Ontologies computationnelles et neuropsychologie

Xavier Aimé

INSERM, U1142, LIMICS, F-75006, Paris, France ;
Sorbonne Universités, UPMC Univ. Paris 06, UMR_S 1142, LIMICS, F-75006, Paris, France ;
Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), F-93430, Villetaneuse, France.
xavier.aime@inserm.fr

Abstract. La mémoire sémantique, composante de la mémoire à long terme, est la mémoire des mots, des idées, des concepts. Certaines démences, avec au premier rang d'entre elle la démence sémantique, et dans un moindre mesure la maladie d'Alzheimer, se traduisent entre autre par une atteinte de la mémoire sémantique. Après nous être intéressé aux modèles classiques de la mémoire sémantique, nous proposons dans cet article d'utiliser le modèle des ontologies computationnelles, modélisation formelle et relativement fine provenant de l'Intelligence Artificielle, pour étudier (et évaluer) la mémoire sémantique.

Keywords: Ontologie, neuropsychologie, mémoire sémantique, modèles.

1 Introduction

Certaines démences, avec au premier rang d'entre elle la démence sémantique, et dans un moindre mesure la maladie d'Alzheimer, se traduisent entre autre par une atteinte de la mémoire sémantique. Cette dernière, composante de la mémoire à long terme, permet l'acquisition et la rétention de connaissances générales (factuelles) sur le monde, sans référence nécessaire aux conditions d'acquisition [17]. C'est la mémoire des mots, des idées, des concepts. L'objet de cet article s'articule autour de la notion de connaissance, de celle contenue dans notre mémoire sémantique (en collaboration avec la mémoire épisodique) à sa modélisation en Intelligence Artificielle (IA). Ainsi, nous nous intéresserons plus précisément au domaine des ontologies computationnelles, qui sont des « spécifications explicites d'une conceptualisation partagée » d'après la définition consensuelle de Grüber [11]. La connaissance propre à un individu est issue principalement de son mode de perception, de représentation et d'analyse de la « réalité ». Une ontologie computationnelle correspond à une modélisation formelle, manipulable et intelligible – par des agents tant logiciels qu'humains – de connaissances propres à un domaine particulier et à un ou plusieurs individus donnés. Une ontologie computationnelle peut donc être vue comme une formalisation de représentations conceptuelles mentales. Il existe aujourd'hui différentes modélisation de la mémoire sémantique : distribuée, hiérarchique, en réseau ou encore computationnelle. Nous proposons dans cet article d'utiliser le modèle des ontologies computa-

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

tionnelles, modélisation formelle et relativement fine, pour étudier (et évaluer) la mémoire sémantique.

La suite de cet article est organisée comme suit. Nous présentons tout d'abord la mémoire sémantique et ses différentes modélisations dans la section 2, puis les ontologies computationnelles issues de l'IA dans la section 3. Nous proposons dans la section 4 de combiner ontologie computationnelle et modèle de la mémoire sémantique dans le cadre de la neuropsychologie.

2 La mémoire sémantique et sa modélisation

Comment nos connaissances sont-elles stockées : sous forme de prototypes, d'images mentales, de concepts, de propriétés ? Comment sont-elles structurées : en réseau, de manière hiérarchique ou distribuée ? Tulving [17] sépare nos connaissances en deux entités qui, si elles sont distinctes, sont pourtant fortement inter-reliées : la mémoire sémantique et la mémoire épisodique. Afin d'illustrer la différence entre ces deux types de mémoire et leur imbrication, nous pouvons dire – par exemple – que le souvenir du café que vous avez pris lundi dernier au café Les Frangines à Paris relève de la mémoire épisodique. Par contre, ce qu'est un café, le fait que Paris soit la capitale de la France, ou encore le fait qu'un bistrot soit un lieu où on sert du café, relèvent de la mémoire sémantique. La mémoire sémantique est un ensemble des représentations conceptuelles dont nous disposons sur le monde qui nous entoure et sur nous-mêmes (sémantique personnelle), ainsi que les mots qui les indicent [15]. Elle donne du sens à nos perceptions. Elle regroupe les connaissances générales non personnelles, le sens des mots, les savoirs sur les objets / animaux / végétaux, les savoirs sur ce qui n'existe qu'en un seul exemplaire (personne, lieu, etc.), les connaissances livresques (géographie, histoire, etc.) [3]. Il y aurait par ailleurs une dissociation entre les concepts manufacturés/biologiques, manipulables/non manipulables, suivant les particularités fonctionnelles ou descriptives (perçues par les sens) [13]. La mémoire sémantique peut être modélisée de différentes manière~: distribuée, hiérarchique, en réseau ou encore computationnelle. Le *modèle distribué* est fondé sur l'idée que les concepts sont représentés sous la forme d'ensembles de propriétés (qui peuvent être de type descriptif ou fonctionnel). Ces propriétés, pour chaque concept, peuvent posséder un degré définitoire différent. Le sens d'un concept est donné par un ensemble de propriétés primitives. Dans cette vision ensembliste, les relations entre concepts sont les propriétés partagées. L'inconvénient de ce modèle est la définition même de ces propriétés. Le *modèle hiérarchique*, sorte de modèle en réseau, est fondé sur l'idée d'économie cognitive. Élaborée par Collins & Quillian [7], elle repose sur le temps de réponse aux questions de type « X est-il un Y » avec comme hypothèse que plus ce temps de réponse est grand, plus le chemin séparant les concepts X et Y est grand. Ce modèle suppose également un héritage des propriétés entre les catégories et leurs super-ordonnées. Le sens d'un concept dépend de sa position dans ce réseau sémantique organisé hiérarchiquement. Il est cependant reproché à ce modèle le fait que l'ensemble des membres d'une catégorie y possède la même représentativité. Le *mo-*

dèle en réseau est fondé sur l'idée d'une organisation des concepts en réseau sémantique où les concepts (nœuds) sont reliés entre eux par une ou plusieurs relations (arcs orientés). Il vient en alternative au modèle hiérarchique. Le sens d'un concept est défini par ses relations avec les autres concepts. La distance entre les nœuds du graphe représente la similarité entre les concepts portés par ces nœuds. L'amorçage sémantique correspond alors à l'activation de l'un des nœuds puis à la propagation de l'activation décroissante dans le temps dans le réseau [6]. Il est cependant reproché à ce modèle d'être trop complexe et pas assez contraint. De nombreux *modèles computationnels* ont été élaborés dans le but de reproduire le fonctionnement de la mémoire sémantique, que ce soit pour l'apprentissage, le stockage ou l'utilisation de connaissances. L'intérêt de ces modèles, au-delà de leur aspect descriptif, est bien connu. Ils permettent, une fois implémenté dans des ordinateurs, de faire des simulations en créant des patients artificiels et ainsi vérifier certaines hypothèses. Cree et al [8] en ont établi un inventaire relativement complet. Les auteurs classent les différents modèles en trois groupes : *(i)* les modèles fondés sur un traitement distribué et parallèle, *(ii)* les modèles de mémoire sémantique altérée et *(iii)* les modèles fondés sur la cooccurrence de mots.

3 Les ontologies

Les ontologies puisent leurs racines dans les ontologies philosophiques et les taxinomies d'Aristote. Elles reprennent d'ailleurs le modèle hiérarchique plus contemporain de la mémoire sémantique de Collins & Quillian [7]. Si elles ont pour objectif de modéliser la connaissance d'un groupe d'individus pour un domaine donné, sous la forme d'ensembles structurés de concepts (représentés par leurs propriétés, les termes les dénotant et une liste d'exemplaires) et de relations typées entre ces concepts. Il est possible de considérer ces ontologies soit comme des outils informatiques, situés entre logiciel et base de données, soit comme un possible modèle formel de la mémoire sémantique dans le prolongement des travaux de Collins & Quillian – paradigme que nous adoptons dans nos travaux. La plupart des articles, travaux et autres mémoires évoquant ces ontologies donnent en préambule de leurs propos la définition énoncée par Grüber [11] : « spécification explicite d'une conceptualisation partagée » (ou par Borst [4]} qui parle de formalisation au lieu de spécification explicite). Derrière cette formule, Guarino et al [12]} distinguent trois notions qu'ils tentent d'expliciter. Il s'agit tout d'abord d'une conceptualisation, c'est-à-dire « un résumé, une vision simplifiée du monde que nous souhaitons représenter dans un but donné » (en référence aux travaux de Genesereth et al [10]). Guarino et al [12] définissent ce monde conceptualisé comme « un ensemble totalement ordonné d'états du monde, correspondant à l'évolution du système dans le temps. » Il s'agit ensuite d'une entité formelle, d'une spécification explicite, utilisant un langage – typiquement de la famille des logiques de description » faisant référence aux différents éléments (concepts/classes, instances, termes, relations). Il s'agit enfin d'une conceptualisation partagée, c'est-à-dire l'expression d'un consensus sur la représentation qu'un ou plusieurs individus se font de leur monde, c'est-à-dire « une structure d'un domaine sous réserve que les

parties prenantes comprennent ses termes primitifs dans des points de vue appropriés ». Nonobstant, pour ces auteurs, une conceptualisation formelle peut ne pas être partagée, dans le cas – par exemple – où elle modélise une conceptualisation au niveau de l'individu. Selon Uschold [18], « chaque individu, organisation ou système a une ontologie [souvent tacite] – les choses supposées exister dans le monde et la façon dont ils se comportent... » Ces ontologies personnelles imprègnent et sous-tendent nos discussions, nos décisions et guident nos actions. A l'inverse, une conceptualisation partagée peut ne pas être formelle, si nous la considérons comme un ensemble de connaissances propres à un groupe d'individus, à une organisation, « permettant d'être clair sur ce que les choses signifient et, ce faisant, mettre tout le monde sur la même longueur d'onde » [18]. La catégorisation sous-jacente à ces conceptualisations dépend donc de l'individu et de son contexte. Ce point de vue écologique, défendu entre autre par Rosch [16], a été repris dans les travaux de Aimé [1] par l'introduction de la prototypicalité en Ingénierie des ontologies.

4 Ontologie et mémoire sémantique

Le vieillissement normal et le vieillissement pathologique semblent affecter de manières différentes les systèmes mnésiques. D'après Eustache & Desgranges [9], « selon une conception hiérarchique de la mémoire sémantique, la maladie d'Alzheimer provoque une dégradation partielle et progressive de la mémoire sémantique, affectant d'abord les attributs spécifiques et épargnant les connaissances catégorielles, au moins au début de la maladie. C'est cette dégradation de bas en haut de la mémoire sémantique qui entraine la production de paraphasies sémantiques, des erreurs dans les tâches de définition et une réduction du nombre d'exemplaires cités en fluence verbale. » Les troubles de la mémoire sémantique peuvent être des signes précoces de la maladie d'Alzheimer. Il convient donc de pouvoir les repérer très en amont du développement de la maladie en testant la mémoire sémantique des patients. Dans les tests de la mémoire sémantique de type « appariement », il est demandé au patient de retrouver un ensemble d'informations de nature sémantique relatives à un item cible, et d'associer ces informations à celles de l'un des autres items proposés. Ces informations peuvent être des propriétés communes ou des relations sémantiques. Il existe plusieurs tests d'appariement, parmi lesquels le *Pyramid and Palm Tree Test* (PPTT ; [14]), le *Camel and Cactus Test* (CCT ; [5]), ou encore la *Batterie d'Evaluation des Connaissances Sémantique* (BECS ; [2]). Ces tests ont pour qualité principale d'être assez faciles à administrer, et ne nécessitent que peu d'effort aux patients (appariements image-image, image-terme ou terme-terme). Cependant, la majorité d'entre eux (1) ne demande pas aux patients de préciser la nature de l'association entre les items, et (2) évalue l'association suivant des réponses prototypiques. Or il peut être pertinent de savoir si une mauvaise réponse suivant le protocole est une réponse non conventionnelle mais justifiée, ou une altération de la représentation conceptuelle. La construction par un patient d'une petite ontologie (phase de conceptualisation uniquement) sur un objet donné peut être très riche d'enseignements. Nous pensons qu'un test fondé sur ce processus pourrait apporter plus d'informations sur la qualité de la con-

ceptualisation de l'univers du patient que les tests actuels. Il permettrait au praticien d'évaluer l'étendue – et l'évolution – du réseau sémantique associé à un objet donné pour le patient, et ce tant sur le lexique que sur les propriétés descriptives comme fonctionnelles, les relations avec d'autres concepts, ou encore les instances (qui relèvent de la mémoire épisodique), et ce de manière individuelle. Nos travaux futurs visent à l'élaboration d'un nouveau test de la mémoire sémantique fondé sur cette hypothèse et à son évaluation auprès de patients atteints de la maladie d'Alzheimer.

5 Références

1. Aimé, X.: Gradients de Prototypicalité, mesures de proximité et de similarité sémantique: une contribution à l'Ingénierie des Ontologies. Thèse Université de Nantes (2011), http://tel.archives-ouvertes.fr/tel-00660916
2. Belliard, S., Moreaud, O.: GREMOIRE : tests et échelles de la maladie d'Alzheimer et des syndromes apparentés, chap. GRESEM Batterie d'évaluation des connaissances sémantiques (BECSGRECO), pp. 83–85. Solal, Marseille (2008)
3. Belliard, S., Bon, L., LeMoal, S., Jonin, P.Y., Vercelletto, M., LeBail, B.: La démence sémantique. Psychologie & NeuroPsychiatrie du vieillissement 5(2), 127–138 (2007)
4. Borst, W.: Construction of Engineering Ontology. Ph.D. thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands (1997)
5. Bozeat, S., Lambon Ralph, M.A., Patterson, K., Garrard, P., Hodges, J.R.: Nonverbal semantic impairment in semantic dementia. Neuropsychologia 38(9), 1207–1215 (2000)
6. Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychological review 82(6), 407 (1975)
7. Collins, A.M., Quillian, M.R.: Retrieval time from semantic memory. Journal of verbal learning and verbal behavior 8(2), 240–247 (1969)
8. Cree, G.S., Armstrong, B.C.: Computational models of semantic memory. The Cambridge Handbook of Psycholinguistics p. 259 (2012)
9. Eustache, F., Desgranges, B.: Les systèmes de mémoire dans la maladie d'Alzheimer. Psychologie française 42, 391–402 (1997)
10. Genesereth, M., Nilsson, N.: Logical Foundations of Artificial Intelligence. Morgan Kaufmann, Los Altos (1987)
11. Gruber, T.: Toward principles for the design of ontologies used for knowledge sharing. In: Guarino, N., Poli, R. (eds.) Formal Ontology in Conceptual Analysis and Knowledge Representation. Kluwer Academic Publishers, Deventer, The Netherlands (1993)
12. Guarino, N., Oberle, D., Staab, S.: Handbook on Ontologies, chap. What is an ontology?, pp. 1–17. International Handbooks on Information Systems (2009)
13. Hillis, A.E., Rapp, B., Romani, C., Caramazza, A.: Selective impairment of semantics in lexical processing. Cognitive Neuropsychology 7(3), 191–243 (1990)
14. Howard, D., Patterson, K.: The pyramid and palm trees test: A test of semantic access from words and pictures. Bury St Edmunds: Thames Valley Publishing CO (1992)
15. Laisney, M., Desgranges, B., Eustache, F., Giffard, B.: L'altération du réseau lexico-sémantique dans la maladie d'Alzheimer et la démence sémantique à travers le prisme des effets d'amorçage sémantique. Revue de neuropsychologie 2(1), 46–54 (2010)
16. Rosch, E.: Cognitive reference points. Cognitive Psychology (7), 532–547 (1975)
17. Tulving, E.: Episodic and semantic memory. Organization of memory pp. 381–402 (1972)
18. Uschold, M.: Making the case for ontology. Applied ontology 6(4), 377–385 (2011)