développé au long de ce siècle et dont j'aime espérer que mon oeuvre fait intégralement partie. C'est ainsi que je vous ai parlé du fleuve constitué par le traitement de l'information, la cybernétique et les idées systémiques... et de ce que cela peut signifier pour le monde de la science contemporaine.

Enfin, avant de m'embarquer dans mes commentaires proprement dits, je voudrais exprimer ma gratitude pour l'esprit très exceptionnel qui a marqué toute cette rencontre scientifique. Lorsque je suis arrivé ici, je connaissais déjà quelques-uns d'entre vous et certains sont de vieux amis. Mais je crois que j'ai gagné ici de nombreux nouveaux amis et j'espère que ces amitiés pourront continuer, car nous avons beaucoup de choses en commun. Cette rencontre constituera désormais pour moi un souvenir plein de chaleur.

Il y aurait peut être une autre façon de vous dire la même chose, en vous assurant que vous avez été trop généreux à l'égard des oeuvres de Herbert Simon et pas assez critiques pour divers défauts qu'elle dissimule ! ... mais je ne le ferai pas...

J'ai organisé mes commentaires autour de questions qui correspondent en gros aux diverses sessions du colloque, mais je jouerai un peu avec cette règle car certains thèmes croisent plusieurs séances. Je me propose donc de vous parler de quelques questions en psychologie ; puis de quelques autres en management, en organisation et, plus généralement, en systémique ; ensuite, de quelques questions en économie puis en intelligence artificielle... pour revenir enfin un peu plus sur la science des systèmes.

Ce qui correspond à peu près au découpage adopté dans le colloque. J'essaierai de consacrer 5 minutes à chaque sujet, mais certains me retiendront plus que d'autres, tant la liste est considérable. J'ai tenté, par ailleurs, d'incorporer dans mes commentaires les questions écrites que nombre d'entre vous m'avez transmises. Je vous redis, à nouveau, que si vous avez l'impression que j'ai éludé votre question ou ne l'ai pas traitée de façon adéquate, de me demander d'y revenir tout à l'heure, lors de la discussion générale. A l'occasion, j'aurai à nommer tel ou tel auteur à titre d'exemple, prenant ainsi le risque de vous offenser — soit en vous faisant dire des choses que vous n'avez pas dites, soit en ne mentionnant pas votre

² Cf. Quelques remarques historiques sur la science de la cognition (H.A. Simon).repris dans le document "Introuvable N° 2)

nom — mais il faut bien que je prenne ce risque.

Psychologie

Il y a eu beaucoup de discussions sur la simulation par ordinateur et ses relations avec la psychologie cognitive. Je pense, par exemple, aux préoccupations de M. Tiberghien quant aux «risques de la science de la cognition» tels qu'il les décrit dans son texte à propos des problèmes méthodologiques posés par cette science. Toute science doit se débattre avec les procédures de formulation de ses théories.

Depuis Newton, les sciences physiques ont découvert un outil puissant qui est encore leur principal outil de formulation de théories. Nombre de ces théories physiques s'expriment par des systèmes d'équations différentielles ou, pour les cas où elles doivent déboucher sur des calculs sans pouvoir disposer d'un temps continu, elles y substituent de petits intervalles de temps discrets et formulent la théorie sous forme d'équations aux différences. Ainsi, elles font appel à des systèmes d'équations différentielles et à des systèmes d'équations aux différences. Je n'ai pas besoin de vous détailler les avantages de cette double formulation : c'est la disponibilité de ces méthodes qui a rendu possible les succès de la physique, y compris celui de la physique quantique moderne.

Sans doute, en sciences économiques, avons-nous disposé aussi de telles théories mathématiques ; mais dans les autres sciences sociales nous sommes beaucoup plus démunis. Pour l'essentiel, nous avons dû théoriser avec des mots. Et bien que les mots d'une langue naturelle soient des outils très puissants pour un certain nombre d'objectifs, ils ne s'avèrent pas très satisfaisants pour décrire les interrelations complexes de systèmes à multiples composants. Même si l'on peut décrire de tels systèmes à l'aide du langage naturel, ces descriptions ne nous fournissent pas les moyens d'analyser le comportement du système et, en particulier, les interactions de ses composants.

Les théories en forme de programmes informatiques

Les programmes sont aujourd'hui des langages formels permettant d'exprimer les théories psychologiques sous la forme d'équations aux différences, puisque les ordinateurs modernes procèdent par cycles et que chaque cycle réalise une opération. L'opération réalisée dépend de l'état du système et de ses intrants à l'instant considéré. L'état d'un ordinateur à un instant donné, quel que soit cet état, et ce qui entre dans ce ordinateur à cet instant déterminent ce qu'il fait et déterminent donc ses extrants et son état à l'instant suivant. Formellement donc, dans le sens le plus rigoureux de ce terme, ce programme informatique est un système d'équations aux différences.

Une activité importante de l'entreprise dite du traitement de l'information va consister, dès lors, à utiliser le nouveau langage des ordinateurs pour décrire les processus mentaux humains par de telles équations aux différences. Mais il s'agit là d'un type fort peu classique d'équations aux différences, puisqu'elle ne sont pas construites à l'aide de nombres réels mais avec des structures beaucoup plus générales de symboles.

On se pose parfois la question de savoir si un programme informatique peut être une théorie (il ne s'agit pas ici de se demander si ladite théorie est correcte, ce qui est une tout autre question). J'ai du mal à comprendre cette question. Si les équations aux différences et les équations différentielles peuvent être des théories, il «en va exactement de même d'un programme de traitement de l'information écrit dans un

langage de programmation tel que LISP, OPS ou PROLOG ..., quel que soit le langage de programmation que vous préférez. C'est exactement le même type d'objet formel et peut dès lors être utilisé comme une théorie. De plus, il peut être testé empiriquement et expérimentalement comme n'importe quelle théorie ; parce que si vous disposez d'un tel programme, vous pouvez étudier le comportement du système de traitement de l'information qu'il prédit, en le faisant fonctionner dans un contexte donné (celui de l'environnement du problème) ; puis vous pouvez tester ces prédictions en regard de tel comportement humain.

Je ne vois rien d'anormal — j'allais dire rien de «révolutionnaire» — dans cette méthodologie, sinon peut-être la généralisation de variables réelles à des variables ayant des structures symboliques. C'est pour cette raison que j'ai décidément un peu de mal à comprendre des questions, comme celles de M. Hoc, demandant si les programmes peuvent être des théories. En outre, je ne vois aucune innovation, à une petite exception près que je mentionnerai plus loin, dans les méthodes que nous proposons, en psychologie du traitement de l'information pour tester, valider ou falsifier ces modèles. Les questions méthodologiques sont exactement les mêmes que celles que nous rencontrons dans n'importe quelle autre science. Le seul problème est de trouver un ensemble de données qui puissent être comparées avec les prédictions du système.

Le système dont nous parlons ici est un système dynamique, qui change au fil du temps, même si ce n'est pas très rapidement. Le cerveau humain est fait de neurones ; et les neurones fonctionnent par des processus électromagnétiques et chimiques, qui sont des dispositifs plutôt lents. On ne penserait jamais aujourd'hui à essayer de construire un ordinateur à partir de neurones, parce que ce ordinateur serait tenu pour incroyablement lent. Il ne se passe rien, je le soutiens, dans un système neuronal, «en une milliseconde (un millième de seconde) - En dix millisecondes, il commence à se produire un petit quelque chose, et beaucoup plus en cent millisecondes, mais ce ne sont pas des vitesses très impressionnantes. Le problème de la vérification empirique et expérimentale consiste à repérer ce qui se passe dans un système dont le comportement doit être enregistré toutes les cent (ou quelques centaines de) millisecondes.

La seule modeste innovation qui est apparue dans la recherche par simulation — ce n'est pas vraiment une innovation, mais plutôt une ré-invention et une révision d'une technique utilisée depuis longtemps — tient à l'usage abondant des protocoles (ici, des verbalisations de gens en train de résoudre des problèmes) considérés comme l'une des sources de données empiriques pour tester le comportement du système. Il y a sans doute quelques questions en cours sur la manière d'interpréter les données de ces protocoles. Certains d'entre vous connaissent l'étude que Anders Ericsson et moi-même avons publiée sur ce sujet dans *Psychological Review* en 1980. Celle-ci a été développée depuis dans un livre intitulé «Protocol Analysis» (M.I.T. Press, 1984). Il reste encore quelques questions sérieuses sur la manière d'obtenir une bonne analyse de protocoles, mais je ne vois là aucune difficulté insurmontable...

On a posé, au cours de cette rencontre, quelques questions sur la manière dont on pouvait, à partir de telles données, établir une connaissance de la structure interne du processus considéré. Vous pouvez observer le comportement du système, en y incluant ses propres verbalisations, et vous pouvez les enregistrer et les analyser en tant que données ; mais comment pouvez-vous savoir ce qui se passe dans la tête, dès lors que vous ne disposez pas d'observations neuro-physiologiques directes de ce qui se passe effectivement dans la tête?

Il s'agit bien là, sous une forme déguisée, tout simplement du vieux problème de l'induction. Nous partons de quelque phénomène observable et nous postulons quelques structures et processus à l'aide desquels nous pouvons interpréter ces phénomènes, avec précaution mais efficacement ; et nous pouvons alors voir si les prédictions résultant de ces mécanismes hypothétiques correspondent en fait aux observations que nous pouvons faire. Nous ne sommes jamais certains, bien sûr, même si le comportement observé est conforme à la théorie, que d'autres mécanismes ne seraient pas susceptibles d'expliquer tout

aussi bien le phénomène.

Rarement cependant, dans les sciences du naturel, on regrette de disposer de 7 mécanismes explicatifs possibles, ou de 3 seulement, qui tous parviennent à expliquer avec succès le même phénomène. Il arrive même quelquefois, que nous n'en ayons que deux, tels le modèle matriciel de Heisenberg et le modèle ondulatoire de Schrödinger, qui semblent expliquer correctement un phénomène. (Comme ces deux derniers modèles s'avèrent mathématiquement équivalents, cela ne fait pas problème). Mais, d'une façon générale, en recherche scientifique, on ne dispose *d'aucune* hypothèse ou mécanisme susceptible d'expliquer le phénomène étudié : ils constituent justement un grand mystère. Si bien que lorsque vous disposez *d'une* bonne hypothèse — et même si vous êtes très conscients des problèmes théoriques de l'induction ; même si vous êtes très conscients que cette hypothèse qui explique vos données ne prouve nullement qu'il n'y en aurait pas d'autres qui les expliqueraient tout aussi bien, ni qu'elle continuera à les expliquer quand vous disposerez de nouvelles données —, alors, je vous l'assure, vous considérerez que vous détenez une forte bonne explication ; et vous la maintiendrez jusqu'à ce qu'elle ne marche plus ou jusqu'à ce que quelqu'un d'autre soit assez brillant pour en proposer une autre. Et si quelqu'un d'autre avance une autre hypothèse, vous essaieriez alors de monter une «expérience cruciale», ce qui est plutôt rare en science. Si bien que, la plupart du temps, nous sommes heureux de trouver ne serait-ce qu'une seule explication.

Que je vous raconte, à ce propos, une brève anecdote. En 1900, Planck découvrait la loi du rayonnement du corps noir, un des fondements de la théorie quantique, au demeurant fort mystérieuse alors puisque le mécanisme, qu'il construisit pour expliquer ce phénomène, était incompatible avec la physique classique. Si vous étudiez cette histoire en détail, vous apprenez que Planck découvrit en fait cette loi par un exercice d'ajustement d'une courbe. Il avait quelques données et il cherchait la fonction la plus simple qui s'ajustât à ces données. En pratique, on peut reproduire ce processus sans avoir des connaissances très poussées en mathématiques ; on peut, en tout cas, établir un scénario très plausible de la façon dont il s'y prit. Il le fit en un après-midi, aussitôt qu'il eut vu ses données. Puis, il passa près de 3 mois à rationaliser la loi, en imaginant des agencements hypothétiques de mécanismes physiques qui, s'ils étaient vrais, expliqueraient les données. Il n'avait aucune assurance qu'il n'existait pas beaucoup d'autres mécanismes qui auraient pu aussi expliquer ces données ; il n'avait même pas pensé qu'il pût y en avoir ; il n'avait pensé qu'à celui-là. Nous savons aujourd'hui que ce fut le bon!

Dans la pratique scientifique, vous n'abandonnez pas une théorie qui explique un phénomène ; et c'est ce que fit Planck en 1900. Incidemment, Kuhn a fait remarquer que Planck n'avait pas bien compris sa propre théorie; ce furent Einstein et Ehrenberg qui, en 1905, durent la lui expliquer, mais ceci est une autre histoire!

Ainsi donc, vous pouvez constater que, dans les simulations par ordinateur, nous mettons en oeuvre les procédures habituelles de la science. Les méthodes de simulation nous procurent non seulement des prévisions du comportement humain mais nous donnent aussi une théorie de l'organisation interne des processus de pensée, au niveau du traitement de l'information. Jusqu'à ce que quelqu'un vienne proposer un meilleur modèle des processus mentaux, qui soit aussi puissant pour expliquer les comportements effectifs, j'aurai bien du mal à me convaincre que je ne devrais pas m'en tenir à ce que je détiens — ce que certains voudraient appeler une métaphore; mais ce que j'appelle une théorie scientifique — sur la manière dont les choses se passent.

Les processus du développement humain

J'en viens à d'autres sujets. Il y a eu de nombreuses discussions sur les relations de l'apprentissage avec les modèles des processus de développement et, à leur tour, sur les relations de ces derniers avec les modèles piagétien du développement. Je n'ai pas grand chose à dire qui n'ait déjà été dit ici. La question reste ouverte, me semble-t-il, de savoir si ce sont les mêmes processus qui rendent compte de l'apprentissage à court terme et du développement à long terme; et je pense que chacun de nous pourrait élaborer un scénario plausible de l'une ou l'autre interprétations.

La prudence, (et la commodité) suggérerait que nous essayions de comprendre les processus du développement en recourant aux mêmes mécanismes que nous utilisons pour comprendre les processus d'apprentissage, bien que cela puisse être pourtant faux. Dans le cas des systèmes dynamiques, il est souvent vrai que les processus à long terme sont commandés par d'autres règles que les processus à court terme. C'est au fond la leçon principale de la théorie des systèmes quasi-décomposables³. Un même système peut sembler très différent selon que l'on observe seulement les changements à long terme ou qu'on ne s'en tient qu'aux modifications à court terme. Ces deux types de changements peuvent être gouvernés par des mécanismes totalement différents. Il s'agit là d'une question ouverte, qui le restera probablement longtemps, tant il est difficile d'obtenir des observations directes des processus de développement même si on peut disposer de toutes sortes d'indications indirectes. Bien entendu, c'est l'une des contributions majeures de l'Ecole genevoise que de nous avoir fourni des observations sur ce que le système semble être aux diverses étapes du processus de développement.

Emotions et motivation

On s'est demandé, lors des discussions, dans quelle mesure la théorie du traitement de l'information pouvait prendre en compte les émotions et motivations aussi bien que la cognition. Avec mon optimisme habituel, je pense que la voie est ouverte pour de tels développements en cette direction; mais, jusqu'à ce jour, je n'ai pas repéré beaucoup de progrès ni même beaucoup d'efforts en ce domaine. Je pourrais compter sur mes doigts le nombre de tentatives sérieuses visant à traiter l'émotion et la motivation dans le cadre général du traitement de l'information.

Cependant, en neuro-physiologie, il existe dorénavant bon nombre de connaissances et même de théories partiellement validées, relatives à l'émotion et la motivation, qui peuvent être incorporées à la théorie du système de traitement de l'information. Notamment l'idée que la fonction principale des émotions — du moins, des émotions du type surprise, colère, peur, anxiété, etc. — est d'interrompre l'attention et de déplacer celle-ci vers ce qui a suscité ces divers types d'émotion.

Si je peux me permettre une analogie un peu audacieuse, avec un système informatique fonctionnant en temps-partagé, je dirais qu'un tel système doit lui aussi prendre des décisions en fonction de ce qui se passe à l'entour et doit pouvoir être interrompu. On peut même concevoir un programme qui intégrerait des émotions, autrement dit des occasions d'interruption, dans un système en temps partagé qui aurait à faire face à des sollicitations en temps réel.

Je n'en dis pas plus sur cette question, sinon que j'ai écrit une étude à ce propos «Motivational and emotional controls of cognition» (publiée depuis dans «*Models of Thought*»⁴). C'est dans cette direction

³ NdT. Voir "L'architecture de la complexité" dans «La science des systèmes, science de l'artificiel» («The Sciences of the Artificial», 1962, 1969). Traduction française, 1974, Paris, Epi.

⁴ NdT. *Models of Thought*, vol. 1, , 1979, .yale University Press, , p.29-38.

que j'essayerais de chercher.

Les relations avec les Neuro-sciences

Je pense avoir déjà été assez clair sur ce sujet. Selon moi, la théorie du traitement de l'information n'est pas une théorie neuro-anatomique, pas plus que la chimie du 19^e siècle n'était une théorie de physique atomique.

Il est clair que l'une des tâches scientifiques de la prochaine génération (ou de la suivante) — outre de développer les avancées de la psychologie du traitement de l'information et les progrès en neuro-anatomie et neurophysiologie — consistera à établir un pont entre ces deux domaines. Cela se fera en son temps mais je ne crois pas que cela arrivera dans les cinq prochaines années.

Résolution de problème et formulation⁵ de problème

Il y a eu beaucoup de discussions sur les relations entre la résolution et la formulation des problèmes. Peut-on considérer la formulation comme une résolution de problèmes ? Lors de la formulation du problème, on doit, d'une façon ou d'une autre, aboutir à des buts ou à des critères d'évaluation des solutions qui ne sont pas nécessairement donnés à l'avance, tout en établissant une représentation du problème. Pour la psychologie du traitement de l'information, le défi ici est de montrer que ce sont fondamentalement les mêmes types de processus qui interviennent, d'une part, pour trouver un itinéraire entre une situation de problème et sa solution, et, d'autre part, pour trouver une voie menant d'une situation à la formulation d'un problème. On devra se demander comment on peut utiliser l'analyse «moyens-fins», par exemple, pour déterminer les objectifs ; ou quel rôle pourraient jouer les processus de re-cognition⁶ dans la formulation des buts ou dans l'élaboration de la représentation d'un problème.

Il y a peu de littérature là-dessus, à commencer par quelques textes exploratoires qui suggèrent comment cela pourrait être fait sans aller très loin dans les détails. Mon propre article «The structure of ill-structured problems»⁷, dans «*Models of Discovery*», est un essai dans cette direction.

Le domaine de l'architecture, sur lequel nous avons entendu beaucoup de choses ici, est un très joli terrain pour explorer les questions de la formulation de problèmes, car un problème architectural ne se pose pas habituellement de lui-même comme un problème bien structuré. Lorsque les gens commencèrent à considérer ce qu'ils allaient faire de la Grande-Motte, je suis sûr qu'ils n'avaient pas une vision très claire ni très tranchée de l'ensemble des buts et des contraintes. Une grande part du processus a consisté à développer cela. C'est pourquoi je veux espérer que l'on va bientôt se mettre à consacrer beaucoup d'efforts à étudier le processus architectural ainsi que les autres processus de conception complexe du

3. NdT : On propose de traduire «Problem-setting» par «formulation de problème», en suggérant d'entendre «formulation» au sens large : repérage, identification, caractérisation, mise en forme.

4. NdT. On propose de traduire — dans ces contextes — «recognition» par le néologisme «re-cognition» : on tire ainsi parti de l'ensemble de connotations attachées au mot «cognition» et on réduit les nombreuses connotations «parasites» du mot «reconnaissance».

⁷ NdT : «la structure des problèmes faiblement structurés» ou encore «la structure des problèmes mal définis» (1973, 1977).

même genre, afin de voir comment ils fonctionnent en fait. De façon plus spécifique, pouvons-nous utiliser dans ce cas les mêmes processus que ceux qui sont ordinairement mis en oeuvre dans la résolution du problème? Bien que certains aient suggéré comment cela pourrait être fait, de toute évidence cela n'a pas encore été fait. Lorsque je dis «suggéré», je pense à des travaux exploratoires tels que ceux de Charles Eastman et Omer Akin à Carnegie-Mellon⁸.

Sur une échelle beaucoup plus restreinte, c'est-à-dire pour des problèmes qui ne sont pas aussi ouverts que ceux de la conception architecturale, je vous renvoie aux chapitres de «*Models of thought*» sur le programme appelé UNDERSTAND, un programme qui part d'une description verbale d'une tâche et qui en établit une représentation, y compris une représentation du but. Il est vrai que le but est donné sous forme verbale : le programme ne se demande pas lui-même de créer ce but, mais en revanche il se charge lui-même de tous les autres aspects de la représentation.

Une des raisons de mon optimisme quant à la possibilité de modéliser la résolution de problèmes faiblement structurés tient tout simplement au fait que, d'un point de vue positiviste, la formalisation d'un but revient à la formalisation de buts selon des critères dont on dispose déjà, fût-ce implicitement. Je ne vois pas ce que veut dire «formuler un but» si l'on n'a pas déjà un but, qui n'est certes pas ce but précis mais quelque ensemble plus large ou plus fondamental de critères.

Ainsi, la *formulation* des buts ressemble en fait beaucoup plus à une évocation des buts. Lorsque l'on considère le problème de la conception de la Grande-Motte, par exemple, on ne prétend pas découvrir subitement un nouvel ensemble de valeurs humaines qu'il s'agit d'atteindre ; mais on évoque des buts possibles : «Oh, si nous faisons ici un parc industriel ? Et si nous faisons là un village de loisirs ?» ... Il ne s'agit pas en fait de buts nouveaux : les gens y ont pensé bien avant et ces buts ne sont sans doute pas sans liens avec les désirs humains les plus fondamentaux et avec les images que les êtres humains se forgent sur leur manière de vivre.

Aussi, me semble-t-il, devrions-nous parler d'évocation et de définition de buts plutôt que de formulation de buts issue d'un auteur indifférent. Comme le disait Archimède : pour soulever le monde, il suffit d'un levier et surtout d'un point d'appui pour ce dernier. De même, pour penser à des buts, il faut déjà disposer de buts à partir desquels on puisse penser.

J'ai été très intéressé par les commentaires de M. Quilici sur les brevets. Le domaine des brevets est particulièrement intéressant pour l'exploration des processus de l'invention et de la création. Des recherches ont été réalisées sur l'histoire des découvertes scientifiques mais, à ma connaissance, peu de travaux récents se servent des brevets comme matériaux de recherche. Une des raisons pour lesquelles je pense qu'ils sont de bons matériaux tient au fait qu'ils sont très bien documentés. Vous ne pouvez pas obtenir un brevet si vous ne fournissez pas une documentation soignée sur ce que vous avez fait, bien que vous ne soyez pas obligé d'indiquer le processus que vous avez réellement suivi

Management et organisation

J'en viens maintenant à la gestion et à l'organisation. J'ai été très intéressé par la contribution de M. Demailly et sa description des modalités d'analyse des organisations en termes de structures de décision et de «patterns» de communication. Il y a là une approche qui peut sembler évidente, en un certain sens, mais que fort peu de gens pratiquent. M. Demailly a décrit quelques démarches pour ce faire et dont il est évident qu'elles aient eu un grand succès ; et je me disais que j'y voyais là un exemple qui allait connaître

⁸ NdT : au sein du «Design Research Center» de l'Université Carnegie-Mellon (Pittsburgh).

beaucoup d'imitations.

La démocratie industrielle

M. Marchesnay s'est intéressé à ma théorisation de la démocratie industrielle et spécialement à mon scepticisme quant à la démocratie industrielle. Il s'agit là de problèmes essentiels de politique sociale et d'organisation sociale, auxquels je n'essaierai pas de répondre en une minute. Mais j'aimerais clarifier ma position sur ce point, ne serait-ce que pour n'être pas accusé de sentiments anti-démocratiques.

Je pense qu'il y a là quelques sérieuses questions de définitions sur lesquelles je n'avais sans doute pas été assez clair dans mon article, dès lors que le propos de celui-ci était assez polémique. A l'origine, cet article était écrit pour un auditoire de syndicalistes suédois. J'argumentais contre une conception qui me semble très largement acceptée aujourd'hui dans les démocraties occidentales — y compris mon pays et une grande partie de l'Europe — et qui repose sur la conviction que l'essence de la démocratie est fondée sur l'égalité numérique de l'influence de chacun. Bon nombre des modes de scrutin dont nous disposons — et la représentation à la «proportionnelle» en est un — reposent sur cette idée d'égalité arithmétique. Si vous examinez les arguments qui étayaient de tels modes de scrutin, et en fait tous les modes de scrutin, vous constatez qu'ils reposent tous sur l'idée que le problème fondamental de la démocratie est d'obtenir quelque forme d'égalité numérique précise à propos de l'influence que chacun peut exercer sur toute chose.

C'est sur cette question que je revenais dans mon article, en réexaminant les comptes-rendus de la convention qui élaborait la Constitution des Etats-Unis, convention composée d'un groupe d'hommes plutôt sages. Pour le moins, la structure qu'ils conçurent a démontré qu'elle était viable pour longtemps. En étudiant leurs délibérations, je m'aperçus qu'ils ne pensaient pas du tout à l'égalité numérique. Ce qu'ils avaient à l'esprit était la nécessité de prévenir, autant que possible, toute grande accumulation de pouvoir où que ce soit dans la société et, en particulier, des pouvoirs si forts qu'ils ne pourraient être équilibrés par aucun autre contre-pouvoir. Les rédacteurs de la Constitution n'avaient pas la notion d'égalité numérique ou de démocratie arithmétique : ils recherchaient seulement des moyens pratiques de prévenir la centralisation du pouvoir.

L'argument de mon papier était que la solution du problème du contrôle du pouvoir peut être très différente selon qu'on parle d'un Etat national — parfois même d'une communauté civique pouvant fixer des impôts ou disposant d'autres attributions de ce type — ou selon que l'on parle d'une organisation industrielle et commerciale ou même d'une organisation universitaire. Pourquoi ces solutions pourraient-elles être très différentes ? Parce qu'une organisation universitaire ou une organisation d'affaires, à la différence des organisations d'Etat, n'ont pas les mêmes occasions d'utiliser un pouvoir effréné dans leurs relations avec la société ou même vis-à-vis de leurs propres membres. L'Etat peut vous mettre en prison. Votre employeur, peut, dans une certaine mesure, vous rendre la vie difficile mais il ne peut pas vous mettre en prison. L'Etat, lorsqu'il veut mobiliser les ressources de la société, peut vous faire payer des impôts ; mais une entreprise peut seulement faire appel aux ressources de la société en produisant et vendant des biens que celle-ci semble réclamer.

Ainsi sur ce point, l'administration de la preuve revient à ceux qui pensent que les procédures de scrutin ou les participations par scrutin (participations à des commissions ou autres choses du même genre) jouent un rôle aussi central, au sein d'une Université ou d'une entreprise, pour une répartition adéquate du pouvoir et une prise en compte pertinente des besoins et désirs humains, que les mêmes procédures qui jouent au niveau de l'Etat. C'était là un véritable message.

Qui a besoin du vote des étudiants pour décider quels devraient être les professeurs à conserver dans

une Université ? Il existe d'autres moyens d'associer les étudiants à l'action, même s'ils sont restreints. Telles étaient mes excuses pour mon attaque de la démocratie industrielle. Sur d'autres points, je suis un bon Américain de l'aile gauche du parti démocrate ; mais sur celui-ci, je suis réactionnaire...

Les ordinateurs et la productivité

Quelqu'un a demandé pourquoi les ordinateurs n'avaient pas augmenté la productivité depuis 25 ans. S'il avait demandé pourquoi ils n'avaient pas augmenté cette productivité d'une manière formidable, je serais d'accord avec lui ; ce n'est pas ce qu'il a fait. Je ne sais pas si les ordinateurs ne l'ont pas augmentée pour tout le monde, mais bien des entreprises semblent fort heureuses d'avoir converti toutes sortes d'opérations en procédures informatisées. Ce qui m'incite à penser qu'elles ont dû économiser un peu d'argent dans l'affaire, ce qui démontre à l'évidence un gain de productivité.

Mais l'ordinateur n'a sûrement pas eu un effet drastique. Selon moi, une raison majeure de cela tient au fait que le taux d'impact de l'ordinateur sur la société dépend du taux des investissements en capital-ordinateur. Ces investissements ont été modestes par rapport aux investissements en capital dans beaucoup d'autres domaines. La croissance de la productivité, si l'on met de côté les effets du choc pétrolier et d'autres problèmes des dix dernières années, a été à peu près la même que celle que l'on observait avant l'introduction des ordinateurs.

C'est une des raisons pour lesquelles je ne vois personnellement aucun miracle dans la présente révolution informatique, même si on y inclut les systèmes-experts. On ne pourra réaliser des gains importants qu'à la condition de consentir de gros investissements en capital-ordinateurs, au détriment d'autres formes d'investissements. En attendant, je prévois que la productivité continuera de croître à un taux des plus modestes.

On s'est demandé, plus particulièrement, si ces retards en croissance de productivité n'étaient pas dus à l'inadéquation de nos organisations hiérarchisées. Sur ce point, je ne pourrai que dire que, dans le passé, les organisations ont eu surtout à faire face à la rareté de l'information appropriée, des informations dont les gens avaient besoin pour prendre leurs décisions et que, dans ces dernières années, tout a formidablement basculé du fait de l'ordinateur et des moyens de communication.

Le problème des organisations, aujourd'hui, est de préserver l'attention humaine afin qu'elle ne se perde pas dans des informations triviales ; mon exemple favori est celui du journal quotidien (j'espère que tous ceux qui sont dans cette salle ont renoncé à lire leur quotidien ou regarder la télévision pendant cette semaine à la Grande-Motte !). Nous avons à repenser nos organisations en termes de gestion de l'attention, dès lors que l'attention humaine est une ressource bien plus rare que l'information. Quand nous essayons d'améliorer des systèmes de décision ou des systèmes de communication, le plus difficile à faire n'est pas de «déboucher» des canaux d'information mais d'y insérer des filtres intelligents, il s'agit là de l'un des domaines où l'intelligence artificielle peut jouer un rôle éminent à l'avenir.

Science fondamentale et science appliquée

On s'est posé des questions sur les applications pratiques de l'intelligence artificielle et ma dernière remarque constitue un exemple de ce que peuvent être de très importantes applications, selon moi. En aucun moment de ma vie, je ne me suis senti écartelé entre la science pure et la science appliquée. Si nous

prenons le temps de regarder le monde réel des phénomènes, ce monde est une source extraordinaire de problèmes scientifiques «purs» et fondamentaux.

Mon premier problème scientifique, qui traitait de la théorie de la décision humaine, fut suscité par l'examen du commissariat de police des «Great-Plains», Wichita, Kansas. Patrouiller, pendant une semaine dans les voitures de police de cette ville, me valut quelque très bons sujets de recherche qui, au fond et dans un certain sens, sont encore ceux qui m'occupent pour le reste de ma vie. De bons problèmes de recherche fondamentale sont ainsi nés des préoccupations très pratiques d'un service de police. En essayant d'améliorer les opérations de police, on parvient à comprendre bien des phénomènes fondamentaux de l'administration et du comportement humain.

Par ailleurs, je pense que l'on peut aussi affirmer que «comprendre» consiste précisément à savoir comment améliorer. C'est pour cela que nous développons les sciences fondamentales parce que nous pensons que si nous comprenons les systèmes de traitement de l'information, les applications suivront. Elles ne suivront pas sans efforts — des travaux considérables de développement et d'application doivent être poursuivis — mais la recherche fondamentale a déjà donné et continuera de donner des repères essentiels pour la solution de problèmes pratiques...

L'économique

Nous avons eu ici beaucoup de discussions sur la science économique. Aussi, me limiterai-je à quelques commentaires.

La rationalité limitée

Observons d'abord que l'approche de la rationalité limitée, celle du «satisfecum»⁹, trouve ses applications les plus immédiates en micro-économie. Il y a une raison à cela, c'est que si vous êtes disposé à formuler des hypothèses globales à partir d'une rationalité non-limitée, vous pouvez en tirer de nombreuses conclusions qui vous conduiront quasi-directement à des modèles macro-économiques. Pour faire cela, vous n'avez pas vraiment besoin de quitter votre fauteuil, puisque cet exercice ne demande pas des masses de travail empirique!

La difficulté avec la théorie de la rationalité limitée, et c'est ce qui explique principalement qu'elle n'attire pas beaucoup les grands économistes, est qu'il s'agit d'une théorie «faible» en ce sens qu'elle ouvre un très grand nombre de possibilités. Si les gens ne savent pas comment est effectivement fait le monde dans lequel ils vivent, ils peuvent imaginer dès lors beaucoup de mondes possibles et ce seront les limites de leurs capacités de computation et d'information qui détermineront principalement le choix, parmi ces mondes qu'ils imaginent, de celui qu'ils tiennent pour réel ; ou encore le choix, parmi ces mondes possibles, de celui dont ils vont se servir pour leurs computations.

Ainsi, avant même que vous puissiez commencer à utiliser la théorie de la rationalité limitée à des fins de théorisation économique, vous devez trouver des éléments sur les processus effectifs auxquels recourent en effet les gens pour simplifier leur représentation du monde afin d'agir sur lui. Et, pour faire

⁹ NdT : la traduction de «bounded rationality» par «rationalité limitée» semble aujourd'hui généralement acceptée. En revanche, celle de «satisficing» (expression forgée par H.A. Simon, différente de «satisfying» qui correspond à «satisfaisant») fait encore problème. Pour traduire cette expression, on propose de retenir volontiers le néologisme «satisfecum», initialement introduit par Daniel Carré et J.L. Le Moigne (cf. «Auto-organisation de l'Entreprise», Paris, 1977).

cela, vous devez entreprendre des quantités de recherche empirique. La différence la plus fondamentale entre la théorie du «satisfecum» et les autres versions théoriques de la rationalité me semble tenir aux différences quant aux demandes qu'elle requiert en matière de travail empirique.

L'un de vous s'est interrogé sur les relations de la théorie de la rationalité limitée avec l'œuvre de Keynes. Si vous étudiez Keynes puis son principal concurrent du moment, la théorie des anticipations rationnelles, vous vous apercevez que ces théories sont vraiment très semblables. L'une et l'autre acceptent l'idée d'un homme rationnel au comportement maximisateur. L'œuvre de Keynes est remplie de cette idée d'homme qui maximise de même que celle de Lucas, bien entendu. L'ennui est que l'on ne peut établir des cycles économiques avec des gens qui seraient constamment rationnels, au sens de la maximisation classique des utilités subjectives espérées.

Comment Keynes et Lucas obtiennent-ils quand même des cycles économiques? En insérant une petite graine d'irrationalité dans leurs systèmes.

C'est comme les perles de culture : vous trouvez une huître et vous y placez un petit grain de sable et, peu après, vous obtenez une jolie perle. En ce qui le concerne, Keynes place son grain de sable dans l'équation du marché du travail. Les travailleurs, dit-il, ne font pas la différence entre les salaires nominaux (ou monétaires) et les salaires réels. Une fois que vous avez greffé cette imperfection sur la rationalité, toute la théorie des cycles économiques de Keynes suit (je sur-simplifie car Keynes faisait aussi une hypothèse non-classique sur les taux d'intérêt).

Où Lucas place-t-il son grain de sable ? il le place dans le comportement des entrepreneurs. Les entrepreneurs seraient rationnels mais pas *parfaitement* rationnels, car ils seraient trop ignorants pour connaître la différence entre les mouvements généraux des prix dans la société et les mouvements divergents des prix dans leur propre industrie. C'est une autre sorte d'illusion monétaire, d'où découlent les cycles économiques de Lucas.

Si bien que la différence entre les théories keynésiennes et les théories des espérances rationnelles ne tient pas au fait que les unes acceptent l'homme rationnel classique et les autres pas. Toutes deux acceptent l'homme classique jusqu'à un certain point, puis elles insèrent leur grain de sable pour obtenir de l'imperfection dans la rationalité. Mais elles ne reviennent jamais, sauf cas fortuits, vérifier où est vraiment logé le grain de sable. Elles ne démontrent jamais si les salariés souffrent réellement d'une illusion monétaire, en procédant à quelques enquêtes sur les travailleurs, ou si les entrepreneurs souffrent de leur propre forme d'illusion monétaire. Toutes deux n'ont jamais entrepris de travaux empiriques en micro-économie, plutôt que de débattre entre elles dans les tribunes et de donner des avis contradictoires aux présidents. C'est là que la théorie de la rationalité limitée adopte un point de vue différent. Elle insiste sur les études empiriques pour établir les limites effectives de la rationalité.

Rationalité limitée et théorie marxiste

Je voudrais dire un mot de la relation entre la théorie de la rationalité limitée (qui est le type de théorie que j'ai épousé) et les théories marxistes. J'aurais à faire une réserve sur les remarques qui ont été faites à ce propos : quelqu'un a suggéré que peut-être une différence existait entre le long terme (Marx) et le court terme (la rationalité limitée). Je me permets d'attirer l'attention sur le dernier chapitre de «*The New Science of Management Decision*»¹⁰ qui propose un modèle de la dynamique à long terme des effets

¹⁰ NdT : traduit en français sous le titre *Le nouveau management la décision par les ordinateurs*. Ce chapitre IV de l'édition 1977/1980 est intitulé «Les effets économiques de la Technologie de l'information». 69, 99-118. Article repris dans "Models of Thought".

de l'automatisation. Vous y verrez que l'analyse est très non-marxiste sous certains aspects (bien que marxiste pour d'autres), en particulier dans son interprétation de l'histoire et de quelques hypothèses fondamentales qui subsistent sur l'histoire.

Dans mon modèle, il n'y a pas de surplus de capital et le taux de formation du capital est juste en ligne avec le taux de croissance de la productivité, si bien que le capital par unité de production s'avère constant, ce qui est historiquement vérifié. Partant de cette donnée historique, prise pour ce qu'elle vaut, le modèle s'élabore en fonction de cette hypothèse et aboutit à la conclusion que le système ne s'effondre pas dans un surplus de surproduction ; et que le travail, loin d'être exploité, accroît progressivement sa part dans le revenu, ce qui fut le cas aux Etats-Unis dans les cent dernières années. Mais je présume, M. Munier, que vous avez quelques doutes sur la transposition de ces mêmes faits à la France, à moins que nous puissions disposer de quelques données empiriques à ce sujet. Mais je voulais seulement mettre en exergue ce chapitre, qui pourrait clarifier les relations, s'il y en a, entre mes vues et celles de Marx.

L'incertitude

On s'est demandé comment l'on traiterait de l'incertitude dans la théorie du «satisfecum». Une des hypothèses qui me semble intéressante repose sur le fait que les gens ne calculent pas souvent les distributions en probabilité des événements. Ils peuvent le faire lorsqu'ils négocient une assurance-vie, mais ce sont en réalité les compagnies d'assurances qui établissent ces probabilités et les gens qui s'assurent ne le font pas.

Les gens, face à l'incertitude (celle qu'ils veulent bien prendre en considération), suivent, la plupart du temps, une stratégie de type minimax, laquelle peut être présentée en termes de «satisfecum» de la façon suivante. S'il y a un grand nombre d'issues possibles et si l'on ne sait pas ce qui va se passer, si ce n'est que toutes les occurrences sont possibles, alors il s'ensuit une stratégie qui ne conduise pas au désastre quoi qu'il arrive. Face à l'incertitude, le «satisfecum» évite le désastre.

C'est ce que montrait, de façon un peu plus formalisée, mon article de 1955 «*A behavioral model of rational choice*»¹¹, la première de mes tentatives pour formaliser un tel modèle économétrique. Par exemple, appliquée au cas de l'assurance vie, la politique du «satisfecum» minimax, pour un jeune homme ayant des enfants mais pas de propriété, est de déterminer un montant de sa police d'assurance qui ne pèse pas trop sur sa famille, selon qu'il meure ou reste en vie. Cette assurance-vie ne devrait pas être un désastre, s'il restait en vie ; elle ne devrait pas être un désastre, s'il venait à mourir.

*Le satisfecum et la théorie de la quête ou de l'exploration*¹².

Quelques questions ont été soulevées à propos des relations entre la théorie du «satisfecum» et la théorie de l'exploration ("search") de Stigler ou d'autres théories contemporaines et néo-classiques de la quête. On se rappellera que, pour l'économie néo-classique, l'une des façons de traiter de l'incertitude est

¹¹ NdT : «un modèle comportemental du choix rationnel», *Quarterly Journal of Economics*, 1955, 69, 99-118. Article repris dans "Models of Thought".

¹² NdT : on devrait traduire «search theory» par «théorie de la recherche», ce qui conduirait à une confusion avec «research theory». Le mot «search» doit être entendu ici dans le sens de l'enquête, de l'examen, de l'investigation ou de l'exploration. En français, «exploration» est peut être trop fort ; «quête» paraît plus neutre. (Ph. Mongin, dans son étude (cf. p. 538), s'est résigné au français et traduit «search» par search !...).

de dire : «Ah, oui, il y a l'incertitude et les gens vont chercher à la réduire jusqu'au point où la valeur marginale qu'ils peuvent espérer d'une quête supplémentaire d'information sera égale au coût marginal de cette quête d'information». Ils font donc du «satisfecum», en s'arrêtant à ce point et non au point optimum. Mais ce point est aussi un optimum, si l'on prend en compte le coût de la quête.

Et cela n'est-il pas du «satisfecum» ? Non, ce n'est pas du tout du «satisfecum». Pourquoi ? Parce que, loin d'avoir simplifié le problème de la computation, loin d'avoir pris en compte la rationalité limitée des êtres humains, les néo-classiques ont tout simplement fait exploser le problème de la computation. Désormais, les décideurs devront non seulement connaître les distributions en probabilité des alternatives et de leurs conséquences, mais devront aussi connaître la productivité marginale de la quête d'information. Si vous construisez un modèle de ce genre de quête, avec ses aspects computationnels et informationnels, ce modèle sera toujours plus compliqué que le modèle classique d'optimisation en situation certaine.

L'argument fondamental, et la clé qui distingue mon interprétation de la rationalité limitée de toutes ces excroissances de la théorie néo-classique, est que la théorie de la rationalité limitée tient compte, avec le plus grand sérieux, des limites des capacités humaines de traitement de l'information. Les gens ne pratiquent pas volontiers des computations sophistiquées. Et il s'avère que, si vous vous déclarez satisfait d'un «assez bon» résultat, vous pouvez vous économiser des quantités considérables de computation.

Certains d'entre vous connaissent déjà mon exemple des aiguilles dans le tas de foin. Quelqu'un a jeté au hasard quelques aiguilles dans un tas de foin ; certaines d'entre elles sont plus pointues que les autres. Vous venez de perdre un bouton à votre veste et devez le recoudre ; il vous faut donc chercher une aiguille dans le tas de foin.

Je vous donne alors le choix entre deux méthodes d'exploration. En exploration optimisante, vous cherchez la meilleure aiguille dans le tas de foin. En quête «satisfecum», vous vous contentez de chercher une aiguille, quelle qu'elle soit, qui vous permette de recoudre votre bouton. En conséquence, si vous cherchez la meilleure aiguille dans ce tas de foin, la quantité de quête nécessaire croît linéairement avec la taille de ce tas. Comme le monde est un énorme tas de foin, il s'ensuit que la quantité de quête ne peut être qu'énorme.

Si, en revanche, vous vous contentez de retenir la première aiguille qui fera l'affaire, l'exploration ne dépendra que de la densité des aiguilles dans le tas de foin et pas du tout de la taille de celui-ci.

Dans sa forme la plus simple, la théorie de la rationalité limitée est une théorie du «comment vivre» dans un monde infini, tout en ne disposant que de très modestes moyens de computation ; moyens qui ne dépendent pas de la taille du monde réel, mais seulement de votre environnement local et de ce que vous pouvez y faire. Et je crois qu'il s'agit-là d'une différence vraiment fondamentale.

L'intelligence artificielle

Je n'ai pas encore beaucoup parlé de l'intelligence Artificielle, ni de la science des systèmes, qui sont en quelque sorte le parapluie, ou le toit, sous lesquels nous sommes assis ici. L'une des raisons pour lesquelles je n'en ai pas encore parlé réside en ce que les contributions qui ont été faites sur l'intelligence artificielle, celles de MM. Pitrat et Laurière, en font un très bon examen et fournissent d'excellents exemples de ce qu'est ce domaine. Si bien que je puis me consacrer à quelques questions plus spécifiques.

Intelligence Artificielle et Cybernétique

M. Dupuy a posé une question, à laquelle j'ai brièvement répondu hier soir¹³, sur les relations entre

cybernétiques et intelligence artificielle. il est certain que l'intelligence artificielle s'est développée à partir des foisonnements de la cybernétique, autour des années 50. Je pense que ce qui les a séparées fit des différences d'accentuation, notamment sur deux points : l'idée du traitement de l'information symbolique et l'utilisation de l'architecture de Von Neumann pour la conception des ordinateurs. Ces deux points ont contribué à séparer ces deux domaines, sans hostilité ni conflit, mais parce que l'un et l'autre se consacraient à des tâches différentes. Depuis, ces deux domaines ont connu des histoires différentes.

Les aspirations « maximalistes » de l'Intelligence Artificielle

Par ailleurs, M. Andler semblait plutôt malheureux, hier, à propos de la position qu'il a appelée «maximaliste», selon laquelle les ordinateurs pourraient exécuter toutes les opérations que les gens effectuent avec leur tête. Comme je l'ai dit hier soir, ma position est certainement «maximaliste», et en dépit des reproches de M. Andler et de ses instructions quant à la manière de me réformer, je crois très improbable que je me conforme à ses directives.

Il ne m'a pas encore persuadé. Je suis sûr qu'il ne renoncera pas à sa position, pas plus que n'y renonceront d'autres adeptes tels qu Dreyfus et Weitzenbaum ; en fait, ils sont en train de faire carrière dans cette tâche de missionnaire.

Je m'étonne que cette discussion ait atteint une telle intensité, un tel niveau polémique. Quand une discussion devient polémique, on est alors tenté d'en venir à des arguments «ad hominem». Ce qui détourne de l'essentiel de la discussion et me conduit à me demander pourquoi il en va ainsi. Et j'ai une hypothèse à ce propos — ce n'est qu'une hypothèse — qui a néanmoins quelque fondement historique.

Je pense qu'il existe, dans le monde, des idées fortes qui s'avèrent très proches des images que nous avons de nous-mêmes — des représentations fondamentales que nous nous faisons de la condition humaine et de la nature humaine. L'idée de Copernic, selon laquelle la Terre n'était pas le centre de l'univers mais qu'elle dansait autour du soleil, choqua une idée qui était particulièrement centrale dans la théologie de cette période et entraîna des réactions plutôt vives de la part des autorités, entre autres. (Fort heureusement, les autorités en place n'ont pas encore donné leur verdict sur l'intelligence artificielle et personne n'a encore été jeté en prison).

On a retrouvé le même phénomène à propos de la révolution darwinienne. Darwin n'a pas simplement proposé une nouvelle doctrine scientifique ; il avançait l'idée que la place de l'homme dans le cosmos était très différente de celle qui lui était faite jusqu'alors selon les prescriptions des croyances religieuses orthodoxes de l'Occident. Et il y eut des polémiques hautes en couleur autour des problèmes du Darwinisme. Le fait qu'une polémique atteigne ce niveau d'intensité ne signifie pas que les arguments de l'une ou l'autre parties soient faux, ou que ceux de l'une soient meilleurs que ceux de l'autre. Je me borne à évoquer le ton de la discussion.

Je crois — et je ne pense pas que ceci surprendra Dreyfus — qu'une large part de la forte température, qui échauffe les discussions sur l'intelligence artificielle, provient de ce que celle-ci défie la thèse du caractère unique de l'homme dans le Cosmos. Ainsi, vous avez abandonné l'idée que l'on est unique parce que nous sommes au centre de l'univers ; vous avez renoncé à l'idée que l'on est unique parce que nous serions une espèce spécialement créée; mais, selon vous, nous serions encore la seule espèce capable de penser — enfin pas tout à fait (les chiens et les chats pensent un petit peu) — disons que nous serions la seule espèce capable de penser de grandes choses. Et vous vous retrouvez devant un nouveau défi : il se peut que, dans des circonstances appropriées, une autre espèce, appelée ordinateurs,

¹³ NdT : cf. «Quelques remarques historiques sur la science de la cognition», par H.A. Simon.

soit en mesure de penser de grandes choses.

Ce défi a de quoi être très inconfortable. Evidemment, il le sera d'autant plus que nous nous limiterions à considérer le caractère unique de l'homme comme la seule cause de notre valeur dans le monde. Il peut y avoir d'autres points de vue, par exemple des points de vue panthéistes, assez populaires de nos jours, dès lors qu'ils sont en rapport avec l'environnementalisme. D'après cette façon de voir environnementaliste, l'intéressant pour nous n'est pas que nous soyons uniques au plan cosmique ; c'est que nous faisons partie d'une énorme structure très élaborée et que nous participons à son élaboration. Ce qui est vraiment étonnant dès lors, ce ne sont plus les mystères, notamment celui qui ferait que notre esprit serait impénétrable. Ce qui est vraiment étonnant, dans tout cela, serait que tous ces systèmes compliqués — dont nous-mêmes ainsi que tout ce qui nous entoure, y compris les cieux — seraient intrinsèquement organisés et obéiraient à des règles de fonctionnement. Ce faisant, il devient tout aussi passionnant de rechercher l'organisation des processus de pensée de l'homme que de découvrir l'organisation du mouvement des planètes autour du soleil. Il s'agit là d'une autre approche possible de l'univers de l'homme, qui nous permet d'envisager calmement les perspectives de l'intelligence Artificielle.

Le développement de l'Intelligence Artificielle

Il s'est forgé une histoire mythique de l'intelligence artificielle. Je pense que la source originelle d'une bonne partie de cette histoire mythique fut le livre de Dreyfus ainsi que les essais qu'il publia antérieurement et furent repris par d'autres auteurs, notamment le célèbre rapport de Lighthill.

Lighthill était un physicien qui se sentit, du jour au lendemain investi d'une «vision» sur l'intelligence artificielle en Angleterre et y provoqua une polémique qui aboutit pratiquement à la destruction du Laboratoire d'intelligence artificielle d'Edinburgh. Et l'on trouverait ailleurs d'autres traces de cette histoire mythique. Quel est ce mythe ?

Il repose d'abord sur l'idée que, dans les premiers jours de l'intelligence Artificielle, les gens tinrent des propos excessifs en sa faveur ; et qu'au bout d'un certain moment, ils proclamèrent leur déception, ce qui entraîna un cycle d'enthousiasme puis de désappointement.

Je pense que c'est mythique. Tout au long de l'histoire de l'intelligence artificielle, comme au long de l'histoire de l'astrophysique, il en est ainsi : non seulement les chercheurs avancent des démonstrations, mais ils formulent aussi des spéculations. Il y a les théories des états stationnaires (steady-state) du cosmos et il y a les théories du «big-bang» pour ce même cosmos (ces dernières sont plutôt populaires, de nos jours). Il y a eu toutes sortes de théories du cosmos. Mais personne n'accuse les astrophysiciens de pousser des clameurs excessives ; on dit qu'ils spéculent. C'est une grande chose que de spéculer. C'est une des choses dont l'esprit humain tire grand plaisir. En 1958, dans la revue *Operations Research*, Alan Newell et moi publiâmes quelques spéculations : elles ne furent pas toutes fausses ni toutes exactes ; Aucune d'elles ne fut extravagante.

Venons-en à la déception qui aurait suivi des ambitions excessives. C'est un mythe que de parler d'un arrêt des recherches en Intelligence Artificielle dans les années 60. Je ne pense pas qu'on puisse étayer cela sur des bases empiriques. De plus, s'il y a eu une pause, ce n'en est certainement plus une maintenant. L'Intelligence Artificielle a été, au sens plein de cette expression, un paradigme scientifique en progrès, dans lequel nous en venons à connaître, d'année en année, comment faire des choses et construire des systèmes que nous ne savions pas faire et construire l'année précédente. De nos jours, comme cela a été illustré dans l'exposé de M. Laurière et de quelques autres, nous savons comment construire des systèmes-experts — c'est-à-dire, des systèmes qui, dans certains domaines, sont capables de performances proches d'un être humain, qu'il soit expert ou professionnel. Mais vous me direz «cela ne va pas durer, il y aura d'autres déceptions». D'accord, c'est possible, mais pourquoi devrions-nous attendre et voir ? Plutôt que d'attendre et voir, pourquoi pas passer à l'action et aller de l'avant dans la recherche?

Pour l'instant, rien ne pèse sur la prévision de ce que sera l'objectif final ; mais, il est amusant de l'envisager et je suis tout à fait désireux de l'atteindre. Je ne vois aucune limite. Je ne vois aucun profond mystère pesant sur la personne humaine à propos du traitement de l'information, mystère qui établirait les limites des explications de l'esprit humain en termes de traitement de l'information. En disant cela, je ne veux pas dire qu'il n'y a pas maintes recherches à faire avant que nous soyons capables d'expliquer une gamme plus vaste de systèmes.

Mais rien n'assure vraiment que j'ai raison ou tort à ce propos. Personne ne va pouvoir manger, à présent, ou devoir jeûner, avoir un toit ou être sans toit, ou disposer d'autres éléments pratiques de la vie, en fonction de cette prévision. Ce que nous avons à faire est d'en entreprendre les étapes suivantes et de voir jusqu'où nous pouvons mener ce paradigme. C'est aux études empiriques de décider cela, non aux polémiques ou aux débats. Alors, il faut mettre le paquet sur les études empiriques.

Entreprenons aussi des discussions, car le progrès scientifique a besoin d'être guidé. Nous avons besoin de décider si nous nous engageons dans telle ou telle direction. Alors, engageons la discussion, mais menons-la de manière civile. Pourquoi devrions-nous contester les motivations d'autrui ? Pourquoi devrions-nous contester l'intelligence d'autrui dans nos débats ? Aussi, ma seule suggestion à M. Andler serait (j'ai fait auparavant cette même suggestion à Dreyfus, dans une lettre ouverte que Pamela McCorduck a reproduite dans son livre sur l'histoire de l'Intelligence Artificielle) gardez votre sang froid («*cool it*») !

Conclusion

Je crois que j'ai parlé assez longtemps. Permettez-moi un paragraphe de conclusion. J'ai lu quelque part qu'un optimiste est quelqu'un qui commence à remettre ses souliers quand il entend le conférencier dire «et finalement», mais je m'arrêterai réellement dans un ou deux paragraphes.

Dans l'ensemble des questions dont nous avons parlé durant ces quelques jours, comme dans l'ensemble des questions sur lesquelles j'ai travaillé durant ma carrière scientifique, il y a quelques thèmes centraux qui semblent les agencer entre elles. Il y a l'idée d'une science de l'artificiel, que j'ai utilisée comme un concept organisateur dans le petit livre qui porte ce titre. Il y a l'idée d'une science des systèmes complexes, qui a été «l'étoile du berger»¹⁴ de la cybernétique durant ces derniers temps — la cybernétique se considère aussi comme une science des systèmes —. il y a l'idée des systèmes de traitement de l'information en tant que classe, centrale et capitale, de systèmes dans le monde. Il y a l'idée d'utiliser ces larges concepts pour illuminer la théorie en sciences sociales, en sciences économiques, en théorie des organisations et en sciences de gestion.

Je crois que ces idées constituent des heuristiques valables pour guider l'exploration et la pensée. Il serait présomptueux pour quiconque de prétendre que, au sens fort du mot «théorie», ces idées sont plus que des heuristiques, qu'elles représentent plus qu'un ensemble d'idées directrices allant nous permettre des va-et-vient entre ces champs spécifiques — emprunter des idées à l'un, les modifier puis les appliquer à l'autre. Et je pense que c'est en ce sens que, en dépit de la diversité des sujets que nous avons abordés au long de ces trois jours, et malgré peut-être la diversité de nos façons de parler, ce sont-là quelques thèmes unificateurs qui sont tout à fait centraux pour l'une des grandes directions des entreprises scientifiques de la société occidentale actuelle.

¹⁴ NdT : traduction de «guiding light».

NOUVEAUX DEBATS SUR CES COMMENTAIRES (et sur quelques autres) (texte établi à partir de l'enregistrement des débats de la séance de clôture du colloque de la Grande-Motte, 4 févr. 1984).

Je vais répondre brièvement à quelques questions qui m'ont été posées.

Question

Quelle est votre position sur le concept des «idées innées», par exemple le concept de Chomsky selon lequel les hommes ont quelque part en eux-mêmes, dès leur naissance, une théorie complète du langage ? Par ailleurs, «quelle est votre position sur le déterminisme»?

Réponse

Comme ces questions sont plutôt vastes, il se peut que je ne les couvre pas entièrement. Je supposerais que si nous avons des idées innées, ces idées doivent être construites dans les structures de notre système nerveux central. Elles doivent être transmises génétiquement par ces structures.

A un certain niveau d'abstraction, il semble clair que nous avons des idées innées. C'est-à-dire que nous disposons d'une machinerie pour penser qui nous permet, par l'expérience et l'interaction avec le monde, d'atteindre des niveaux humains de pensée qui sont différents de ceux d'un chimpanzé ou d'un chien ou même d'un dauphin. Il semble clair qu'il y a quelque chose de construit dans notre système nerveux qui nous permet d'interagir ainsi dans notre environnement, de devenir humain et d'acquérir ces capacités, y compris celles d'utiliser un langage et de penser sur des problèmes. En ce sens, je puis être très «chomskien» et je puis accepter le concept «d'idées innées».

Mais la théorie de Chomsky porte en elle une autre ambition : elle affirme qu'il existe quelque chose de spécifique à propos du langage et que ce trait spécial est construit dans l'homme, dès sa naissance. Je suis très sceptique sur cette assertion, car il me semble que ce dont on a besoin, pour acquérir le langage, c'est principalement un système nerveux capable de manipuler une information relationnelle symbolisée, autrement dit les mêmes capacités dont nous ferions usage pour interpréter une scène visuelle — pour décomposer cette scène en ses différents composants.

Je vois de nombreux objets humains dans cette salle, sans parler des tables ou autres objets de ce genre. L'aptitude à stocker en mémoire une représentation des aspects de cette scène me semble relever tout à fait du même type de capacité nécessaire au stockage de la structure profonde, ou même de la structure superficielle, des composants linguistiques.

Pourquoi dis-je ceci ? D'abord, pour une raison empirique : il est très facile de proposer aux gens des tâches, qu'ils effectuent facilement et ces tâches consistent à comparer des intrants visuels à des intrants linguistiques. Nous faisons cela tout le temps, mais Clark et Chase en ont fait une démonstration expérimentale très simple. Ils présentaient un stimulus contenant une étoile et une croix, ainsi qu'une phrase disant «l'étoile est au-dessus de la croix» ou «l'étoile est sous la croix». Les sujets devaient répondre "vrai" ou "faux". Pour répondre, les sujets devaient trouver un code commun à l'information présentée tant sous forme linguistique que visuelle. Je ne sais si ce codage commun s'apparente plus aux

éléments de langage ou aux éléments figuratifs ; et je ne pense pas que nous ayons à en décider maintenant. Ce qui est important, c'est qu'il existe un mode de codage commun, auquel on peut recourir pour le visuel comme pour le linguistique. Dès lors, la situation est entièrement symétrique. Il n'y a rien d'unique en ce qui concerne le langage.

Une autre forme de confirmation nous vient de l'intelligence artificielle. On a, dorénavant, construit plusieurs systèmes, dont l'un par Laurent Siklossy (lequel travaille actuellement à l'Université Libre d'Amsterdam), qui sont capables d'acquiescer un langage dès lors qu'on leur présente, d'une part, des éléments linguistiques et, d'autre part, des images schématiques de scènes qui décrivent les éléments linguistiques. A l'instar de l'expérience de Clark et Chase, l'expérimentateur présente, dans cette procédure, une image et une phrase. Ces programmes ont ensuite la capacité d'apprendre le langage en ce sens que, lorsqu'on leur présente une nouvelle scène en l'absence de toute phrase, ils peuvent construire une phrase qui décrit correctement la scène. C'est comme cela que les enfants font avec leur premier langage et, comme je l'ai dit, Siklossy a élaboré une telle procédure d'apprentissage ; John Anderson a expérimenté sur elle et on trouve un ou deux autres cas de ce genre dans la littérature.

C'est pour cette raison que je dis «oui, il y a des idées innées». Appelons-les «représentations sémantiques». Il s'agit essentiellement de l'aptitude à construire, en mémoire, des structures de listes et des structures relationnelles. Mais le langage utilise, lui aussi, cette aptitude générale. Et je pense qu'il ne se sert pas d'une représentation différente et à part. Sous cet angle, je ne suis pas «chomskien».

Certes, lorsque Chomsky dit «langage inné», il veut dire aptitude innée à manipuler un langage. En clair, l'enfant ne naît pas avec le «Chinois» inné ou le «Japonais» inné ou le «Français» inné ou «l'Anglais» inné. L'enfant devrait savoir où il est né pour faire cela. L'idée chomskienne est bien plus abstraite que cela. Je pense qu'il parle surtout de la machinerie sémantique, et dans ce cas, je ne crois pas que nous ayons sur ce point un réel différend.

Un point très important dont nous n'avons pas discuté relève de la question «maximaliste». Une des principales objections de Dreyfus porte sur l'affirmation selon laquelle nous pouvons simuler «l'homme dans sa totalité». Ce qui nous ramène au problème «Corps-Esprit». Un des aspects fondamentaux de la condition humaine est que nous sommes attachés à des corps ; et que bien des choses qui nous sont importantes dans la vie sont associées à ce qui arrive à nos corps et nos organes des sens. Si bien que pour simuler l'être humain, nous devrions y inclure les interactions entre les processus cognitifs et les processus émotionnels, (lesquels sont eux-mêmes des processus biologiques), ainsi que les processus sensoriels. L'ordinateur devrait avoir accès au monde et aux autres, de la même manière que nous le faisons nous-mêmes. Il devrait être tout aussi capable d'avoir des expériences, de recevoir des intrants sensoriels provenant du corps et de s'en occuper.

Les expérimentations dans ces directions n'ont pas été poussées très loin. Une tentative de simulation de ce genre de phénomènes, dans un cadre très limité, est le programme PARRY du psychiatre Kenneth Colby, qui simule très grossièrement l'activité cognitive d'un patient paranoïde.

Ce qui rend le patient paranoïde, ce n'est pas simplement qu'il a des idées étranges sur la Mafia qui essaierait de le capturer, mais c'est aussi le fait que ces idées sont évoquées lorsque les émotions de PARRY sont perturbées et que ces émotions peuvent être agitées par des choses que vous dites à PARRY. PARRY peut devenir offensé ou irrité ou inquiet par ce que vous lui dites. Quand je dis «offensé, irrité, inquiet», je veux dire que le système dispose de mécanismes de rétro-action qui sont stimulés de la même façon que l'anxiété et la colère sont éveillées chez un patient paranoïde ou même quelqu'un de normal ; et ces mécanismes agissent sur le système cognitif de la même façon que les émotions, une fois éveillées, agissent sur le patient paranoïde. C'est-à-dire qu'ils produisent les symptômes de la paranoïa en faisant sortir le patient de la conversation normale pour le ramener à ses idées paranoïdes sur la Mafia. Certes, ce programme est très sommaire, très préliminaire ; mais il suggère une direction dans laquelle nous pourrions engager un essai à long terme pour réaliser le programme maximaliste.

Question

Quel a été l'impact, s'il y en a eu un, de la science de la conception (science of design), proposée dans «*The Sciences of the Artificial*», sur les Universités américaines et particulièrement les écoles d'ingénieurs?

Réponse

Comme je le dis dans le chapitre consacré à la science de la conception, durant la marée haute de la science autour des années 60 — depuis la seconde guerre mondiale jusqu'à aujourd'hui ou presque — la science a pratiquement chassé l'ingénierie des écoles d'ingénieurs aux Etats-Unis. Tout était analyse, rien n'était synthèse, dès lors que l'analyse pouvait être menée rigoureusement et que la synthèse ne pouvait l'être qu'intuitivement. Le ordinateur nous a donné les moyens de parler avec rigueur du processus de synthèse et cela a eu un très grand impact sur notre école d'architecture de Carnegie-Mellon, qui utilise les ordinateurs de multiples manières tant à des fins de synthèse que d'enseignement de la synthèse.

Cela a aussi conduit au développement d'un Institut de Recherche sur la Conception¹⁵ au sein de l'école d'ingénieurs, qui permet de rassembler les membres des facultés d'ingénierie : plus particulièrement intéressés par les processus de synthèse. J'estimerais que la courbe de progression du nombre de ces membres est passée de 2 à 15 au cours de la dernière décade.

Je suis moins bien informé quant au développement de ce mouvement dans les autres écoles d'ingénieurs, ou quant à son émergence ailleurs. Je n'attribue pas ce développement à mon livre «*The Sciences of the Artificial*». Je l'attribue au fait que l'intelligence artificielle produit des systèmes de synthèse qui, naturellement, pénètrent tôt ou tard les processus éducatifs. Ils servent de propagande par l'exemple. N'était-ce pas Georges Sorel, le célèbre syndicaliste, qui disait à ce propos «Pas de propagande du mot, mais une propagande de l'action exemplaire»? Eh bien, les ordinateurs et leurs programmes de computation sont la propagande par l'exemple de la science de la conception.

Question

Dans quelle mesure les ordinateurs peuvent-ils simuler la pensée humaine s'ils ne sont pas connectés aux sens humains de la vision et de l'audition? Par ailleurs, la métaphore et l'analogie ne sont-elles pas cruciales pour la pensée humaine ? Comment pouvez-vous les simuler?

Réponse

Votre supposition est tout à fait correcte. Les ordinateurs sont privés d'intrants sensoriels et il est clair qu'il faudra que cela change, pour l'avènement de certaines choses que j'ai évoquées. L'étude de l'organisation de la reconnaissance (pattern recognition) visuelle et auditive par ordinateur est un domaine important de la recherche informatique, et ceci depuis ses débuts. Ce n'est pas un domaine où j'ai été directement impliqué, mais c'est à l'évidence un secteur-clé comme celui de l'exécution de tâches motrices. Lorsque l'intelligence artificielle a démarré, nombre de gens pensaient que ces domaines allaient progresser très rapidement. Ils estimaient que ce qui était difficile serait «la pensée profonde» que les

¹⁵ NdT : traduction de «Design Research Institute».

individus mettent en oeuvre dans leur cerveau central.

Ce qui se montre faux ; il s'avéra plus facile de simuler un individu en train de résoudre un problème — tel un ingénieur ou un professeur confrontés à un problème — que de simuler une personne regardant une scène complexe et lui conférant un sens.

Nous aurions dû deviner qu'il en serait ainsi ; si nous avions été biologistes, nous aurions dû le deviner. La partie du cerveau qui intervient dans la reconnaissance des formes est celle que nous partageons avec tous les mammifères. Elle est allée évoluant depuis 200 millions d'années de manière très impressionnante ; pour la simuler, nos progrès sont extrêmement lents. La partie du cerveau afférente à la «pensée profonde», remonte peut-être à 2 millions d'années et n'a pas eu beaucoup de temps pour évoluer. Ce qui fait qu'elle est probablement assez primitive.

Mais celui qui a posé la question a soulevé un autre problème celui que l'un des principaux composants de la pensée humaine est la métaphore ou l'analogie : se saisir de représentations dans un domaine et les appliquer ailleurs. Tout d'abord, ne soyons pas trop romantiques à notre égard. L'usage réussi de la métaphore et de l'analogie intervient peu souvent et non sans peine. Il y a une longue tradition, en Psychologie, à propos des transferts d'apprentissage : autrement dit, se servir d'un apprentissage dans un domaine et l'appliquer dans un autre. Nous savons de ces expériences qu'un important transfert ne peut avoir lieu que dans des conditions très particulières. Il y a encore bien des mystères sur la nature de ces conditions.

De sorte qu'on ne devrait pas considérer que les êtres humains le font aisément. Le calcul n'est pas réinventé chaque jour, dès lors qu'on pense que le calcul est une métaphore ou une représentation susceptibles d'être appliquées dans de nombreux domaines. Nous avons besoin d'en savoir plus à ce sujet. Il s'agit d'un secteur actif de la recherche en intelligence artificielle, bien qu'on n'y ait enregistré peu de progrès. A Stanford, Kling conçut un programme (il y a de ceci 10 ans) qui utilisait la similitude comme base de démonstration de théorème. Dans notre Université, Newell et Moore ont entrepris quelques travaux sur un programme nommé MERLIN qui établissait des correspondances pour essayer de trouver des analogies entre les choses. Nous avons encore beaucoup à apprendre à ce sujet, mais ma propre évaluation des aptitudes humaines à réaliser tout cela est peut-être plus modeste que ne le souhaiterait celui qui a posé la question.

Question

N'est-ce pas une particularité des sciences humaines qui fait que les gens, que vous décrivez comme des systèmes artificiels, aient, pour environnement, d'autres systèmes artificiels les sociétés ?

Réponse

Je pense qu'il existe une distinction importante entre les configurations de comportement qui résultent d'une conception consciente et celles qui sont le résultat de l'interaction de deux systèmes, dans lesquels chaque action peut être intentionnelle mais où la forme de l'interaction n'est pas organisée conjointement. Il s'agit d'une idée que bien des gens ont du mal à comprendre. J'ai décrit à certains d'entre vous une expérience qui m'est arrivée, il y a plusieurs années, lorsque je donnais des cours aux étudiants en architecture de l'Institut de Technologie de l'Illinois, à l'époque où Mies Van Der Rohe en était le Président. Je leur enseignais l'économie urbaine. Comme ils étaient très imbus du rôle de l'architecte comme aménageur (ou planificateur), ils pensaient à ce qu'un système devait être et à la manière de l'être — ce qui relève de la science de l'artificiel — sans faire aucune référence à leur expérience selon laquelle une telle configuration pouvait naître d'interactions du «marché» ou de la «main invisible» d'Adam Smith. Et il n'était pas facile de faire passer cette idée dans leurs débats.

Dans «*The Sciences of the Artificial*», j'ai essayé de souligner que la frontière entre l'artificiel et le naturel est des plus vagues. Par exemple, un champ de blé est un objet très artificiel. Dans la nature, les champs de blé ne poussent pas ; ils ne poussent que là où il y a des fermiers. Et pourtant là, la croissance du blé est un processus biologique. Résoudre ce paradoxe apparent réside en reconnaître que les objets artificiels sont toujours créés au sein des contraintes externes de l'environnement et des contraintes internes qui leur sont propres...

Le cas du marché apparaîtrait comme un curieux type d'artéfact : vous avez une série d'artéfacts individuels fonctionnant selon leurs propres contraintes. Ces artéfacts interagissent les uns avec les autres puis chacun d'eux détermine l'environnement des autres.

Je pense que cette idée, selon laquelle il y a des ensembles particuliers de situations où l'environnement de chaque objet artificiel constitue un autre objet artificiel, mérite d'être développée ; ce que je n'ai pas fait dans «*The Sciences of the Artificial*». Evidemment, si je la développais, je ne me limiterais pas aux systèmes sociaux, puisque c'est toute l'écologie qui est concernée ici. Sur ce point, nous avons pu écouter l'exposé de M. Godron. Ici aussi, on peut considérer que chaque plante constitue un système naturel ou un système artificiel, mais le fait qu'elle interagisse avec un environnement donné déclenche des «propriétés émergentes» et nous devons disposer de théories sur ces systèmes émergents.

Ainsi, nous avons besoin de théories de l'écologie, tout autant que nous avons besoin de théories de l'économie, et, bien entendu, pour cette raison évidente, il nous faut quantité de communications entre ces deux disciplines. Les théories de l'écologie ressemblent énormément à certaines théories économiques, ainsi que l'ont remarqué des gens comme Lotka depuis plusieurs années.

Mais, merci pour l'idée. Je ne sais pas s'il y aura encore une nouvelle édition de ce livre¹⁶, mais s'il y en a une, je veillerai à ce que ce problème soit abordé. Mais je pense qu'il ne s'agit pas d'une distinction entre les sciences sociales et les autres : il s'agit simplement du fait que nous avons une hiérarchie de systèmes avec différentes combinaisons de l'artificiel et du naturel à différents niveaux ainsi que différentes combinaisons de conception intentionnelle et de configurations résultantes à différents niveaux.

Je ne peux m'empêcher de faire une dernière remarque. Une grande partie des propriétés les plus importantes des protéines sont des propriétés qui ne découlent pas directement de la séquence des acides aminés dans les protéines, mais proviennent indirectement du fait que la séquence chimique provoque le pliage de la protéine selon une forme particulière. C'est l'interaction entre ces protéines ainsi formées, leurs propriétés tertiaires (ainsi qu'on les appelle habituellement), qui détermine leur activité biologique. On peut même faire des substitutions dans la séquence, aussi longtemps qu'on ne perturbe pas la forme de la protéine. On trouve donc bien la même forme de système — que j'appellerai «troisième système» — jusque dans le «système naturel» des protéines.

Question

Cette question porte sur la présentation du paradigme du système de traitement de l'information. En se mettant à la place de l'enseignant qui, face à ses étudiants, essaie d'en donner les concepts-clés en quelques pages. On a le sentiment qu'on doit partir de «*Human Problem-Solving*», à vous de dire si vous êtes d'accord ou non. Dans votre récent article, à propos des relations entre art et science¹⁷ il y a 3 ou 4

¹⁶ NdT : il s'agit de «*The Sciences of the Artificial*». Note postérieure 2001: Il y eu en effet en 1996 une 3^e édition complétée d'un chapitre nouveau abordant partiellement ce thème. Cette nouvelle édition n'est pas encore traduite en français .JLM.

¹⁷ NdT «L'unité des arts et des sciences : la psychologie de la pensée et de la découverte». Traduction française d'une adresse à l'Académie Américaine des Arts et des Sciences, (mars, 1982), publiée dans *AF CET-Interfaces*, n^o

pages — qui disent la même chose — mais dont on peut penser qu’elles ajoutent quelques éléments nouveaux à votre édifice. Avez-vous à l’esprit un texte plus définitif ou considérez-vous qu’il est naturel que nous nous adaptions intelligemment à votre nouveau monde?

Réponse

Bien, il est très difficile de répondre. Ce court essai sur l’art et la science constitue une distillation de vues plus récente que le livre, mais je pense que vous savez tous que les textes que l’on produit tout au long de 47 années ne forment pas un monolithe. Si l’on me demandait d’écrire un essai qui rassemblerait tout ceci en une synthèse unique qui serait auto-cohérente et agencerait tous les textes les uns avec les autres, je plaiderais mon incapacité à le faire pour cause de rationalité limitée. J’aurais tendance à regarder cette collection d’écrits non pas comme un grand corps de travaux se parlant à lui-même ou se révisant sans arrêt, mais plutôt comme un ensemble de rapports d’avancement de travaux que j’ai lancé dans le courant des idées et que d’autres- peuvent utiliser pour construire dessus comme j’ai moi-même bâti sur leurs contributions. Je ne suis pas crispé — peut-être pas assez crispé — du fait que ce que je me vois en train de dire le Mardi ne soit pas tout à fait pareil à ce que je raconte le Vendredi. Je ne veux pas dire que l’on ne doit pas essayer d’être cohérent, mais je ne me sens pas déprimé lorsque je n’y parviens pas.

Question

Voyez-vous dans vos notes quelques autres points que vous aimeriez développer un peu?

Réponse

J’ai seulement noté une chose que j’aimerais dire à propos d’une intéressante partie de la discussion de Vendredi matin (séances consacrées aux sciences de la conception et à l’intelligence artificielle). Bon nombre d’exposés se terminaient, selon moi, sur un thème commun que l’on pourrait qualifier de dialectique entre le processus et le produit de la conception et de la pensée. Le processus produit un produit qui, à son tour, devient un intrant dans le processus. Il y a eu récemment quelques discussions sur cette dialectique entre notre groupe de l’Université Camegie-Mellon et l’Ecole de Genève. Il est clair que nous avons besoin de comprendre et d’être capables de caractériser, en termes à la fois de processus et de produits, les systèmes dont nous traitons. Mais j’ai été intéressé par la manière dont ce jeu interactif émergeait, selon des voies tout à fait différentes, d’au moins trois des exposés de vendredi matin.

Je n’ai pas répondu à la question de savoir si les programmes d’apprentissage peuvent être plus que des programmes par «essais et erreurs» et ma réponse est oui. La vieille idée du perceptron, essayant des alternatives jusqu’à ce qu’il trouve une forme satisfaisante, n’est plus très viable aujourd’hui. Il y a maintenant d’autres approches de l’apprentissage, et des schèmes d’apprentissage, qui sont d’une grande puissance pour l’extraction de l’information à partir des situations. Hier après-midi, l’un des exposés mentionnait «l’apprentissage à partir d’exemples». On construit dorénavant des systèmes adaptatifs de production qui apprennent à partir d’exemples.

De tels programmes peuvent examiner un exemple déjà traité, imprimé dans un manuel ou fourni par un

enseignant, montrant la manière de résoudre une équation. En n'en sachant pas plus que ce qu'il y a dans cet exemple, et en utilisant des techniques du type de celles du GPS (General Problem Solver), le programme est capable d'extraire de cet exemple l'information lui permettant de se reprogrammer lui-même et d'exécuter des tâches relevant de cette classe générale de travaux.

Il serait capable de résoudre la plupart des équations linéaires après avoir eu un exemple de la manière de résoudre une équation linéaire. Dès lors, il est clair qu'il ne s'agit pas d'un apprentissage par essai-erreur. Il y a, à présent, un grand nombre de recherches en cours sur l'apprentissage par des systèmes capables d'extraire de grandes quantités d'information à partir de leur environnement, afin de produire des changements appropriés de leurs propres programmes.

Question

Peut-on vous demander comment, dans votre vie personnelle, vous arrivez à rester attentif à l'information qui vous vient des décisions d'autrui ?

Réponse

Ma remarque était la suivante : «il est très important de sélectionner l'information à laquelle nous prêtons attention parce que nous ne pouvons accorder de l'attention qu'à une fraction infime de celle qui est disponible». Comment est-ce que je m'y prends ? D'abord, je mène une vie très désordonnée, qui n'est guère planifiée, si bien que c'est tout juste si j'arrive à lire tant soit peu. Ensuite, et de plus en plus, à mesure que je me fais des amis dans différentes disciplines, ce sont eux qui me disent ce que je devrais lire et il s'avère que ce filtre est très fiable. Les amis sont une bonne source de filtrage ; le plus souvent, ils ne me demandent pas de lire des choses sans intérêt. Si bien que, de plus en plus, je m'en remets à mes amis pour filtrer l'information.

Mais, par ailleurs, je pense que les règles les plus importantes sont celles qui concernent ce qu'il ne faut pas lire, car on peut être assuré que certaines sources seront hautement redondantes, si ce n'est non-pertinentes. Par exemple, je crois que c'est un suicide scientifique que d'essayer de se tenir au courant de ce qui est publié dans un domaine étendu. On doit, bien sûr, connaître la littérature concernant le problème immédiat sur lequel on travaille. Mais il y a toutes sortes de façons de rester en prise avec la littérature.

J'ai observé que, dès lors que j'ai compris les idées centrales d'un domaine donné, et laissant celui-ci de côté pendant 5 ans, je pouvais y revenir ensuite sans qu'il s'y soit passé grand chose ; ou bien, s'il s'est passé quelque chose, je l'étudie comme un nouveau domaine.

On peut se mettre à une diète très sévère en matière de lecture des revues professionnelles courantes de son propre domaine et être sûr de pouvoir trouver ce dont on a besoin en faisant de nouveau appel à ses amis ou, en désespoir de cause, en recourant aux aides bibliographiques ; c'est suffisant pour trouver l'information dont on a besoin.

Cela vaut aussi pour la littérature non-professionnelle. Dans ce cas, il faut se mettre à la diète progressivement. On commence par noter que notre journal quotidien dit aujourd'hui 90% de ce qu'il disait hier. Quoi de nouveau aujourd'hui ? 90% est prévisible ; par exemple, à Pittsburgh, il arrive parfois que des gens en tuent d'autres avec une hache, un couteau ou une autre arme à feu. Maintenant, pour tel jour donné, on ne sait pas qui va se servir d'une hache ou d'un couteau ou d'une arme à feu ; mais on sait que quelqu'un va le faire et, habituellement, ce n'est pas quelqu'un que l'on connaît. Si c'est quelqu'un que l'on connaît, un ami vous le téléphonera et vous n'aurez plus à le lire dans les journaux. Tant et si bien que les journaux quotidiens sont hautement redondants.

D'abord donc, vous abandonnez votre quotidien, la radio et le journal télévisé au profit d'un magazine hebdomadaire. Vous trouvez alors la même redondance dans le magazine hebdomadaire et vous essayez un magazine mensuel, en découvrant que celui-ci vous tient suffisamment au courant. Aux Etats-Unis, nous disposons d'un merveilleux livre annuel appelé «*The World Almanac*», qui est publié pour les journalistes de quotidiens. Il contient toutes sortes de faits utiles, si bien que vous pouvez étendre votre diète jusqu'à ce que vous ne lisiez plus que «*The World Almanac*». Et ceci vous économise une énorme quantité de temps que vous pouvez consacrer à d'autres informations plus intéressantes.

Question

Je me souviens d'une réponse de Jean Piaget à une question similaire : «j'ai un grand avantage sur les autres, car je n'ai pas à lire mes propres écrits». J'en viens à ma question, qui sera la dernière de ce Colloque. Lorsque, avec André Demailly, nous avons commencé à penser à cette rencontre en cherchant lesquels de nos collègues francophones s'intéressaient au paradigme simonien, nous avons repéré quelques domaines où il n'y en avait pas et notamment le domaine de la logique mathématique. Ou nous n'en avons trouvé que très peu. Or, à l'examen de votre pensée, on constate que le recours à la logique va croissant, d'année en année. Comment voyez-vous cela? Pensez-vous que ce point est important ? Et, plus généralement, voyez-vous d'autres disciplines qui auraient dû être représentées au cours de cette rencontre ?

Réponse

Je ne suis pas très au courant des recherches en logique en France. Je connais quelques personnes et quelques travaux mais je n'ai pas une vue d'ensemble de la situation. Dans d'autres pays, il existe plusieurs domaines de la logique sur lesquels travaillent plusieurs groupes de logiciens ; je trouve que leurs travaux sont tout à fait pertinents.. - et j'essaie de les comprendre et de travailler avec eux. Je pense à Dana Scott, dans notre propre département des sciences de la computation, bien que ses travaux portent principalement, à présent, sur des problèmes formels de science de la computation qui ne sont plus directement liés à mes propres travaux. Notre interaction s'est faite, bien avant, à propos de travaux sur les fondements de la science et l'axiomatisation des théories.

Il y a deux domaines de contact entre mes travaux et la logique. L'un d'eux est bien connu en science de la computation et concerne l'utilisation possible des logiques modales, tout particulièrement des logiques qualifiées de «non-monotones». Je trouve que c'est intéressant et j'ai même écrit un peu sur ce sujet, en m'opposant principalement aux idées de McCarthy. Mon dernier article sur ce thème a été publié dans *Artificial Intelligence*, en 1983.

Mon argument central est en diptyque:

1 — je ne pense pas que les systèmes de résolution de problèmes gagnent à être vus comme des systèmes de raisonnement déductif — ce principe de résolution ne me semble pas récapituler l'ensemble du raisonnement inductif humain —.

2 — les logiques standard (dès lors qu'on a recours à une logique) ont la plupart des vertus et très peu de vices des logiques spécialisées qui ont été proposées par ailleurs.

Il apparaîtra peut être que je me trompe sur les deux tableaux, mais c'est une position en laquelle je crois pour l'instant.

La logique formelle, prise dans un sens très large — et non comme technologie —, a une formidable influence métaphorique (ou analogique, si vous préférez) sur tout le développement du traitement de l'information. Mais cela n'implique guère les activités courantes des logiciens.

L'autre domaine de contact de mes travaux avec la logique, dont il n'a pas été beaucoup discuté ici, est l'axiomatisation des théories scientifiques. Dans ce domaine, les considérations théoriques sur les modèles sont tout à fait essentielles et m'amènent à avoir beaucoup de contacts avec des logiciens, tant aux Etats-Unis qu'en Finlande, où il a pu se créer des groupes de gens intéressés par ce problème... ainsi qu'en Pologne, qui a une vieille tradition dans ce genre de réflexion. Je ne crois pas avoir, en France, de tels contacts en ce domaine.

Conclusion

Laissez moi vous dire combien ce rassemblement a été pour moi une expérience très émouvante : voir une lignée d'idées qui virent le jour voici près de 40 ans — ce ne sont pas seulement les miennes mais tout ce flux d'idées dont je vous ai parlé la nuit dernière — en train de croître et prospérer en une grande plante luxuriante ; voir toutes ces directions dans lesquelles les gens en France poursuivent la recherche en science des systèmes, en sciences de l'artificiel, dans leurs manifestations les plus variées..., me va au fond du cœur ; il s'agit, pour moi, d'une expérience très chaleureuse et exceptionnelle dont je me souviendrai longtemps.