

LES TRAITEMENTS SUR ORDINATEUR DANS LES SCIENCES HUMAINES

(Quelques données récentes sur le problème)

Par R. DEVAUGES

Un symposium sur le thème des "Méthodes de calcul dans les Sciences de l'Homme" a été organisé en juillet dernier par le Centre International de Calcul (ICC) de l'ONU, à son siège à Rome. Ce colloque a été immédiatement suivi par une série de journées d'étude plus spécialisée consacrée à "La théorie des Graphes et ses applications", également dans les Sciences humaines. Ces manifestations faisaient partie d'une série de colloques organisés chaque année depuis 1962 par le Centre International de Calcul pour "promouvoir le développement de certaines branches (de l'Informatique) et pour favoriser la coopération internationale entre experts" (1). Nous sommes partis des travaux présentés à cette réunion pour dégager les tendances principales de ce qui mérite d'être appelé maintenant la recherche sur ordinateurs et pour formuler certains des problèmes que ce mode de recherche paraît poser au point de vue des sciences de l'Homme elles-mêmes.

1 - ORIENTATIONS ACTUELLES DE LA RECHERCHE SUR ORDINATEURS.

Le thème de la réunion dont il est rendu compte ici était interprété de façon restrictive en ce sens qu'il concernait non pas l'ensemble des applications des mathématiques aux sciences de l'Homme mais seulement

celles susceptibles d'être effectuées par le moyen des calculatrices électroniques. Nous n'évoquerons ici que les communications ayant fait l'objet d'un résumé écrit à l'intention des participants au Symposium et dont certaines d'ailleurs n'ont pu, pour des raisons diverses, être présentées en public. Ces communications étaient accompagnées d'exposés plus brefs portant sur des recherches particulières / ^{qui en constituaient} des compléments ou des illustrations : nous ne nous référerons toutefois qu'occasionnellement à ceux-ci, notre objet étant seulement de dégager des tendances et des voies de recherche dans les perspectives qui nous ont parues intéressantes. Un compte rendu exhaustif dépasserait en effet le plus souvent nos compétences dans le domaine des techniques mathématiques employées et ferait en outre double emploi avec la publication future de ces travaux.

Les participants à ce Symposium, au nombre d'une cinquantaine, se composaient de mathématiciens spécialisés dans les techniques de calcul applicables aux Sciences Humaines et de chercheurs de ces disciplines formés, réciproquement, à l'utilisation des mathématiques. En fait, une certaine ambiguïté et une certaine incompréhension ont parfois gêné la communication entre ces divers spécialistes. Celle-ci tenait semble-t-il au recouvrement insuffisant de leurs domaines de compétence. Cette ambiguïté s'est manifestée par exemple dans la façon dont étaient regroupées les communications présentées : par discipline, soit anthropologie, archéologie, psychologie et sociologie (2). Cette classification était due, évidemment, aux chercheurs de Sciences Humaines. Le Professeur BEIGE - parlant en tant que mathématicien - a exprimé le regret qu'elles ne l'aient pas été, au contraire, en fonction des méthodes de traitement utilisées. Choisisant ici une troisième classification - plus conforme selon nous aux préoccupations des chercheurs de Sciences Humaines - nous essaierons de regrouper ces communications selon les types de problèmes qu'elles abordent. Nous les classerons, dans ce but, en trois catégories : la première concernera les usages relativement simples des ordinateurs tels que l'analyse de contenu ou le classement de l'information. Dans la seconde entreront les traitements mathématiques généraux et dans la troisième la description de modèles particuliers, associée souvent à des utilisations spécifiques des ordinateurs telles que la simulation. Comme on le verra d'ailleurs, ces di-

verses catégories se recouvrent partiellement.

I-1 - Du langage naturel au langage symbolique.

Les recherches du premier type se rencontrent seulement en anthropologie et en archéologie, c'est à dire dans des disciplines offrant un champ particulièrement riche et solide au traitement exhaustif de l'information, mais se trouvant actuellement, pour des raisons diverses, au début de l'utilisation de ces procédés. Une partie des communications concerne des méthodes d'analyse de contenu destinées à classer les données recueillies en dehors de toute interprétation ou hypothèse préalable. Ces procédures se rencontrent généralement en archéologie où le problème est d'analyser et de classer les objets recueillis au cours des fouilles d'une manière "objective" c'est à dire établie en dehors des théories interprétatives disponibles et permettant plusieurs classifications possibles. C'est le cas des travaux des chercheurs allemands C. ANKEL et R. GUNDLACH et du suédois C.A. MOBERG. Dans la même catégorie se situent - en anthropologie - les analyses de Mrs V.R. BRICKER (USA) cherchant à établir la signification psychosociologique de l'humour dans une population Maya du Yucatan en identifiant divers thèmes culturels, puis en analysant, dans les différents textes, leur fréquence d'apparition. Quelques autres travaux se regroupent à un niveau d'élaboration plus élevé, puisqu'il s'agit, par le classement objectif des données, de faire apparaître des structures. C'est le cas des recherches d'analyse structurale de P. MARANDA (USA), visant au moyen d'un langage analytique convenable à mettre en relief l'"arête" et certains mécanismes structuraux des mythes analysés, des procédures de restitution directe sous forme cartographique des différences de champ magnétique relevées dans un site archéologique, présentées par I. SCOLLAR (Allemagne) et R. LININGTON (Italie) et des méthodes décrites par B. SOUDSKY (Tchékoslovaquie) permettant de déterminer au moyen d'une calculatrice les stratigraphies d'un site à partir des observations relevées sur les pièces recueillies.

I-2 - Les techniques mathématiques générales

Les quelques communications rangées dans la seconde catégorie proposée ici concernent des traitements mathématiques très généraux, rendus accessibles grâce à l'utilisation des grands ordinateurs. Plusieurs d'entre elles proposent l'application à des phénomènes organisés de structures ma-

thématiques particulièrement aptes à rendre compte de cette organisation : en anthropologie, J.P. BOYD (USA) applique la théorie des monoïdes aux structures de parenté, "l'ensemble des relations définissant le monoïde, (permettant) de définir un ensemble de représentations qui est lié au problème de l'analyse componentielle des systèmes de parenté". En archéologie, V. ELISEEFF (France) met en évidence le fait "que l'ordre des transformations survenues dans l'évolution chronologique des vases archaïques chinois peut être représenté par un scalogramme". Il est certain dans ce dernier exemple que l'application d'un modèle de structure cumulative paraît particulièrement adéquat pour l'étude des phénomènes se déroulant dans le temps et accumulant en quelque sorte leur passé derrière eux. Dans la même catégorie des descriptions de procédures générales nous classerons en psychologie l'exposé de J.C. LINGOES (USA) sur le développement récent des programmes de calcul pour l'analyse non-métrique. Ces procédés, fondés sur des hypothèses minimales sont en effet beaucoup plus "naturels" dans l'état actuel de nos connaissances des phénomènes humains, que ceux reposant sur des "métriques" trop fortes.

D'un caractère plus complexe et nous rapprochant des modèles dont il sera question plus loin, dans la mesure où les techniques proposées visent à nous faire pénétrer dans le domaine de l'explication, sont les communications conjointes de B. ROY et M. BARBIT pour la France et celle de H.C. SELVIN pour les USA. Les deux premiers présentent des méthodes de choix lorsque plusieurs critères sont en jeu. Il s'agit, au moyen de l'outil d'analyse proposé, de "déduire de préordres complets induits par les critères sur les objets entre lesquels on doit choisir une relation de préférence majoritaire et (de) choisir parmi les éléments d'une partie stable de cette relation". La seconde communication pose le problème voisin de "séparer les effets de la variable indépendante de ceux des facteurs provenant de causes extérieures", c'est à dire des variables parasites ou aléatoires. Dans les expériences, cette séparation est essentiellement une question de "schéma", c'est-à-dire de plan d'expérience. Dans le cas d'enquêtes, où l'on ne peut intervenir sur les sujets, elle doit au contraire s'opérer entièrement au moyen de ce que l'auteur appelle l'analyse tabulaire, c'est à dire en comparant les tableaux de contingence des pourcentages. Mais cette

méthode ne peut, au mieux, dépasser 4 ou 5 variables ; elle nécessite en outre des échantillons importants pour donner des résultats significatifs et elle nous fait finalement assez mal pénétrer dans le mécanisme causal des phénomènes (3). L'apparition des grands ordinateurs permet le recours à ce que SELVIN appelle l'analyse linéaire. Celle-ci permet de "tenir compte de facteurs extérieurs aussi nombreux qu'il est nécessaire", 50 ou 100 selon l'auteur, tout en travaillant sur des échantillons relativement restreints - donc plus économiques - de 100 ou 200 individus. La multiplication des variables d'analyse permet de pénétrer beaucoup plus profondément que l'analyse tabulaire classique dans le déterminisme des phénomènes.

I-3 - Les usages spécifiques des ordinateurs : modèles et simulations.

Les communications réunies dans la troisième catégorie proposée ici, concernent des applications plus complexes et plus spécifiques des ordinateurs, puisqu'il s'agit cette fois de recourir aux "modèles". L'archéologue R.G. CHENHALL (USA), dans une communication sur laquelle nous reviendrons par ailleurs, définit avec clarté le rôle de ces modèles dans la recherche : "Récemment, écrit-il, un mathématicien décrivait sa tâche essentielle comme consistant à examiner une situation physique ou un fragment du "monde réel", et à développer un modèle mathématique pour le représenter. Ce modèle, poursuit-il, consiste en une série d'équations ou de relations et il doit présenter une ressemblance suffisante par rapport à la situation originale pour que, lorsque des questions trouvent une réponse dans les termes du modèle, l'on trouve les mêmes réponses aux mêmes questions dans le monde réel". Le rôle de l'ordinateur dans cette tâche est alors d'être un "outil mental" auquel on demande de "traiter efficacement une vaste quantité de données". Les modèles utilisés doivent alors répondre non plus à une seule mais à deux catégories d'exigences, c'est à dire être "à la fois (a) représentatifs de la réalité (...), (b) capables d'être convertis dans la langue des ordinateurs pour un traitement par des techniques mathématiques".

Les communications groupées dans cette troisième catégories, présentent plusieurs degrés de complexité dans l'utilisation des modèles ou, plus exactement, dans les phénomènes auxquels ils sont appliqués et c'est par ce caractère que nous les ventilerons à leur tour en sous-catégories. La construction de modèles applicables à des phénomènes stables, statiques et

structurés, apparaît surtout, comme il est naturel, chez les archéologues. A.B. COULT (USA) étudie la signification des diverses analyses automatiques pour les théories anthropologiques, et les types principaux d'opération qu'elles autorisent : classements complexes et objectifs, analyses de corrélations sur un grand nombre de variables et enfin opérations de formalisation sur des systèmes naturels ou théoriques où la machine permet parfois de déceler des erreurs. La communication de J. CUISENIER (France) porte, comme précédemment celle de BOYD, sur l'analyse des structures de parenté. Elle propose cette fois un modèle exploitable sur ordinateur et permettant de reconstituer et de caractériser au moyen de quelques nombres caractéristiques les structures de parenté. Cette reconstitution et cette analyse s'opèrent sur un univers de parenté déterminé réparti en tranches synchroniques, par le moyen de consignes données à la machine et permettant de suivre et de distinguer les individus à travers les homonymies et les changements de nom.

L. FREY, également en France, étudie d'un point de vue purement formel les passages parallèles des Evangiles synoptiques. L'analyse des relations d'ordre entre ces passages permet d'interpréter les interversions entre certains de ces passages et de reconstituer les sources les plus importantes. Elle offre des éléments pour choisir entre les théories en présence relativement à l'ordre chronologique de ces Evangiles et aux sources qui les auraient inspirés. R. JAULIN (France) propose une analyse formelle de la géomancie reposant sur l'étude des systèmes dits "en activité" et du système-étalon et des correspondances entre les deux. Appliquant ensuite ces "caractéristiques formelles à l'ensemble des significations géomantique (il formule l'hypothèse) que le signifié géomantique a été organisé en raison des propriétés géométriques, arithmétiques et algébriques du système". CHENHALL déjà cité, s'efforce de définir le rôle du modèle mathématique en archéologie. Selon lui, les étapes relatives à la traduction du langage matériel en langage symbolique, à l'exploitation des données sous cette forme symbolique et à l'interprétation finale des résultats ont fait beaucoup de progrès. Par contre, "la nécessité majeure est maintenant de clarifier et de standardiser les concepts et les termes utilisés dans l'identification, la description et la classification archéologi-

que". Les progrès sont donc à réaliser au niveau de la phase qualitative et spécifiquement archéologique de l'analyse des matériaux. Il convient en particulier - et nous avons déjà trouvé cette préoccupation chez plusieurs auteurs - d'admettre, au niveau de la codification, plusieurs systèmes de classification possibles "pour ne pas limiter les efforts de recherche".

Un autre groupe de communications - en Sociologie cette fois - proposent des analyses de groupes de variables - variables de milieu et de comportement - en vue de prédire des comportements "simples" dans des situations socialement codifiées. H. ROSENTHAL (USA) propose de perfectionner un modèle mathématique applicable aux coalitions en introduisant les relations culturelles entre les "joueurs" (ici les partis politiques). "Le modèle permet une confirmation des résultats de la coalition sans avoir recours aux informations au sujet de la personnalité des candidats ou des autres particularités de la circonscription électorale". Concernant également le domaine de la sociologie électorale, mais dans un esprit très différent A. KALMAN (USA) présente les résultats d'une étude-pilote visant à définir le modèle analytique des comportements au vote législatif. Au lieu des analyses statistiques sur une seule dimension réalisées jusqu'ici on espère que le modèle réalisé permettra, à partir de corrélations multidimensionnelles entre des comportements antérieurs et des caractéristiques diverses, d'effectuer des prévisions pour les votes futurs. L'auteur a présenté les résultats obtenus jusqu'ici comme imparfaits, mais prometteurs. Dans un domaine plus général K. RANIO (Finlande) a présenté la description "d'une série de simulations de l'interaction sociale dans laquelle cette dernière est considérée comme un processus séquentiel de choix individuels". Le modèle de simulation a été appliqué à différents domaines où apparaissent divers aspects de l'interaction sociale : contacts sociaux aboutissant à certaines structures d'amitié, contacts entre groupes, résolution de problèmes dans des groupes de petite taille. Afin de tester le réalisme du modèle "des comparaisons sont faites avec des données empiriques".

Nous avons groupé à part les communications concernant l'application des modèles non plus comme dans le cas précédent à la prévision de comportements à partir d'éventualités connues à l'avance, mais à la définition de lois de comportement générales. Le psychologue suédois G. KARLSSON propose un modèle pour l'étude de l'influence dans les groupes de petite taille. Le principe de ce modèle est qu'un comportement en plus ou en moins,

est déterminé par une attitude x . L'influence d'une personne i sur une personne j , supposée constante pendant une période donnée, s'ajoute à x_j ; si l'action de i est positive, s'en retranche si cette action est négative. L'auteur discute des cas d'application de ce modèle pour différentes structures et différentes valeurs. Toujours en psychologie, N.H. FRIJDA (Hollande) présente une tentative plus ambitieuse qui est la construction d'un modèle de fonctionnement à long terme de la mémoire humaine, utilisable sur une calculatrice. Dans ce modèle "le stockage des informations se présente comme un réseau d'informations avec, malgré tout, des liaisons entre les noeuds (ou éléments) de type qualitativement différent". La restitution simulée de l'activité de la mémoire "consiste à trouver un élément qui soit déterminé par un autre au moins et sa relation avec l'élément à restituer". Les auteurs espèrent que ce modèle, qui est actuellement en cours d'élaboration, dont le principe même a provoqué de vives controverses avec certains participants du Symposium, permettra une fois au point de "résoudre avec réalisme des tâches simulées de reconnaissance des tâches imitant la résolution des problèmes et des activités de libre association". Le sociologue américain B. LIEBERMANN étudie le problème, classique pour les économistes, de découvrir "comment les préférences individuelles peuvent être condensées ou amalgamées à un choix social". Il propose pour cela un modèle d'analyse permettant d'aider à "faire comprendre comment les propriétés individuelles plongent réellement dans un cadre social".

Les trois communications dont il va être question maintenant ont été groupées en dernier parce qu'elles nous ont paru réaliser l'utilisation la plus complexe et la plus élaborée, à la fois des modèles et des ordinateurs.

Elles concernent toutes les trois l'activité d'un sujet s'appliquant à la résolution de problème. S'il s'agit encore comme dans le cas des travaux de FRIJDA, de simuler un comportement psychologique, ce n'est plus cette fois dans un domaine général, mais lorsque ce comportement est appliqué à la résolution d'une tâche complexe bien définie. A. NEWELL (USA) prend pour sujet d'étude les protocoles. Il définit le protocole comme "un enregistrement du comportement dans le temps". Dans les expériences étudiées ici, le sujet obéit à la consigne de penser

tout haut. La traduction des protocoles en termes de théorie de l'information présente des difficultés. L'auteur propose ici : "de se servir d'espaces de problème et de ce qui y correspond : le graphe du comportement de problème en tant que technique d'analyse de données au premier échelon". Il peut par ce moyen, fractionner les comportements et formuler des règles abstraites de réponse aux divers stimulus. C'est à partir de ces données au premier échelon que sont établis les programmes. Les travaux de F. KLIX (Allemagne) reposent également sur une application de la théorie des graphes. Il s'agit cette fois de mettre les sujets en présence d'un problème inconnu qu'ils doivent résoudre. Une procédure de résolution optimale étant définie, on étudie "les mécanismes qui sont à la base des propriétés créatrices du comportement d'un sujet recherchant un but". On traduit par un graphe toutes les démarches possibles. On peut, à partir de là, constater le progrès d'un sujet "dans l'efficience de la pensée par le passage d'une concaténation à une autre". Ce comportement est simulé dans un ordinateur, ce qui permet de montrer "qu'il existe des propriétés générales du traitement humain de l'information, classées comme stratégies de la pensée, et qui comprennent de grandes capacités pour élucider et élaborer les informations venues de l'environnement". Par l'analyse de séries statistiques de résultats, le même processus peut être étudié sur des groupes. L'étude du jeu d'échecs, considéré comme un type de problème dont les conditions sont des éléments discrets et opératoires susceptibles de se déplacer dans un certain milieu statique", constitue le sujet de la communication de V. POUCHKINE (URSS). Les difficultés rencontrées par les programmes cybernétiques appliqués jusque là par les Américains à ce problème provenaient du fait qu'ils reposaient sur l'hypothèse du labyrinthe. En fait "le processus de la solution commence par la création des modèles informatifs des pièces d'échec (...). Le processus de la formation des connexions entre ces pièces amène le joueur d'échecs à la création soit du système opératoire et informatif, soit du modèle d'une certaine position incluant des pièces déterminées de cette position et leur mouvement sur l'échiquier". C'est ce processus informatif qui "sert de base à la création d'une stratégie raisonnable dans ce milieu relativement infini".

2 - PERSPECTIVES ET PROBLEMES

Des enseignements aussi riches que nombreux peuvent être tirés de ce symposium qui aura présenté, dans un large domaine, l'état actuel des recherches de formalisation en matière de Sciences Humaines. Nous proposons ici de retenir quatre points qui nous ont paru essentiels. Les conclusions tirées pour les Mathématiciens par le Professeur BERGE, dans son exposé de clôture, constitueront le premier de ces points. Nous considèrerons ensuite les conclusions qui nous paraissent à retenir pour les Sciences Humaines : l'une concerne le rôle de stimulants et d'instruments d'expérience que les ordinateurs peuvent jouer pour la recherche, l'autre les conséquences pour la formation des chercheurs au caractère très mathématisé de ces méthodes. Nous essaierons enfin et pour terminer, de définir quelques linéaments "réalistes" d'une doctrine provisoire à l'usage de chercheurs dont le rôle est d'être spécialisés en sciences humaines davantage qu'en recherche mathématique.

2-1 - Le point de vue des mathématiciens.

Dégageant - à l'intention surtout des mathématiciens - les enseignements de cette réunion, le professeur BERGE a développé deux séries de considérations. La première concernait l'utilisation même du langage mathématique. Les qualités principales de ce langage sont la logique, la rigueur dans le raisonnement, et aussi l'habitude de diviser et de sérier les difficultés ; son utilisation ne peut donc être que salutaire pour les sciences de l'Homme. Elle n'est pas toutefois sans soulever des difficultés. L'une de celles-ci tient à la multiplicité des langages mathématiques utilisables et à la difficulté de trouver le mieux adapté : le "mythe" actuel par exemple est la théorie des ensembles. Un autre risque serait de tomber dans l'excès consistant à abandonner trop vite le langage "naturel" au profit d'une terminologie mathématique qui n'apporterait rien de vraiment nouveau aux problèmes spécifiques étudiés. Dans tous les cas, c'est la nature du problème concret et la visée du chercheur qui doivent déterminer le choix du cadre opératoire et non l'inverse.

La seconde série de considérations concernait l'aspect particulier des techniques mathématiques étudiées ici, et que le professeur

BERGE a appelé "l'étude des phénomènes mécanisables" et qui constitue selon lui l'apport le plus original de ce Symposium. Jusqu'ici les problèmes de psychologues, des sociologues, des archéologues etc... étaient analysés au moyen d'une pluralité de langages naturels peu communicables et dont la structure logique était parfois mal dégagée. La nécessité de traduire ces problèmes dans un langage normalisé et de les interpréter au moyen d'algorithmes contraint à préciser les théories et les concepts et suscite souvent la découverte de phénomènes nouveaux. La simulation joue dans cet ordre d'idée un rôle particulièrement original. La tâche des mathématiciens - et c'est la conclusion du professeur BERGE - ne se borne pas à répondre aux questions et à satisfaire aux besoins des sciences humaines. Certes, ces dernières sont pour eux génératrices de problèmes nouveaux. Mais, une fois leur demande satisfaite, les problèmes mathématiques qu'elles ont suggérés les dépassent et proposent aux mathématiciens des tâches spécifiques et plus générales qui ne doivent plus rien aux suggestions extérieures.

2-2 -- Le point de vue des Sciences Humaines.

Les centres d'intérêt qui se dégagent, pour le chercheur des Sciences Humaines, de cet ensemble de communications sont d'un tout autre ordre d'idée. Il faut d'abord remarquer que celles-ci, qui ont pour métaphysique ou plutôt pour métalogue commune les théories cybernétiques, représentent les tendances les plus avancées dans la mathématisation des phénomènes humains.

Au-delà des problèmes concrets auxquels elles apportaient une contribution, ces recherches font apparaître une conception sinon nouvelle, du moins beaucoup plus affirmée et plus audacieuse de la recherche dans les sciences du comportement. Il a déjà beaucoup été écrit sur ce sujet et nous voulons seulement ici évoquer les réflexions inspirées par la relecture des notes et des documents rapportés de ce colloque. En premier lieu, il apparaît que la nécessité de traduire le langage "naturel" de l'observation dans le langage "symbolique" de la mathématique exerce une influence incomparablement stimulante sur le travail intellectuel du chercheur. La perspective de devoir construire un "modèle", c'est-à-dire de se mettre en mesure de reproduire - ou de faire reproduire par l'ordinateur - les mouvements qu'il a observés dans le réel, l'oblige à mieux délimiter ses concepts et les réseaux de relations qui les relient, en d'autres termes à définir plus rigoureusement ses

hypothèses, à expliciter la structure logique de ses mouvements de pensée et à chasser les zones d'ombre, les prudences ou les ambiguïtés inconsciemment entretenues ; par là elle oblige le chercheur à se critiquer et, du même coup, à se mettre en situation de découvrir des conséquences inattendues de sa théorie. L'effort logique exigé par la formalisation dégage cette vérité première que toute activité de théorisation - qui est le propre de la science - est aussi une activité de mécanisation, la construction d'un "modèle analogique" du réel, ou plutôt de l'observé.

Si l'on accepte ces prémisses, le rôle de l'ordinateur apparaît sous un jour nouveau. Ce n'est plus, comme nous le disions au début, une super-machine à calculer ; c'est d'une part une super-machine à raisonner - reprenant avec une ampleur, une rapidité et une précision hors d'atteinte des ordinateurs biologiques que nous sommes (4), les raisonnements que nous avons bâtis - et d'autre part une machine à reproduire le réel, à le simuler, conformément du moins à nos hypothèses. Un mauvais usage des ordinateurs serait un usage dogmatique : à partir d'observations insuffisantes nous construisons un modèle et nous légiférons à partir des conséquences de ce modèle sans nous soucier de leur adéquation au réel ou des changements non-prévus de celui-ci. Mais on peut dès maintenant concevoir d'une façon encore un peu futuriste mais qui est dans la ligne des travaux présentés ici, une autre conception, expérimentale celle-là, des ordinateurs. Celle-ci a d'ailleurs été suggérée par la communication de SELVIN. Comme nous l'avons déjà souligné, cet auteur a fait remarquer que, dans les enquêtes de terrain, on ne pouvait guère recommencer - et surtout un grand nombre de fois - les mêmes observations. L'expérimentation proprement dite - qui est le propre des sciences de la nature - y est alors pratiquement impossible. On en est réduit pour dégager les variables pertinentes, les structures et les dynamismes à s'appuyer uniquement sur l'analyse du matériel existant. C'est dans cette analyse que l'ordinateur peut jouer un rôle considérable, beaucoup plus grand même que dans les sciences de laboratoire. Que l'on traduise en modèle les hypothèses formulées par l'analyse des matériaux d'enquête et qu'on les confie à l'ordinateur, la simulation devient le substitut de l'expérimentation impossible. En poursuivant les résultats obtenus par l'application de la théorie proposée, elle

nous fournit le moyen d'en comparer les résultats avec les situations réelles et de modifier le modèle jusqu'à obtenir dans le domaine du contrôlable, des résultats assez satisfaisants pour nous autoriser à extrapoler, au delà du domaine de l'observable. Même si la perspective de remplacer ainsi l'observation par le calcul est encore utopique, aucun obstacle logique ne nous empêche de penser que cela est "probablement possible" (5) et de prendre en tous cas cet idéal pour visée.

2-3 - L'effet de suggestion dû au modèle et l'observation du concret.

L'emploi de ces méthodes hautement élaborées et dont la fécondité et la force de renouvellement ne sont pas près d'être épuisées, n'est cependant pas sans poser aux chercheurs de sciences humaines des problèmes délicats et dont certains tiennent à leur séduction même. Le premier de ces problèmes nous paraît être celui de l'interaction qui ne peut manquer de se produire dans l'esprit du chercheur entre les servitudes impliquées par l'usage de ces "outils mentaux" et l'observation du réel.

Il y a certainement un danger qui réside dans la tendance que semblent avoir toujours manifesté le chercheur dans son activité scientifique à se représenter le monde par analogie avec les machines que sa société savait fabriquer. Par exemple la conception mécaniste classique se représentait les mouvements de la nature comme des systèmes de câbles et de poulies : elle ne parvenait pas, pour cette raison à se dégager des processus de causalité linéaire. L'interaction n'a été clairement conçue et utilisée qu'à une période très récente, du jour où l'on a su construire des mécanismes auto-régulateurs, des circuits en feed-back. Et l'on sait le rôle que joue aujourd'hui la théorie des communications empruntée de façon très concrète aux ingénieurs, pour rendre compte des phénomènes de communication entre les personnes. Très clairs dans ce cas, ces mêmes modèles sont peu commodes lorsqu'il s'agit de se représenter des phénomènes de diffusion à l'intérieur des groupes ou entre les groupes. On voit par là combien l'emploi quasi-métaphorique de schémas et d'usages empruntés à des domaines externes, aussi fécond qu'il soit dans des cas limités, n'en est pas pour autant exempt de dangers. Il n'est pas douteux - et cette tendance apparaît clairement dans certaines des communications analysées ici - que les exigences du langage machine et de la structure logique qui lui a été donnée influencent parfois exagérément le langage

naturel dans lequel on codifiera par la suite l'observation des phénomènes. La suggestion, l'attitude mentale ainsi créée a pour effet de conditionner rétroactivement - nous avons là un bel exemple de feedback - et par là de déformer plus ou moins consciemment, la perception que le chercheur aura du réel, en privilégiant les faits entrant dans le modèle et en minorant au contraire - en les rejetant dans le purgatoire des variables parasites ou des "erreurs" - les faits qui ne s'y plient pas. On ne peut certes reprocher à l'homme de penser avec les outils mentaux dont il dispose mais on ne peut s'empêcher au passage de signaler cette tendance et, de souhaiter que soient réservés les droits de la pensée heuristique, la seule qui reste ouverte à l'inconnu. Le rôle de celle-ci, malgré son faible rendement technique, est de demeurer ouverte à l'invention et de relativiser à leur tour les techniques puissamment séduisantes mais enfermées dans leur cohérence interne, que nous propose le "modélisme" scientifique.

2-4 - Phase d'observation et phase de formalisation dans la recherche.

En deçà, en quelque sorte, du problème que nous venons d'évoquer, il en existe un autre, beaucoup plus modeste et plus concret, mais qu'il faut tout de même bien formuler : c'est celui du niveau de formation mathématique nécessaire aux chercheurs de sciences humaines. Nous avons signalé au début de ce travail l'incompréhension provoquée par la spécialisation très poussée d'un certain nombre des techniques proposées. Le résultat est que, sans doute, bien peu des chercheurs présents au colloque de Rome étaient en mesure de comprendre la totalité des communications présentées. Ce qui était vrai des gens venus des sciences humaines l'était aussi d'ailleurs - nous l'avons constaté - de certains des mathématiciens présents. Et c'est là que réside précisément le problème qu'il faut bien soulever. Les techniques mathématiques appliquées aux sciences humaines s'étendent dans des directions très variées - souvent entièrement nouvelles- et dans lesquelles elles atteignent rapidement un niveau élevé, surtout si on le compare à celui de la statistique scolaire. Elles sont en outre en état de rapide et constante évolution.

On peut prévoir dès maintenant qu'il faudra donner un jour aux futurs anthropologues, sociologues ou psychologues une formation mathématique de base aussi poussée que celle des physiciens. Certes, cette formation se fera dans une voie différente, adaptée aux "ensembles pauvres" - sur le plan mathématique - qui constituent leur domaine. Mais, dans la situation transitoire où nous sommes actuellement, la prolifération dans des voies multiples et la complexité rapidement croissante des recherches formelles conduisent à exprimer la crainte que le chercheur de Sciences humaines qui veut se tenir au courant ne court le risque de s'essouffler à vouloir acquérir la maîtrise de techniques d'une complexité rapidement croissante et qui, en outre, courent demain le risque d'être dépassées ou abandonnées. On peut alors se demander si, au moins dans la phase d'évolution accélérée des techniques de formalisation que nous sommes en train de traverser aujourd'hui, il ne faut pas se résigner à laisser se former deux branches de spécialisation l'une vers le concret - terrain ou laboratoire - l'autre vers la méthodologie ou, plus exactement, vers la technologie, c'est à dire vers les procédures de formalisation et de traitement "extérieur" des données. Poussée à l'extrême, cette conception de la recherche n'est sans doute pas exempte de risques : celui en particulier de voir se cloisonner les deux domaines et s'instaurer une coupure entre chercheurs "d'observation" et chercheurs "de formalisation". Il y a sans nul doute là une question difficile, mais qui n'est pas insoluble et relève de l'organisation et de la répartition des tâches. L'articulation devrait sans doute être conçue de façon extrêmement souple et, surtout, se faire au sein d'une équipe dans laquelle chacun participerait - avec simplement des degrés de responsabilité divers - à l'ensemble des tâches. Les uns y seraient plus spécialisés dans l'observation des faits et la découverte des symptômes, qui sont les liens entre ces faits et les théories spécifiques : psychologique, anthropologique, etc... Les autres auraient davantage la charge des techniques d'exploitation et du dialogue avec les mathématiciens et les programmeurs. Ils détermineraient avec ces derniers les cadres mathématiques les plus propres à traduire les formes et les mouvements aperçus dans l'observation du réel ainsi que les algorithmes permettant de traduire ces formes et éventuellement de simuler ces mouvements par le moyen des calculatrices. Ils utiliseraient ainsi pleinement/la perspective expérimentale suggérée il y a un instant, les modèles et les techniques de simulation

qui leur sont naturellement associées, et mettraient au mieux, au service de la recherche, les possibilités énormes de clarification et de multiplication de notre activité intellectuelle proposées conjointement par les mathématiques et les ordinateurs.

R.D.

(1) Brochure sur le Centre International de Calcul. On désigne par le néologisme d'Informatique le traitement automatique de l'information par ordinateur sous tous ses aspects, c'est à dire non seulement l'enregistrement et la restitution des données mais toutes les formes de calcul, de transcription graphique de simulation de modèles etc... opérées sur ces machines.

(2) On s'est étonné de l'absence de l'Economie et de la Linguistique à cette réunion d'autant que des références très fréquentes - implicites il est vrai - étaient faites à leurs théories et à leurs procédures de formalisation.

(3) Cette critique vaut pour l'analyse multivariée de LAZARFELD que nous avons évoquée dans le n° 5 de ce Bulletin.

(4) Du moins en tant que machine à raisonner et en laissant de côté l'aspect heuristique de la pensée, c'est à dire la possibilité de trouver quelque chose de nouveau et qui ne soit pas dans les données.

(5) Voir dans cette perspective l'ouvrage de Aurell DAVID : La Cybernétique et l'humain. - Gallimard, coll. "Idées", 1963