

Vers une sémiotique des espaces

Maryvonne LONGEART et Gilles ÉMOND

Abstract. The direct digitalization of textual documents makes their automatic manipulation and exploration possible but mostly at a superficial level. Our approach is different in that it seeks to capture their underlying conceptual structure within an essentially spatial framework such that the manipulation of this framework in itself should also constitute a manipulation of the knowledge contained within the original text. Our research rests largely on the hypothesis that any form of knowledge representation, be it done with the aid of computers or otherwise, is essentially a process of semiosis as defined by Charles Sanders Peirce. To this day, most efforts in the field of knowledge representation that are based on the works of Peirce, such as Sowa's conceptual graphs, expand upon his existential graphs but do not view these from a semiotic perspective, as did Peirce. Because of this, the essentially spatial nature of any such method of diagrammatization, through which they are able to function as true iconic signs, has been underestimated.

Keywords: Knowledge representation, semi-osis. **Mots-clés :** Représentation des connaissances, sémiosis.

La transposition directe des textes sur support électronique permet l'exploitation automatique mais superficielle des données textuelles. Si la gestion documentaire (traitement de texte, repérage automatique de l'information, concordances...) en est facilitée, l'analyse du contenu gnoséologique des textes y trouve peu de nouveaux appuis. L'analyse de texte assistée par ordinateur a souvent pris la forme de l'analyse lexicale statistique ou encore de l'utilisation de techniques d'hypertexte et de repérage automatisé de l'information. Notre approche se veut plus fondamentale en ce qu'elle propose de saisir la structure conceptuelle sous-jacente au texte matériel de telle façon

✉ M. LONGEART; Département d'informatique; Université du Québec à Hull; C.P. 1250, succ. B; Hull; Québec J8X 3X7 (Canada).

Fax : +1 819 773 1638

E-mail : longeart@uqah.quebec.ca

✉ G. ÉMOND; *Pareidos* Informatique; 101, rue St-Jean-Bosco; Hull; Québec J8Y 3G5 (Canada).

E-mail : emondg@uqah.quebec.ca

que la seule manipulation spatiale de cette structure explicite constitue un traitement efficace de la connaissance qu'il véhicule. L'ambition est de réaliser un outil susceptible de se prêter à la représentation et à la manipulation de systèmes conceptuels arbitrairement complexes. Posé en ces termes, le problème de l'accès au contenu du texte est pour l'essentiel un problème de représentation des connaissances au sens de l'intelligence artificielle¹. Notre recherche a été profondément influencée par l'hypothèse que la représentation des connaissances, en informatique comme dans tout autre domaine, est un processus de sémiologie tel que défini par Charles Sanders Peirce². Jusqu'à présent, les recherches en représentation des connaissances qui se sont inspirées des travaux de Peirce et plus particulièrement de ses graphes existentiels (une représentation diagrammatique de la logique) se sont peu intéressées à sa sémiologie. C'est peut-être ce qui explique le fait que la présentation proprement spatiale des graphes conceptuels de Sowa³ par exemple, un langage de représentation des connaissances inspiré pourtant des graphes existentiels de Peirce, soit souvent considérée comme accessoire et on va même jusqu'à lui préférer une présentation linéaire. Pour Peirce cependant, la spatialité des graphes existentiels, loin d'être un pur artifice de présentation didactique, était une caractéristique essentielle de sa théorie. Les graphes existentiels (et les graphes en général) sont des représentations topologiques. Il nous est apparu que l'ordinateur, avec ses capacités d'interface graphique et interactive, était l'outil par excellence pour explorer les conditions de possibilité et évaluer la puissance d'un langage de représentation des connaissances essentiellement spatial.

La représentation de textes sur ordinateur comme sémiologie

Dans le mythe biblique, Adam désigne et étiquette les autres créatures : le récit de la création est aussi un récit de l'origine du langage, langage vu essentiellement comme une nomenclature. À cette version traditionnelle du mythe, Umberto Eco, s'appuyant sur la sémiologie de Peirce, propose de substituer une version pragmatiste. Ainsi, dit Eco, « la cérémonie baptismale

¹ Une personne ne connaissant pas la problématique de la représentation des connaissances en intelligence artificielle pourrait vouloir consulter BRACHMAN and LEVESQUE (1985) qui contient une bonne revue de la littérature jusqu'en 1985 et reproduit des articles devenus des classiques dans le domaine.

² *Collected Papers of Charles Sanders Peirce*, in BURKS (1960).

³ SOWA (1984); MINEAU, MOULIN and SOWA (1993).

ne baptise pas des choses mais des contextes non pas des individus destinés à subir des histoires dont leur nom ignore tout, mais des histoires à la lumière desquelles nous pouvons découvrir la définition qui identifie leurs actions»⁴. Si un certain objet peut être nommé *rose*, c'est que le terme *rose* peut être mis par exemple dans le contexte du terme *fleur cultivée* qui peut être lui-même mis dans le contexte de toute une histoire qui raconte comment on cultive les roses... Cette nouvelle version de l'origine du langage a ceci de particulier qu'elle suggère d'une part que l'histoire et donc le langage, ou du moins un langage, préexiste à son instauration et d'autre part que la nomination n'est pas un diktat mais une inférence et un mode d'emploi. En ce sens, le récit révisé de l'activité nominative d'Adam n'est pas autre chose qu'une représentation mythique de la théorie de la sémiotique illimitée de Peirce selon laquelle il n'y a pas de commencement à la signification, il n'y a pas d'origine au langage, il n'y a que la reprise contextuelle, l'interprétation, d'une expression antérieure par (dans) une expression nouvelle qui devra être elle-même reprise, interprétée par une autre expression et ainsi de suite indéfiniment. La chose *finalement* nommée, ce que Peirce appelle l'objet dynamique, est l'aboutissement idéal du processus d'interprétation. Elle n'est pas là à attendre son nom pour qu'on puisse raconter son histoire. Son histoire se raconte, moyennant quoi son nom se forme (se transforme) à l'occasion de chaque contexte d'apparition. Il est important de noter cependant que la sémiotique illimitée de Peirce ne fait intervenir aucun sujet conscient puisque l'interprétant d'une expression n'est pas un interprète mais un contexte, donc une autre expression. De la sémiotique au sens de Peirce, nous retenons cette idée particulièrement pertinente pour nos recherches sur la représentation des connaissances en intelligence artificielle : comprendre une expression, c'est être capable d'en produire une interprétation et interpréter c'est accroître le signifié de l'expression par son insertion dans le contexte d'autres expressions. La représentation des connaissances et, plus particulièrement, la représentation de systèmes conceptuels sur ordinateur peut être vue comme un exercice de sémiotique et la recherche sur cet exercice, comme une branche de la sémiotique.

Cette perspective permet de mieux situer quelques essais de représentation informatique de textes philosophiques que nous avons réalisés en utilisant *code*, un logiciel de représentation des connaissances à base de cadres (*frames*)⁵. Notre objectif était d'abord d'explorer la possibilité d'utiliser des

⁴ ECO (1992), p. 306.

⁵ SKUCE, SHENKANG and BEAUVILLÉ (1989).

outils de représentation des connaissances développés en intelligence artificielle pour étudier sur ordinateur des systèmes philosophiques habituellement présentés sous forme de textes écrits. Face à une expérience de ce genre, la première réaction est souvent le scepticisme : si tant est qu'une telle représentation puisse être produite, serait-elle fidèle au texte original ? Cependant, si un texte n'existe pas indépendamment de son interprétation, la question n'est pas de savoir si la représentation serait *fidèle* au texte non interprété, car il nous manque pour cela le deuxième terme de la comparaison, mais si elle était intéressante; et elle serait intéressante si justement elle disait quelque chose de plus que le texte initial, si elle enrichissait le signifié. Cela n'implique pas que n'importe quelle interprétation est bonne et n'importe quelle représentation satisfaisante. Si le texte initial dont on est parti ne préexiste pas à toute interprétation, il préexiste, avec ses interprétations antérieures, à cette nouvelle interprétation et il la détermine. Les interprétations successives s'appuient sur un accord collectif qui fournit à chaque moment de la sémiologie une interprétation initiale qui, si elle n'est ni objective, ni définitive, est du moins intersubjective et en ce sens contraignante. On ne peut donc pas faire dire à un texte n'importe quoi. L'Habitue est notre garde-fou. Outre plusieurs petits textes courts et isolés, nous avons ainsi représenté la chaîne des déductions de l'*Éthique* de Spinoza, le système de concepts mis en place par Kant dans l'*Analytique transcendantale* (première partie de la *Critique de la raison pure*), la théorie des passions chez Hobbes, Descartes et Spinoza et certains éléments de la méthodologie de Hobbes⁶. Ces essais de représentation que nous avons faits en utilisant le système *code* ont montré que certaines parties des systèmes conceptuels, en particulier les classifications, se prêtaient mieux que d'autres à l'exercice de transposition. En effet, le présupposé qui a guidé la conception de *code* (et celui de la majorité des systèmes de représentation des connaissances en intelligence artificielle) est que les relations entre concepts sont pour l'essentiel des relations taxonomiques. Étant donnée l'importance des taxonomies dans les systèmes de concepts ainsi que le rôle primordial joué dans de nombreuses logiques, d'Aristote à Russell, par les principes de classification, cette hypothèse semble trouver sa justification. Cependant, le fait que certains aspects des systèmes conceptuels soient réfractaires à une représentation fondée sur ce présupposé en limite la pertinence. Malgré les limites de l'exercice de transposition, ou peut-être grâce à celles-ci, il est apparu que le principal

⁶ LONGEART and BOSS (1991); BOSS et LONGEART (1993); LONGEART, BOSS and SKUCE (1993).

avantage de la représentation d'un système conceptuel sur ordinateur n'était pas la possibilité (problématique ou triviale) de reproduire intégralement le texte, mais plutôt les moyens que cette transposition fournissait de présenter les rapports conceptuels dans une perspective différente de celle privilégiée par le parcours séquentiel d'un texte. En effet, la représentation informatique pouvait rendre activement présent de façon synchronique dans la représentation un ensemble de relations conceptuelles que le texte écrit ne pouvait présenter que diachroniquement dans le processus de lecture. Tout texte est, en ce sens, un mécanisme inférentiel complexe que le lecteur doit actualiser dans son contenu implicite, pour reprendre une formule de Umberto Eco. Or, approcher de cette limite où la représentation totalement explicite deviendrait à elle-même sa propre actualisation dynamique semblait être précisément ce que la représentation d'un système conceptuel sur ordinateur pouvait rendre possible. Cette représentation informatique d'un système conceptuel n'est donc pas équivalente à sa représentation textuelle et elle se distingue en cela radicalement du simple processus de digitalisation qui permet l'utilisation d'un logiciel de traitement de texte. La représentation informatique produite dans *code* est au texte écrit ce que le texte écrit est à une image. On pourrait dire que nous nous trouvons en présence de deux formes de représentation, de deux systèmes sémiotiques, le texte et l'image dans un cas, le texte et la représentation sur ordinateur dans l'autre, qui ne sont pas équivalents. C'est un exemple de ce que Émile Benveniste a appelé le principe de non redondance entre systèmes sémiotiques : deux systèmes sémiotiques de type différent ne peuvent être mutuellement convertibles. Il n'est pas vrai que deux systèmes de signes distincts peuvent saisir le *même* rapport de signification.

Le graphe comme représentation diagrammatique

La représentation d'un texte dans un système informatique de représentation des connaissances est donc un exercice de transposition entre systèmes sémiotiques différents. Les êtres humains sont particulièrement habiles à passer d'une forme de représentation à une autre et à envisager un problème selon les différentes perspectives qu'offre chacune de ces formes. Le même problème peut être décrit dans les mots du langage ordinaire, formalisé à l'aide de symboles logico-mathématiques ou présenté sous forme d'images ou de diagrammes. On est sensible aux différentes possibilités offertes par

chacune de ces représentations en terme d'opération et d'efficacité et la capacité même de faire varier les perspectives est souvent un élément essentiel du processus de résolution de problème. Dans *code*, le jeu de transposition-interprétation s'est révélé d'autant plus intéressant que le système offre deux types d'interface usager, un ensemble de fenêtres représentant les concepts et leurs relations sous forme de listes, et une fenêtre graphique représentant ces mêmes concepts et leurs relations sous forme de graphe. Il y a donc dans *code* plusieurs représentations dans la représentation. Si en informatique le choix d'une représentation est reconnu comme essentiel, peu de recherches ont été faites sur la coexistence à l'intérieur d'un même système de plusieurs représentations concurrentes ou complémentaires. En 1962, Herbert Gelernter avait présenté un petit programme de résolution de problème de géométrie plane dans lequel les problèmes étaient représentés à la fois axiomatiquement et diagrammatiquement et les deux représentations concouraient à la découverte efficace des preuves. Mais ces recherches sont restées relativement isolées. Dans le cas de *code*, s'il y a coexistence de deux représentations, l'une sous forme de listes, l'autre sous forme de graphe, c'est dans un sens plus faible que dans le système de Gelernter puisque la différence n'existe qu'au niveau de l'interface, les deux représentations étant soutenues, pour l'essentiel, par la même représentation interne. C'est ainsi que la représentation de la structure de l'enchaînement des propositions de l'*Éthique* de Spinoza dans *code* dépassait largement les capacités de représentation graphique (spatiale) du logiciel mais nullement ses capacités de traitement du graphe abstrait (le réseau sémantique) correspondant et aucune des fonctions essentielles du logiciel n'était perdue avec la fermeture de l'interface graphique. *code* n'est pas un système de représentation essentiellement graphique. Si la fenêtre graphique permet de visualiser les concepts et certaines de leurs relations, elle ne permet pas de les manipuler. La représentation sous forme de listes est donc privilégiée. Malgré ce caractère statique et limité de la représentation graphique dans *code*, le graphe, quand il est accessible, présente visuellement certains aspects des structures conceptuelles que le parcours du texte, même sous forme de listes n'aurait pas permis de saisir facilement. Ainsi, par exemple, la différence de *forme* des arbres représentant la classification des passions chez Hobbes, Descartes et Spinoza n'a pas d'équivalent textuel et si elle peut s'expliquer, elle ne peut pas s'*exprimer* directement dans un texte. Cela tient au fait qu'un graphe n'est pas un signe isolé mais un signe complexe qui n'est pas facilement décomposable en un ensemble de traits pertinents. (Par opposition à une simple variante, il faut entendre par trait pertinent une caractéristique élémentaire de l'expression telle que si on la modifie, le

contenu de l'expression s'en trouve également modifié). On ne peut pas mettre en relation les éléments de la représentation graphique avec des éléments correspondants de la représentation textuelle parce que les éléments de la représentation graphique ne sont pas isolés, et peut-être même pas isolables.

Dans la classification peircéenne des signes, les graphes sont des diagrammes et les diagrammes sont des icônes. La classification des signes en symbole, indice et icône est l'apport le plus connu de la sémiologie de Peirce. Un symbole dénote son objet en vertu d'une convention (c'est le cas des mots de la langue par exemple); un indice dénote son objet en vertu d'un rapport de contiguïté (le symptôme d'une maladie par exemple); un icône dénote son objet en exhibant la même qualité ou la même configuration de qualités que son objet (un schéma par exemple). Tout signe est dans une certaine mesure à la fois symbole, index et icône mais, dit Peirce, « les plus parfaits des signes sont ceux dans lesquels les caractères iconiques, indexicaux et symboliques se marient le plus également possible »⁷. Peirce distingue trois types d'icônes, les images, les diagrammes et les métaphores. Un diagramme, en tant que signe iconique, se définit comme un genre de signe dans lequel les relations entre les parties d'une chose sont représentées par des relations analogues entre les parties du signe lui-même. Cependant, si le graphe est diagrammatique, ce n'est pas en vertu d'une vague ressemblance « naturelle » avec son objet mais en vertu d'une transformation topologique conventionnellement réglée qui met en relation les points de l'espace de l'expression et les points d'un modèle de rapport qui n'est pas nécessairement lui-même d'ordre spatial. Comment peut opérer cette mise en relation et que devrait être un langage de représentation qui serait diagrammatique au sens de Peirce ?

Nous venons de voir qu'un graphe, comme diagramme, est une représentation topologique qui exploite les deux dimensions du plan selon un code iconique qui associe conventionnellement des configurations spatiales (des ensembles de points sélectionnés selon certains critères de contiguïté) à des caractéristiques de l'objet représenté de telle sorte que la structure du diagramme comme signe (l'organisation de ses parties) corresponde à la structure de l'objet. Cependant, si globalement il est relativement aisé d'établir cette correspondance (on va pouvoir dire par exemple qu'un rectangle relié à un autre rectangle par un segment de droite représente une relation entre deux concepts), le code graphique dont on dispose est relativement pauvre et le graphe n'est vraiment représentatif (significatif) que parce qu'il est annoté. Par exemple, dans la représentation graphique de *code*, un nœud isolé (non

⁷ *Collected Papers of Charles Sanders Peirce*, in BURKS (1960), 4.448.

relié à d'autres nœuds) représentant un concept est une forme élémentaire, un simple rectangle annoté. On peut donner de cette figure un nombre infini de variantes. La taille, la position sur la page, la forme même, peuvent varier. Le rectangle peut être remplacé par une autre figure, à la limite, il peut même disparaître (une option offerte par l'interface graphique de *code*). C'est là un cas extrême puisque toutes les variantes iconiques sont possibles et l'ensemble des traits pertinents est vide. On ne s'étonnera donc pas que la partie proprement iconique du nœud ne représente rien en elle-même. Le fardeau de la représentation est porté tout entier par l'annotation, donc par le texte. Ainsi, c'est le nom inscrit dans le rectangle représentant un nœud ou accolé au segment de droite représentant une arête qui porte l'essentiel du contenu représentatif du graphe. Or ce nom n'est pas un trait graphique mais textuel. Par ailleurs, la disposition spatiale des éléments de la représentation dans le plan laisse une très grande latitude quant à la forme de la représentation et seulement quelques-unes de ces dispositions légitimes (permises par la syntaxe de la représentation) sont effectivement expressives. Ainsi, il suffit de changer la longueur des segments de droite représentant les relations dans le graphe des passions de Spinoza pour que la belle régularité du diagramme disparaisse rendant toute comparaison immédiate avec d'autres graphes impossible. Autrement dit, les codes graphiques sont trop faibles. Les traits pertinents dont on dispose sont à la fois trop peu nombreux par rapport aux variantes permises et pas assez fins (ils ne permettent pas un niveau de détail suffisant).

Les idées principales qui se dégagent de cette analyse et que nous travaillons à mettre en œuvre sont les suivantes :

- 1) Un langage essentiellement spatial de représentation des connaissances doit permettre de ramener les relations *logiques* à des relations *topologiques* (ce qui était déjà le programme des graphes existentiels de Peirce). Le code iconique doit être tel qu'il permette de saisir les deux caractéristiques essentielles de l'espace, la contiguïté et la multidimensionnalité, et ce, à différents niveaux de détail.
- 2) La représentation doit pouvoir se passer d'annotation textuelle. Or, on ne peut faire l'économie du texte que si le code iconique dont on dispose est suffisamment fort. Pour cela, il faut pouvoir déterminer quels sont les traits significatifs dans la forme de la représentation et ces traits significatifs doivent prévaloir sur les variantes non significatives.
- 3) Le langage doit permettre non seulement de représenter diagrammatiquement n'importe quel ensemble de connaissances mais aussi de le manipuler. Le diagramme doit être tel que l'on puisse guider et examiner sa trans-

formation dans un autre diagramme par une série d'étapes déterminées et observables. Ces opérations de transformation doivent être spécifiées.

Dans ce qui suit, nous illustrerons certaines de ces idées et donnerons un aperçu général de notre système de représentation diagrammatique des connaissances.

Éléments d'une représentation diagrammatique des connaissances

Come on, my Reader, and let us construct a diagram to illustrate the general course of thought; I mean a System of diagrammatization by means of which any course of thought can be represented with exactitude.
C.S. Peirce (*Collected Papers*, 4.530)

Dans le cadre du présent article, donnons-nous un objectif plus modeste. À défaut d'un diagramme illustrant le cours général de la pensée, construisons un diagramme illustrant un triplet dont le troisième terme est z , le premier x et le deuxième y . Remarquons que le groupe de mots :

un triplet dont le troisième terme est z , le premier x et le deuxième y (1)

constitue déjà une première représentation, une représentation textuelle du triplet. Cette représentation est principalement, pour ne pas dire exclusivement, symbolique au sens de Peirce. En effet, c'est par pure convention que (1) représente un triplet. Il n'y a rien dans la *structure* de la représentation textuelle qui corresponde à la *structure* du triplet.

Dans le symbolisme mathématique au contraire, la position relative de x , y et z sur la feuille représente l'ordre des coordonnées :

(x,y,z) . (2)

On voit que dans cette deuxième représentation, la structure du symbole (la disposition de ses parties) correspond à la structure de son objet (la relation d'ordre entre les coordonnées du triplet). C'est en ce sens que d'une façon générale Peirce affirme qu'une équation mathématique est, dans une certaine mesure, un diagramme de la forme de la relation qu'elle représente, diagramme qui repose, comme tout diagramme, sur un certain nombre de conventions, mais aussi sur une certaine iconicité⁸. L'iconicité de la représentation (2) a cependant l'inconvénient de reposer sur une simple habitude de

⁸ *Collected Papers of Charles Sanders Peirce*, in BURKS (1960), 4.530.

lecture-écriture (de gauche à droite) et une convention qui associe l'ordre de lecture-écriture des symboles compris entre les deux parenthèses à l'ordre de l'ensemble ordonné correspondant. La relation d'ordre n'est pas à proprement parler figurée, mais plutôt désignée. En ce sens, le symbole (x,y,z) est plus indexical qu'iconique.

La représentation (2) est elle-même une abréviation d'une représentation ensembliste plus primitive :

$$\left\{ \left\{ \{x\}, \{x,y\} \right\}, \left\{ \{x\}, \{x,y\}, z \right\} \right\} \quad (3)$$

En effet, un couple en tant qu'ensemble ordonné est équivalent à un ensemble contenant deux éléments dont l'un est lui-même un ensemble contenant la première coordonnée et dont l'autre est un autre ensemble contenant la première et la deuxième coordonnées. Ainsi,

$$\begin{aligned} (x,y) &= \{ \{x\}, \{x,y\} \}, \\ (x,y,z) &= ((x,y),z) = \{ \{(x,y)\}, \{(x,y),z\} \}, \\ \{ \{(x,y)\}, \{(x,y),z\} \} &= \left\{ \left\{ \{x\}, \{x,y\} \right\}, \left\{ \{x\}, \{x,y\}, z \right\} \right\}. \end{aligned}$$

Si la représentation ensembliste semble à première vue plus difficile à saisir, elle est cependant plus analytique et en ce sens, supérieure à la représentation symbolique abrégée (2). C'est la structure de la relation d'ordre elle-même qui est saisie dans (3), indépendamment de la contingence de l'ordre de lecture ou d'écriture des éléments. Pourtant, la *place* de ces éléments, c'est-à-dire *là* où ils apparaissent, compte d'une façon plus fondamentale encore dans la représentation (3) que dans la représentation (2). En effet, c'est la co-occurrence des éléments dans un ensemble et l'apparition concurrente du même élément dans des ensembles différents éventuellement imbriqués qui ordonnent les éléments entre eux. Cependant, la représentation ensembliste permet de symboliser les éléments d'un ensemble mais pas les places qu'ils occupent comme telles. Les notions de place et d'adresse ne sont pas des notions de la théorie des ensembles et la structure correspondante n'est donc pas mise en évidence dans la représentation (3).

Pour représenter les places et leurs relations, il faut se donner d'autres moyens. Nous allons donc enrichir la représentation ensembliste d'un nouveau symbole, une boîte représentant une place et pouvant elle-même contenir d'autres boîtes représentant des places englobées dans la première et ainsi de suite. Ainsi, $\{x\}$, $\{x,y\}$ et $\{\{x\},\{x,y\}\}$ sont représentés respectivement par

$$\begin{array}{c} \boxed{x} \\ \boxed{x, y} \\ \boxed{\boxed{x}, \boxed{x, y}} \end{array}$$

et enfin,

$$\left\{ \left\{ \{x\}, \{x, y\} \right\}, \left\{ \{x\}, \{x, y\}, z \right\} \right\}$$

est représenté par :

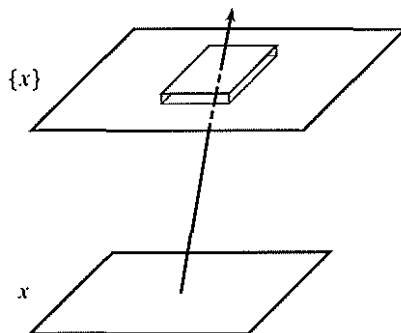
$$\left\{ \left(\left(\left(\boxed{x_1} \right)_4, \left(\boxed{x_2, y_3} \right)_5 \right) \right)_6 \right)_9, \left(\left(\left(\boxed{x_1} \right)_4, \left(\boxed{x_2, y_3} \right)_5 \right), z \right)_7 \right)_8 \right)_{10} \quad (4)$$

Ce faisant, nous sommes passés d'une représentation linéaire à une représentation graphique bidimensionnelle. Nous avons donc ajouté non seulement un nouveau symbole, mais aussi une nouvelle dimension à notre représentation permettant de visualiser la structure des places qui échappait au symbolisme ensembliste. La notion intuitive de place devient du même coup un terme technique dont il va falloir préciser l'acception. Si nous n'avons pas eu besoin d'introduire un nouveau vocabulaire à l'appui des représentations (1), (2) et (3), c'est que nous pouvions supposer que le lecteur connaissait déjà le contexte de ces représentations, à savoir respectivement, le langage ordinaire, le symbolisme mathématique et au moins quelques rudiments de la théorie des ensembles. Il n'en va pas de même pour notre quatrième représentation. Il est donc nécessaire de préciser ce que nous entendons par une *place* dans cette représentation. Nous supposons connue la notion d'adresse telle qu'elle est utilisée en informatique et plus généralement en théorie de l'information : un signe sous lequel est entrée une information et qui permet de retrouver son emplacement sur un support. Appelons *place* un lieu adressable distinct de tout autre lieu adressable. Rien n'est dit de la situation d'une place par rapport à une autre sinon que deux places qui ont la même adresse sont indiscernables. Appelons *place immédiate* une place occupée par une entité, ici une coordonnée du triplet, dans un ensemble dont elle est un élément. Cet ensemble peut lui-même être inclus comme élément dans un autre ensemble dont il occupe à son tour une place immédiate. Par exemple, x occupe une place immédiate dans l'ensemble $\{x\}$ et l'ensemble $\{x\}$ occupe une place immédiate dans l'ensemble $\{\{x\}\}$. À une place immédiate dans un premier ensemble correspond une *place médiate* distincte dans un second ensemble

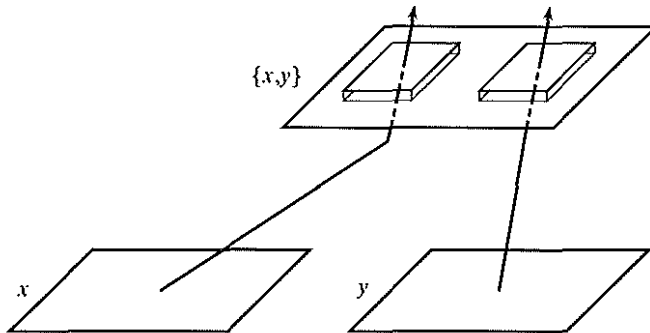
comprenant le premier ensemble comme élément. Par exemple, x occupe une place médiate dans $\{\{x\}\}$, une place distincte de, mais néanmoins déterminée par, la place immédiate que x occupe dans $\{x\}$. On dira que la place occupée par x dans $\{\{x\}\}$ est *englobée* par la place immédiate occupée par $\{x\}$ dans $\{\{x\}\}$ et elle est représentée par une boîte imbriquée dans la boîte représentant cette place immédiate de $\{x\}$ dans $\{\{x\}\}$. Dans la représentation (4), les places immédiates sont numérotées. Remarquons que l'entité x occupe une place immédiate notée (1) dans le singleton et une place immédiate notée (2) dans l'ensemble $\{x,y\}$ et que ces deux places déterminent quatre places médiates dans le triplet, les places dont les adresses sont respectivement,

$$\begin{aligned} &\langle 9\langle 6\langle 4\langle 1 \rangle \rangle \rangle \rangle, \\ &\langle 9\langle 6\langle 5\langle 2 \rangle \rangle \rangle \rangle, \\ &\langle 10\langle 7\langle 4\langle 1 \rangle \rangle \rangle \rangle, \\ &\langle 10\langle 7\langle 5\langle 2 \rangle \rangle \rangle \rangle. \end{aligned}$$

L'organisation des places symbolisée par l'imbrication des boîtes dans la représentation (4) met en évidence un élément de structure important mais complique indûment sa représentation. Une représentation faisant intervenir une troisième dimension permet d'éliminer les parenthèses et les virgules et fait ressortir mieux encore la structure du triplet à l'aide d'un symbolisme purement spatial à la fois plus économique et plus uniforme. Dans cette dernière représentation, les symboles non graphiques sont éliminés. L'entité x est représentée par une surface plane. Cette surface est reliée à une autre surface plus petite, que nous appellerons un *nœud* par analogie avec un nœud d'un graphe, qui se détache d'une troisième surface et qui constitue avec elle la représentation de l'ensemble $\{x\}$:



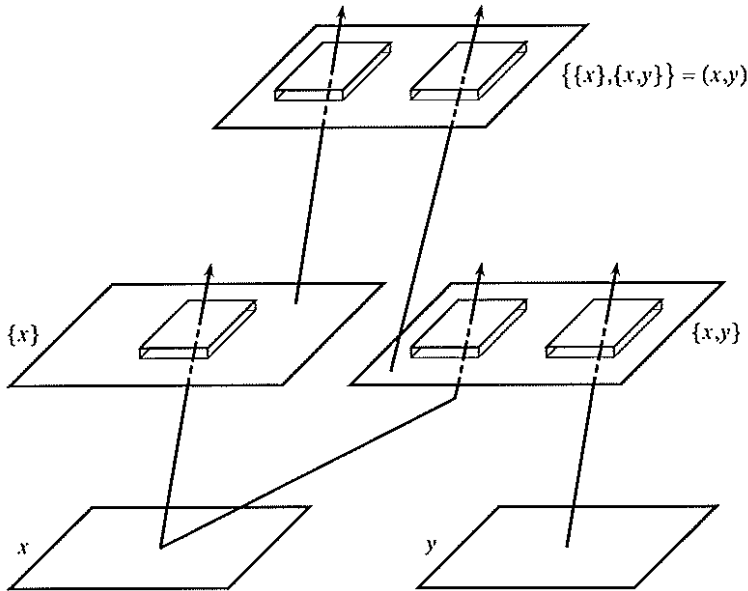
Les boîtes aussi bien que les lettres et les accolades de la représentation (4) sont représentées ici par de simples surfaces puisque la surface vide sous le nœud correspond à la place de x dans $\{x\}$ numérotée (1) dans la représentation (4). La flèche qui relie le nœud-signe à la surface dont il tient lieu semble même ici superflue mais ce n'est pas toujours le cas, comme l'illustre la représentation de l'ensemble $\{x,y\}$:



Cette représentation est éminemment diagrammatique au sens où nous l'avons défini précédemment. Il s'agit bien d'une transformation topologique conventionnelle qui associe les points de l'espace de la représentation aux points d'un autre modèle de rapport (ici, un rapport non spatial d'appartenance d'un élément à un ensemble). Cependant, ce qui fait l'intérêt de cette représentation n'est pas en soi l'économie des moyens dont elle use ni même son caractère purement diagrammatique, mais le fait qu'elle soit à la fois une représentation et la représentation du processus de représentation, donc une représentation de la sémiotique elle-même. Prenons par exemple la représentation de $\{x\}$. On peut établir une correspondance entre les trois surfaces représentant respectivement l'expression de x dans $\{x\}$, x lui-même et l'ensemble $\{x\}$ et la trilogie Signe/Objet/Interprétant telle que définie par Peirce⁹. Dans notre représentation graphique, la surface représentant x (et en tant que telle déjà un signe de x) est l'Objet, la surface plus petite à laquelle elle est liée par la flèche est le Signe de cet Objet (le signe du signe de x) et la troisième surface dont le nœud se détache est l'Interprétant du Signe

⁹ « Un signe, ou *Representamen* est un Premier qui entretient avec un Second, appelé son *Objet*, une telle véritable relation triadique qu'il est capable de déterminer un Troisième, appelé son *Interprétant*, pour que celui-ci assume la même relation triadique à l'égard du dit Objet que celle entre le Signe et l'Objet. » *Collected Papers of Charles Sanders Peirce*, in BURKS (1960), 2.229.

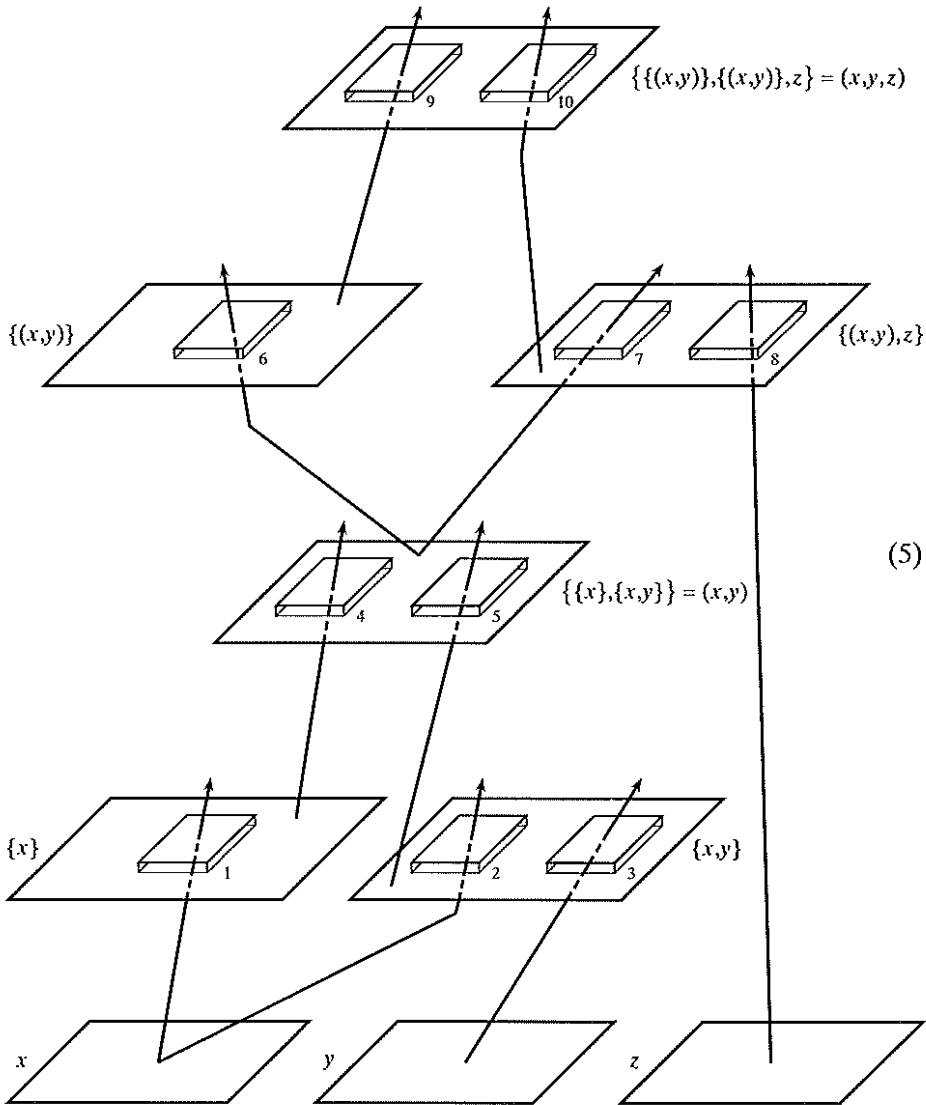
qui peut être lui-même repris comme Objet d'un autre Signe pour un autre interprétant, comme ce serait le cas dans la représentation du couple (x,y) , soit la représentation de $\{\{x\},\{x,y\}\}$:



De même que dans la théorie du signe chez Peirce, l'Objet, le Signe et l'Interprétant sont eux-mêmes des signes (ils ne sont pas de nature intrinsèquement différente), de même, dans cette représentation spatiale, ce sont tous des surfaces distinctes par leur rôle dans le processus de sémiotique et non leur substance. Ce rôle est double. En effet, à cause de la différence qu'il y a entre la relation qu'un signe entretient avec son objet et celle qu'il entretient avec son interprétant, deux sortes de relations sont nécessaires dans le système : des relations illatives qui permettent le passage d'un signe à son interprétant et qui sont, comme leur nom l'indique, des relations de contiguïté (coprésence); et des relations indicatives qui permettent d'associer un signe à son objet (expression)¹⁰. L'exemple très simple développé ci-dessus montre comment une représentation spatiale prend en charge ces deux types de relation. Une étendue plane qui accueille les expressions de plusieurs étendues sous-jacentes met ces expressions, qui sont chacune en relation indicative avec l'objet qui

¹⁰ *Collected Papers of Charles Sanders Peirce*, in BURKS (1960), 4.375.

les émet (type de relation symbolisé par la flèche), en relation illative de contiguïté les unes avec les autres. Il nous reste à donner la représentation complète du triplet produite de cette manière :



Si on se rappelle les exigences énoncées plus haut quant à la représentation spatiale des connaissances, on voit que cette cinquième représentation

satisfait les deux premières, à savoir, ramener une relation logique à des relations topologiques en jouant sur la contiguïté et la multidimensionnalité de l'espace et s'être dispensé des annotations textuelles. Si à chaque niveau de la représentation (5) on a pris la peine d'inscrire les formules correspondantes dans la représentation (3) et de numéroter les surfaces qui correspondent aux places immédiates de la représentation (4), ces commentaires ne servent qu'à faciliter la comparaison des trois représentations.

Sémiotiques des espaces virtuels

On remarquera que, dans la représentation (5), x (de même que y et z) ou plus exactement le plan représentant x dans la représentation (5), n'a pas lui-même de place numérotée. C'est dire qu'il ne correspond à aucune place dans la représentation (4). Quelle est donc sa place dans la représentation (5)? Tout comme l'expression de ce plan (le nœud) occupe une place (la place (1)) dans le plan représentant $\{x\}$, le plan x lui-même occupe une place dans l'expression de cette relation d'expression et qui n'est autre que l'espace de cette relation d'expression. Autrement dit, il y a un *espace virtuel* qui accueille les trois plans x , y et z et ces trois plans occupent une place dans cet espace. Interprétant, tout comme l'expression de ces plans (les nœuds) occupent des places dans les plans Interprétants. Mais alors, n'est-ce pas qu'ils sont eux-mêmes les signes d'autres objets, comme le voudrait d'ailleurs la sémosis sans commencement ni fin, et ces objets ne devraient-ils pas être aussi représentés? Il faut pourtant que la représentation commence, et elle commence effectivement dans notre exemple avec les trois plans représentant x , y et z . Est-ce un commencement satisfaisant? Pouvons-nous sans un renvoi explicite à un ailleurs où x , y et z existeraient avec toutes leurs déterminations, indépendamment de la représentation, parler de x , y ou z dans notre représentation? La représentation est-elle incomplète (par opposition à simplement non complétée) parce qu'elle ne nous permet pas de représenter cet x , y ou z ? De fait, rien ne distingue encore x , y et z , ou plutôt leur signe, dans la représentation (5) sinon leur plan respectif dans la représentation. Allons plus loin et considérons un espace propre à chacun de ces plans, les espaces x , y et z , et disons que ces espaces sont compris dans les espaces propres des plans Interprétants qui accueillent leur expression. Nous constatons que c'est encore par les places distinctes qu'ils occupent dans un espace englobant que nous distinguons les espaces x , y et z les uns des autres. C'est là une indication à la fois de la force et de la faiblesse

de cette représentation. Sa force est de montrer clairement que l'objet dont on voit le signe dans la représentation n'est pas un objet extérieur à la représentation mais un *objet virtuel* qui renvoie à tout ce à quoi il participe dans la représentation, son plan et tous ses nœuds dans les plans d'autres objets virtuels, dans l'espace virtuel de la représentation. Sa faiblesse est de ne pas fournir les moyens de déterminer ces plans de telle sorte que leurs formes (et pas seulement leur situation) diffèrent. En effet, si l'on prend la peine de distinguer les plans de x , y et z , c'est qu'éventuellement leurs différences intrinsèques plutôt que leur différence simplement positionnelle comme première, deuxième et troisième coordonnées du triplet, nous intéressent. Mais quelles transformations de et dans nos espaces virtuels peuvent ménager de telles divergences de forme ? Ces transformations ne peuvent se réduire à des mises en relation de contiguïté et d'indication, il faut ajouter une troisième sorte de mise en relation : celle de la forme à elle-même. Plus exactement, celle de la forme à son histoire. Pour cela, il faut une autre sorte de dimension, à côté des dimensions du plan et des dimensions verticales : des dimensions temporelles. Le pluriel n'est pas ici un abus de langage. Nous avons dans notre système¹¹ dont la représentation (5) n'est qu'une approximation assez lointaine, trois sortes de dimension (planaire, verticale et temporelle) pouvant chacune comprendre plusieurs dimensions. Nous travaillons donc avec des volumes spatio-temporels (plus exactement des hyper volumes) et non des surfaces planes. Cependant, pour que la représentation puisse être visualisée et facilement manipulée interactivement, les dimensions temporelles sont traitées comme une autre sorte de dimension spatiale et seule la *trace* des formes, une projection bidimensionnelle du volume multidimensionnel, a besoin d'être représentée à l'écran de l'ordinateur. Mais si les espaces de x , y et z peuvent se différencier par leur trace résumant l'histoire de leur transformation, il en sera de même pour chaque espace de la représentation et en particulier pour ces espaces qui, en englobant d'autres, les interprètent. Nous avons alors la possibilité non seulement d'avoir des Objets de forme différente, mais aussi différentes sortes d'Interprétants¹².

Si l'on voulait résumer en une simple question les préoccupations de Peirce telles qu'elles s'expriment dans toute son œuvre, de ses études phénoménologiques (ou plutôt phanéroscopiques) à ses études sémiologiques et

¹¹ ÉMOND et LONGEART (1994).

¹² C'est ainsi que l'on peut former dans le système des opérateurs comme ceux de la logique propositionnelle par exemple.

logiques, cette seule question serait « comment toute synthèse est-elle possible ? » Autrement dit, comment passe-t-on d'une idée à une autre, comment un signe qui renvoie à un autre signe (son Objet) se transforme-t-il en un troisième signe (son Interprétant) ? Comment saisir le cours même de la pensée ? Pour passer d'un diagramme comme celui suggéré dans notre exemple au système de diagrammatisation complet préconisé par Peirce, il faut reconnaître l'existence simultanée d'une multitude d'univers virtuels, celui que nous avons évoqué n'en étant qu'un parmi d'autres. Ces univers virtuels sont tous situés les uns par rapport aux autres, comme dans un paysage complexe, une sorte de *méta-univers*. L'univers virtuel originel, l'univers virtuel le plus primitif, occupe le site origine de ce méta-espace et tous les sites occupés par les autres univers dérivables à partir du premier sont situés par rapport à cette origine selon le nombre et le type de transformations qu'il faut faire subir à cet univers originel pour les obtenir. Muni donc d'un univers premier le plus simple qui soit, d'un ensemble de transformations et des conditions de leur applicabilité, tout ce méta-paysage s'ouvre à des explorations interactives, dynamiques et comparatives, à condition bien sûr que l'ordinateur puisse ouvrir les chemins et prendre en charge l'onéreux des calculs détaillés nécessaires¹³.

Cette recherche est subventionnée par le CRSNG (Conseil de recherche en sciences naturelles et en génie du Canada), par les fonds internes de recherche de l'Université du Québec à Hull et par les compagnies ALMERCOS et PAREIDOS Informatique.

Bibliographie

- BRACHMAN, (Ronald J.) and LEVESQUE, (Hector J.) : 1985, *Readings in Knowledge Representation* (Los Altos, CA : William Kaufman).
- BOSS (Gilbert) et LONGEART (Maryvonne) : 1993, « Intelligence artificielle ou philosophie sur ordinateur », *Revue de synthèse*, 4, 2, pp. 255–279.
- BURKS (Arthur) : 1960, *Collected Papers of Charles Sanders Peirce*. (Cambridge : Harvard University Press).

¹³ Notre projet visait initialement des applications dans le domaine de l'analyse de texte assistée par ordinateur, mais il est apparu que des applications plus larges pouvaient être envisagées dans le domaine de la modélisation et du prototypage de bases de données orientées objet, particulièrement en géomatique. Par ailleurs, la représentation et la manipulation en temps réel de structures conceptuelles complexes par le biais d'une interface graphique intelligente exige le recours à l'exécution de processus concurrents sur un matériel à architecture hautement parallèle. C'est dans ce contexte que nous mettons en œuvre notre système sur un ordinateur parallèle de type AVX2.

- ECO (Umberto) : 1992, *Les limites de l'interprétation*. (Paris : Grasset).
- ÉMOND (Gilles), LONGEART (Maryvonne) : 1994, *Document de travail sur les espaces sémiotiques* (Département d'informatique, Université du Québec à Hull, « Rapport de recherche RR-94/03-02 »).
- LONGEART (Maryvonne), BOSS (Gilbert) : 1991, "A Critique of Taxonomical Approach in Structured Object Representation of Abstract Conceptual Systems", in DIDAY (E.) and LECHEVALLIER (Y.) : 1991, *Symbolic-Numeric Data Analysis and Learning* (New York : Nova Science Pub), pp. 359-371.
- LONGEART (Maryvonne), BOSS (Gilbert) and SKUCE (Douglas) : 1993, "Frame-Based Representation of Philosophical Systems Using a Knowledge Engineering Tool", *Computers and the Humanities*, 27, 4, pp. 261-275.
- MINEAU (Guy), MOULIN (Bernard) and SOWA (John) : 1993, *Conceptual Graph for Knowledge Representation* (Springer-Verlag).
- SKUCE (Douglas), SHENKANG (Wong) and BEAUVILLÉ (Yves) : 1989, *A Portable Generic Knowledge Acquisition Environment* (Computer Science Department, University of Ottawa, "Technical Report TR-89-11").
- SOWA (John) : 1984, *Conceptual Structures: Information Processing in Mind and Machine* (Reading, MA : Addison-Wesley).