

### Calculateurs et sciences humaines

Le texte qu'on va lire est la transcription un peu abrégée de la conférence que M. J.-Cl. Gardin a prononcée le 18 février 1969 à l'Université de Neuchâtel. Elle inaugurerait une série de séminaires et de discussions qui devaient s'étendre sur quatre jours, et prenait en quelque sorte l'allure d'un manifeste.

En effet, M. Gardin, directeur du Centre d'analyse documentaire pour l'archéologie (Marseille), s'est illustré par de nombreux travaux qui ont ceci de commun qu'ils portent tous sur l'usage des ordinateurs dans les sciences de l'homme. Il est superflu de dissiper ici les malentendus que pourrait susciter l'usage du terme "ordinateur": M. Gardin le fait lui-même en toute clarté. Il faut en revanche insister sur la nouvelle attitude que représente l'introduction de méthodes mathématiques et formelles au sein des études de lettres.

On a estimé pendant longtemps, en effet, qu'il y aurait comme un sacrilège à tenter de saisir les productions humaines à l'aide de cet instrument que sont les mathématiques, instrument que bien des personnes tiennent pour grossier. Il est vrai que le langage mathématique possède des propriétés qui l'opposent sur bien des points essentiels aux langues naturelles. Il est inflexible là où elles sont toute souples; il est univoque là où elles apparaissent subtilement ambigus; il est explicite là où elles se veulent pleines de sous-entendus. Ces traits toutefois ne sont des qualités ou des défauts que selon l'usage que l'on en fait, et des travaux comme ceux de M. Gardin montrent à l'évidence qu'il en existe un bon. C'est qu'il ne faut ni confondre mathématique et science de la quantité, ni imaginer que les systèmes formels visent à se substituer à toute autre méthode. Il est simplement devenu aujourd'hui possible – en grande partie d'ailleurs grâce à l'évolution même de la logique et des mathématiques – d'aborder les disciplines humaines *more geometrico*, sans sacrifier pour autant et comme automatiquement le fameux "esprit de finesse".

Consciente alors qu'il était vain de parler sans cesse de "sciences humaines" (puisque telle est devenue la façon de dire) sans en tirer en même temps les conséquences méthodologiques, la Faculté des lettres de Neuchâtel a formé le projet de créer un Centre de recherches sémiologiques pour aborder de front les problèmes que pose le traitement scientifique des disciplines "littéraires". Or ceux-ci ne sauraient être traités par un chercheur isolé. Les compétences requises ne sont pas de celles qui se trouvent réunies chez un même homme. On ne peut réclamer du linguiste qu'il soit également algébriste, du philologue qu'il soit logicien, de l'historien qu'il se fasse programmeur. Il y a même, dans cette exigence de travaux en équipe, un signe de "scientificité" non moins certain que celui que constitue la présence d'un ordinateur parmi les chercheurs.

Entendons-nous cependant. Personne ne songe à revenir à l'homme-machine cher à La Mettrie. Personne ne songe même à *réduire* les comportements humains à des modèles logico-mathématiques. Mais, entre une réduction au sens fort du terme et le refus de toute considération formelle, il y a place pour des attitudes intermédiaires dont il n'est plus nécessaire aujourd'hui de prouver qu'elles sont fructueuses. Et puisque il est évident que l'un des traits spécifiques de l'homme est l'usage qu'il fait des systèmes de signes, c'est bien en sémiologie qu'il convient de faire porter l'effort.

L'idée d'une science des signes remonte sans doute à John Locke, mais la sémiologie, en tant que telle, doit son existence à des hommes comme S.C. Peirce, F. de Saussure, Ch. Morris, R. Barthes et autres. Et ce n'est certes pas par hasard si, dès le départ, philosophie, logique, linguistique et même critique littéraire se sont ainsi trouvées liées. La sémiologie ne peut être qu'interdisciplinaire, de même qu'elle ne saurait être absente d'une recherche quelconque qui a trait à l'homme.

On comprend dès lors que si J.-Cl. Gardin a associé son nom à une *Etude des établissements assyriens en Cappadoce par ordinateur* (1961), à une *Analyse conceptuelle du Coran sur cartes perforées* (1963), s'il a publié des articles sur les *Etats et tendances actuels de la documentation automatique* (1964), sur *A typology of computer use in anthropology* (1965), l'essentiel était pour nous la méthode qu'il illustre ainsi en divers domaines. Personne ne pouvait mieux que lui aider à mettre en place quelques-unes des recherches que projette le nouveau Centre. C'est un maître en la matière et c'est à ce titre aussi que l'article qui suit importe à la linguistique appliquée.

Université de Neuchâtel  
Centre de recherches sémiologiques  
CH 2000 Neuchâtel

Jean-Blaise Grize

L'emploi des calculateurs dans les sciences humaines a conduit les chercheurs à poser des questions de méthode dont on croit généralement qu'elles sont liées au calcul électronique. En fait, les problèmes rencontrés sont indépendants des machines et auraient dû être soulevés dans les sciences humaines bien avant l'introduction des calculateurs.

Il existe diverses typologies des problèmes liés à l'emploi des machines électroniques dans les sciences humaines. Certaines sont fondées sur le découpage traditionnel en disciplines: emploi des calculateurs en linguistique, en psychologie, en histoire, etc. Elles sont peu satisfaisantes, car les questions de méthode liées à l'emploi des machines sont largement indépendantes des champs d'application. D'autres classements considèrent le type de données, numériques ou non numériques, que l'on veut traiter en machine. Dans les sciences humaines, sauf dans quelques domaines comme la psychométrie ou la sociométrie, on traite généralement des données non numériques. Nous proposons une troisième typologie, fondée sur la nature des opérations logiques que doit effectuer le calculateur. Nous distinguons ainsi les opérations de tri, de sélection ou de tabulation, et les opérations de construction de modèles.

Dans les disciplines qui étudient des textes — linguistique, philologie, histoire, par exemple — on peut demander au calculateur d'établir des index ou des concordances (index dont les termes sont imprimés avec leurs contextes). Ces tâches sont aisément mécanisables, car elles consistent en des opérations élémentaires de tri, de sélection et de classement. Il existe aujourd'hui des programmes standard pour la fabrication d'index et de concordances sur des textes quelconques, littéraires, historiques, philosophiques ou juridiques. On voit bien que la nature des disciplines concernées n'est pas pertinente.

Au lieu d'étudier des textes en langage naturel, on peut travailler sur des représentations symboliques d'objets, d'images, ou même sur des représentations du contenu sémantique de certains textes. En effet, on peut représenter n'importe quel objet, par exemple une poterie, par un ensemble de symboles désignant chacun un trait distinctif de sa morphologie ou de sa technologie. On établit alors un index des traits distinctifs des objets étudiés. De même, on peut exprimer dans un métatexte, à l'aide d'un langage artificiel *ad hoc*, les traits sémantiques des diverses unités d'un texte; on obtient ainsi, non plus un index littéral des occurrences des mots du langage naturel, mais un index conceptuel, qui indique les notions rencontrées dans l'analyse sémantique du texte étudié.

On utilise aussi les calculateurs pour construire des modèles distinctifs de phénomènes humains, quels qu'ils soient. On distingue les modèles descriptifs (ou statiques) des modèles génératifs (ou dynamiques).

Nous entendons par le terme de modèle toute caractérisation d'un ensemble de faits qui peut être soumise aux opérations de diagnostic et/ou de simulation. Supposons que nous nous intéressons à un ensemble homogène de mythes et que nous disposons d'une bonne exégèse traditionnelle. Elle ne nous permettra probablement pas de diagnostiquer l'appartenance à cet ensemble d'un mythe que nous venons de découvrir, ni de fabriquer nous-même un "faux" présentant, même pour les spécialistes, toutes les caractéristiques des mythes en question. En d'autres termes, une description traditionnelle est insuffisamment explicite; elle ne remplit pas les conditions de diagnostic et de simulation.

Pour élaborer un modèle, on commence par dégager les unités distinctives de l'ensemble des mythes rassemblés. On établit ensuite les règles structurelles qui fondent la mise en relation de ces unités dans le corpus de mythes étudié. Le calculateur permet précisément de trouver mécaniquement les unités distinctives et les règles structurelles en fonction, naturellement, du programme fourni à la machine par le chercheur.

Alors que les modèles descriptifs sont essentiellement statiques, les modèles génératifs visent à simuler le fonctionnement des objets étudiés. Ainsi, par exemple, des psychologues ont élaboré des modèles dynamiques du comportement humain; d'autres chercheurs ont créé des modèles de composition musicale représentatifs d'une période, d'une école ou d'un compositeur.

En fait, les deux catégories de modèles sont étroitement liées: elles présentent deux aspects d'une même entreprise. Les modèles statiques décrivent d'une manière scientifique et différentielle un ensemble de phénomènes quelconques considérés au repos; les modèles dynamiques visent à engendrer ces phénomènes, ce qui est une manière de tester l'adéquation du modèle descriptif initial. On ne peut être sûr de la qualité d'un modèle que si on l'applique sur le plan descriptif et sur le plan génératif.

Le calculateur permet de tester l'adéquation d'un modèle de manière plus commode et plus rapide que les méthodes traditionnelles. On n'en conclura pas, comme on le croit généralement, que le calculateur n'est qu'un auxiliaire dans les sciences humaines. En effet, les exigences du calcul électronique aboutissent toujours à une remise en question des méthodes des disciplines où on l'applique. L'intérêt majeur de l'introduction de l'informatique dans les sciences humaines réside principalement dans cette remise en question fondamentale qu'elle impose à tous les chercheurs.

Examinons maintenant deux exemples d'application des calculateurs dans deux disciplines particulières: l'archéologie et l'histoire de l'art, qui présentent des cas privilégiés. D'une part, ces deux disciplines étudient des données

matérielles; de ce fait, l'élaboration du langage symbolique est plus facile que dans des disciplines comme la psychologie et la sociologie, qui étudient des phénomènes plus abstraits. D'autre part, on peut admettre, sans faire injure aux archéologues et aux historiens de l'art, que les opérations distinctives de leurs raisonnements sont simples et monotones; elles se ramènent toujours à des constitutions de séries, de groupes, de classes, etc. On peut considérer la sériation comme l'opération distinctive de l'archéologie et de l'histoire de l'art. Or, nous l'avons signalé, les calculateurs permettent d'effectuer rapidement et sûrement des opérations de tri.

Passons au second type d'application, un peu plus ambitieux: la construction de modèles. On constate que les modèles archéologiques sont essentiellement des ordinations de séries. Dans la première étape, on s'est contenté d'établir des séries élémentaires; dans la seconde, on constitue un système de séries, une typologie. Une typologie, en effet, en archéologie comme en histoire de l'art, n'est qu'un ensemble ordonné de séries particulières. On voit bien que ces deux disciplines, contrairement à une opinion communément répandue, peuvent bénéficier largement des applications du calcul électronique.

Le premier problème à résoudre, dans l'élaboration d'un modèle, est d'ordre linguistique ou sémiologique, puisqu'il s'agit de construire des langages descriptifs adéquats pour enregistrer l'information archéologique. Si l'on étudie une potiche, il faut représenter ses traits caractéristiques par des symboles que l'on communiquera au calculateur. Il semble, à première vue, que le langage spécialisé des archéologues constitue un instrument adéquat. Tel n'est pas le cas. En effet, il existe autant de langages descriptifs d'objets que de langues naturelles parlées par les archéologues: le japonais ne décrit pas un vase dans les mêmes termes que l'anglais ou le français. De plus, à l'intérieur même d'une langue donnée, on constate d'innombrables usages quant à la manière de décrire le même objet. Il est donc indispensable d'élaborer des langages artificiels, dits aussi langages analytiques, pour exprimer les caractéristiques des objets étudiés, quels qu'ils soient.

Dans la construction de modèles, on rencontre aussi des problèmes mathématiques. Si l'on veut demander à un calculateur de fabriquer une typologie, il faut préalablement poser les opérations constitutives du modèle dans un langage mathématique que la machine peut comprendre. Ces opérations, rappelons-le, effectuent le passage du chaos des séries initiales à un ensemble ordonné. Or, dans toute la littérature archéologique, on ne trouve pas d'explication satisfaisante sur ce passage de la description des objets à la construction du modèle. Les auteurs donnent bien quelques commentaires partiels sur les éléments qui les ont amenés à établir telle ou

telle typologie, mais ils ne proposent jamais un raisonnement complet et explicite qui fournisse toutes les informations nécessaires à l'élaboration d'un programme pour le calculateur. Ici encore, l'apport de l'informatique aux sciences humaines est déterminant, car elle oblige les chercheurs à rendre explicites des démarches jusqu'ici largement intuitives.

On constate que les écrits "scientifiques", en archéologie comme en histoire de l'art, ne résistent pas aux tests de diagnostic et de simulation dont nous avons parlé au début de cet exposé. Supposons qu'un enseignant donne un cours complet sur les vases étrusques sans en présenter la moindre image; ses étudiants seront probablement incapables de distinguer un vase étrusque d'un autre vase ou de fabriquer un "faux" satisfaisant. L'expérience montre que l'information scientifique transmise par la littérature archéologique réside principalement dans les reproductions et les exemples. Seule une toute petite partie de l'information nécessaire est fournie par l'appareil théorique et par les commentaires du spécialiste. D'ailleurs, quand on examine attentivement des ouvrages tenus pour exemplaires, on découvre toutes sortes d'imperfections, d'impropriétés, tant sur le plan sémiologique que logique. Or, il est indispensable que l'étudiant ou le lecteur puisse vérifier la qualité des informations qu'on lui fournit par des tests de diagnostic et de simulation. L'informatique met ainsi en évidence la nécessité d'une transformation radicale des modes d'expression généralement utilisés dans les sciences humaines pour communiquer l'information.

Nous n'avons donné jusqu'ici que des démonstrations négatives. C'est que nous ne possédons pas encore d'exemple spectaculaire d'élaboration, par la machine, d'un modèle assez ambitieux dans le domaine de l'archéologie. Nous ne disposons pas de corpus d'objets publiés avec suffisamment de soin pour que les problèmes sémiologiques puissent être considérés comme résolus; il faudrait que l'informaticien commence par passer cinq ou dix ans à rassembler le matériel nécessaire à sa démonstration. Aujourd'hui, il se trouve malheureusement que les informaticiens n'ont pas envie de se transformer en archéologues, ni les archéologues en informaticiens, ce qui explique les piétinements actuels. Mais, de plus en plus, de jeunes archéologues se mettent à l'étude des systèmes sémiologiques et mathématiques dont nous avons parlé; dans quelque cinq ans, ils auront sans doute élaboré des méthodes mécanisées de construction de modèles.

En conclusion, la question n'est pas de savoir si l'on peut appliquer le calcul électronique dans les sciences humaines ni si on doit le faire. On ne peut que constater l'amorce d'une petite révolution culturelle qui remettra radicalement en question les démarches traditionnelles, et examiner par quelle voie institutionnelle elle pourrait être accélérée. Actuellement, nos efforts

sont freinés par les barrières traditionnelles qui séparent les deux catégories de compétences nécessaires à notre entreprise: celle de l'archéologue et celle de l'informaticien. Il importe donc de précipiter la petite révolution institutionnelle qui permettra, dans les universités, à des archéologues ou à d'autres chercheurs des sciences humaines et à des informaticiens de travailler en commun.

Centre d'analyse documentaire  
pour l'archéologie  
Marseille

Jean-Claude Gardin