



HAL
open science

Une analyse automatique de cartes conceptuelles pour modéliser l'apprenant

Fabien Delorme, Nicolas Delestre, Jean-Pierre Pécuchet

► **To cite this version:**

Fabien Delorme, Nicolas Delestre, Jean-Pierre Pécuchet. Une analyse automatique de cartes conceptuelles pour modéliser l'apprenant. 2005. hal-00005760

HAL Id: hal-00005760

<https://telearn.hal.science/hal-00005760>

Preprint submitted on 1 Jul 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une analyse automatique de cartes conceptuelles pour modéliser l'apprenant

Fabien Delorme, Nicolas Delestre, Jean-Pierre Pécuchet

Laboratoire PSI

INSA de Rouen

{fabien.delorme, nicolas.delestre, jean-pierre.pecuchet}@insa-rouen.fr

RÉSUMÉ. Les hypermédias adaptatifs dédiés à l'enseignement souffrent généralement d'une modélisation soit trop rigide, les destinant à l'apprentissage d'un champ disciplinaire déterminé a priori, soit trop grossier, ne permettant pas la construction d'un modèle suffisamment riche pour adapter finement le document produit. Notre objectif est d'aller à l'encontre de ces problèmes, en proposant une modélisation de l'apprenant plus riche car non basée sur une modélisation par recouvrement. L'idée que nous avons retenue est d'utiliser des cartes conceptuelles comme outil d'évaluation : l'apprenant réalise une carte, qui est comparée aux cartes de référence réalisées préalablement par l'enseignant. Le résultat de cette comparaison permet alors d'établir un diagnostic quant aux connaissances de l'apprenant. Cet approche présente l'avantage d'être automatisable : nous proposons dans cet article un outil d'analyse automatique de cartes produites par des apprenants à des fins de modélisation.

MOTS-CLÉS : modèle de l'apprenant, hypermédia, système adaptatif, intelligence artificielle.

1. Introduction

Nous souhaitons proposer aux enseignants un outil leur permettant de concevoir des cours en ligne, sans toutefois nous limiter à la méthode d'apprentissage dite transmissive. Cet outil devra permettre non seulement une adaptation du contenu, mais devra en plus être configurable par un enseignant, quelle que soit sa discipline. De tels outils existent [DELESTRE 00]. L'un des points noirs de ces systèmes est néanmoins leurs mécanismes simplistes de modélisation du domaine et de l'apprenant, bridant les capacités du moteur d'adaptation. D'autres systèmes, basés sur une modélisation plus riche, sont malheureusement dédiés à une discipline particulière [BRUSILOVSKY & SU 02]. C'est ce mécanisme d'évaluation que nous nous proposons de dépasser notamment grâce à l'utilisation de cartes conceptuelles. À l'aide de cet outil, il est possible de proposer une évaluation plus profonde des connaissances de l'apprenant, permettant de générer un modèle épistémique riche et donc d'adapter le cours d'une manière plus fine [DELORME & al. 04b].

2. Définir un concept à l'aide d'une carte conceptuelle

Barth propose de s'intéresser à la structure des concepts tels qu'ils peuvent être abordés dans un cadre éducatif [BARTH 87]. Dans un tel contexte, l'enseignant est amené à utiliser un nombre assez conséquent de termes ou d'expressions que l'apprenant ne maîtrise pas forcément. Or, l'apprenant comme l'enseignant ne sont pas toujours conscients du fait qu'ils n'associent pas le même signifié à un signifiant donné. On parle dans ce cas-là de *misconception* : l'apprenant est bel et bien capable de donner une définition cohérente au terme qu'utilise l'enseignant, mais cette définition n'est pas celle que l'enseignant attend. Il est donc nécessaire d'amener l'apprenant comme l'enseignant à formuler explicitement la manière dont ils définissent un concept donné.

La solution proposée par Barth suppose de modéliser les savoirs abordés en suivant une certaine méthodologie. L'auteur précise ainsi qu'un concept est une entité composée de trois éléments interdépendants : l'étiquette du concept, qui est le terme employé par l'enseignant ou l'apprenant ; ses attributs, qui sont un ensemble de conditions nécessaires et suffisantes pour qu'un phénomène donné relève de ce concept ; un ensemble d'exemples concrets, qui sont autant de phénomènes vérifiant tous les attributs associés au concept. Considérons par exemple la notion de tri récursif en algorithmique, telle qu'elle pourrait être enseignée dans le cadre d'un cours d'algorithmique. L'étiquette de cette notion sera *tri récursif*. Ses attributs seront les suivants :

- est un *tri*,
- a pour sous-procédures *diviser en deux parties, trier les deux parties et combiner les deux parties*.

Il faudra alors associer à ces deux éléments un ensemble d'exemples. Pour une notion telle que celle qui nous intéresse, un ensemble de traces d'exécution d'un algorithme récursif pourra constituer l'ensemble d'exemples.

Une fois sa définition déterminée, et une fois qu'une première phase

d'enseignement a eu lieu, l'enseignant peut demander à l'apprenant de définir la notion de tri récursif. En fonction de sa réponse, plusieurs situations peuvent être déterminées :

- l'apprenant a donné la définition qu'attendait l'enseignant,
- l'apprenant a oublié certains attributs attendus par l'enseignant, ou a ajouté certains attributs non essentiels,
- l'apprenant a confondu deux notions, donnant la définition d'une autre notion,
- l'apprenant n'a pas été capable de donner une définition.

La réaction de l'enseignant dépendra alors de la manière dont la réponse de l'apprenant aura été classée. C'est une telle classification qui nous intéresse : c'est elle qui jouera le rôle de modèle épistémique de l'apprenant.

Dans le cas qui nous intéresse, c'est un système automatisé qui devra analyser la définition proposée par l'apprenant, et réagir en conséquence. Nous sommes donc amenés à demander à l'apprenant et à l'enseignant de formuler leur définition de manière à ce qu'elle soit aisément manipulable informatiquement, tout en les contraignant le moins possible.

L'idée que nous avons retenue repose sur l'utilisation de cartes conceptuelles.

2.1 La carte conceptuelle comme formalisme de définition de concept

Dans notre contexte, l'enseignant va d'abord être amené à modéliser les notions qu'il souhaite aborder, à l'aide d'une carte conceptuelle. Cette dernière sera ensuite utilisée comme référence pour analyser celle d'un apprenant. Les différences entre les deux cartes sont alors autant d'éléments permettant de diagnostiquer une incompréhension. La carte conceptuelle va donc être utilisée comme un outil permettant de définir une notion donnée : elle sera centrée autour d'un noeud représentant cette notion à définir. Toutes les relations avec d'autres concepts ne serviront qu'à représenter les différents attributs définissant la notion. Ainsi, nous aurons une carte en étoile, tous les liens partant de la notion à définir et l'associant à divers autres concepts.

Une fois cette carte construite, l'enseignant pourra demander à l'apprenant de construire la sienne. Autoriser l'apprenant à utiliser les termes de son choix serait préférable, mais pose des problèmes lors de l'analyse automatique. C'est pourquoi nous proposons de contraindre l'apprenant à sélectionner les concepts et types de liens qu'il peut utiliser dans sa carte parmi une liste déterminée a priori par l'enseignant. Ainsi, nous utilisons la carte conceptuelle non pas comme un outil d'évaluation ouvert, mais comme un outil fermé.

3. Mise en oeuvre

L'environnement que nous sommes en train de développer comprend un outil de construction de cartes conceptuelles par l'apprenant, nommé DIOGEN, un outil analysant les cartes de l'apprenant pour générer le modèle épistémique, nommé MoNIME, et un outil nommé HypARCcIA adaptant l'hypermédia en fonction du

modèle de l'apprenant. DIOGEN a été présenté en détail dans [DELORME & al. 04a]. Des techniques d'adaptation pouvant être mises en oeuvre dans HypARCCIA sont proposées dans [DELORME & al. 04b].

Une fois les cartes envoyées au système par l'apprenant MoNIME essaie de les classer pour générer un modèle épistémique. Pour chaque carte envoyée, un modèle est généré. Comme nous l'avons vu à la section 2, analyser une carte consiste à chercher la carte de référence la plus ressemblante à celle de l'apprenant :

- si sa carte ressemble à la carte de référence attendue, on n'a plus qu'à chercher les éventuels attributs manquants,
- si elle ressemble à une autre carte de référence, l'apprenant a confondu les deux notions,
- si elle ressemble à une carte entièrement vide, ou ne ressemble à aucune carte connue, l'apprenant n'a pas compris la définition qui lui était proposée.

3.1 Recherche des *k* plus proches voisins

L'algorithme de recherche des *k* plus proches voisins [DUDA & al. 01] est un outil classique de classification. Supposons que l'on connaisse, dans un espace à *n* dimensions, un certain nombre de points étiquetés, c'est-à-dire associés à une classe connue. Soit un point *p* dont on connaît les coordonnées et que l'on souhaite classer. En recherchant parmi les points déjà connus celui qui est le plus proche de *p*, et en observant sa classe, on en déduira que *p* relève de la même classe. Pour déterminer quel est le point le plus proche, un calcul de distance euclidienne est suffisant. Si *p* est trop éloigné de tous les points référencés, on le rejettera en indiquant ainsi que l'on est incapable de le classer.

Afin d'appliquer cet algorithme à notre cas, reste principalement à déterminer une méthode permettant de représenter une carte conceptuelle sous la forme d'un vecteur. Formellement, une carte conceptuelle est un ensemble d'attributs $r(c_1, c_2, \dots, c_n)$ où *n* est l'arité de la relation *r*, et *c_i* un concept participant à la relation. Or, lorsqu'ils construisent leurs cartes à l'aide de DIOGEN, les apprenants doivent respecter quelques contraintes : ils doivent se contenter des types de liens (binaires ou *n*-aires) et des concepts déterminés a priori par l'enseignant, et le noeud source d'une relation est forcément le noeud représentant le concept à définir. Le nombre d'attributs différents que les apprenants pourront produire est donc non seulement fini, mais également peu élevé du fait que le noeud source est fixé pour tous les attributs. Ainsi, nous aurons $n = \text{Somme}(r_i \cdot c_i - 1)$ attributs possibles, où *c* est le nombre de concepts et *r_i* le nombre de relations d'arité *i*. Il est donc possible de représenter une carte basée sur cet ensemble de relations et de concepts comme un vecteur *v* avec $v_i = 1$ si l'attribut *i* apparaît sur la carte, et $v_i = 0$ sinon. Une fois la carte de l'apprenant transformée, il suffit de calculer la distance qui la sépare de chacune des cartes de référence, et de déterminer laquelle est la plus proche.

3.2 Prise en compte de la similarité entre attributs

Cette approche présente toutefois un défaut majeur : les attributs sont considérés comme étant sémantiquement indépendants les uns des autres. Or, la plupart des

attributs présentent en réalité des relations de similarité entre eux. Ainsi, l'attribut « est un tri » est relativement similaire à « appelle un tri », ou à « est une procédure », bien que ces attributs ne soient pas équivalents. Dès lors, quand l'attribut « est un tri » est présent, il faudrait faire en sorte que tous les autres attributs présentant une relation de similarité plus ou moins importante avec lui soient considérés comme étant partiellement présents. Le vecteur représentant la carte ne contiendra alors plus des valeurs booléennes, mais des valeurs réelles comprises entre 0 et 1 indiquant que l'attribut est plus ou moins présent sur la carte. La proximité sémantique entre deux attributs dépend de deux aspects : la relation et le ou les concepts reliés. Nous allons donc d'abord nous intéresser aux similarités sémantiques entre les relations, le même raisonnement étant ensuite appliqué aux concepts.

Nous cherchons donc à définir une fonction sim_r telle que $0 \leq sim_r(r1, r2) \leq 1$: si $sim_r(r1, r2) = 0$, alors la présence de $r1$ n'influence pas la présence de $r2$. Si $sim_r(r1, r2) = 1$, cela signifie que $r1 \Rightarrow r2$. Enfin, si $0 < sim_r(r1, r2) < 1$, cela signifie que $r1$ sous-entend partiellement $r2$. Il reste à déterminer comment calculer la valeur de sim_r pour $r1$ et $r2$ donnés. Les relations proposées par l'enseignant sont organisées selon une taxonomie. Il est dès lors possible de dresser un critère de similarité entre les relations : plus le nombre d'arcs à traverser dans la hiérarchie pour passer d'une relation à l'autre est élevé, plus elles sont dissemblables ; en outre, si une relation est présente, alors ses relations parentes sont forcément présentes. Par exemple, si « est un x » est présent, l'attribut « est lié à x » est également présent, car la relation « est lié à » est plus générale que « est un ». À l'inverse, si « est lié à y » est présent, « est un y » sera partiellement présent.

3.3 Résultats expérimentaux

Une expérimentation a été menée dans le cadre d'un cours d'algorithmique au sein de l'INSA de Rouen. Après avoir suivi un cours sur les algorithmes de tris, les étudiants ont été amenés à définir 6 types de tris différents à l'aide du logiciel DIOGEN. Les cartes de 42 étudiants ont alors été analysées et classées par deux experts humains, ainsi que par MoNIME. Les résultats montrent que les deux experts ont évalué les cartes de la même manière dans 84,13 % des cas. La proximité de ce taux avec celui qui est obtenu en comparant une analyse humaine et l'analyse produite par MoNIME (76,98 % et 82,14 % selon l'évaluateur) permet de conclure à l'efficacité de l'algorithme proposé.

4. Conclusion et travaux à venir

L'utilisation de DIOGEN dans un contexte pédagogique réel nous avait permis de confirmer l'idée que les cartes conceptuelles peuvent être utilisées pour évaluer un apprenant. Le développement du logiciel MoNIME et la comparaison des résultats qui en sont issus avec ce que donne une analyse manuelle nous permet maintenant d'affirmer que cette utilisation des cartes conceptuelles est automatisable. Un retour sur les erreurs et omissions des apprenants à destination de l'enseignant serait désormais intéressant. Il est en effet possible que certaines erreurs récurrentes soient le signe d'une mauvaise conception du support proposé par l'enseignant.

5. Bibliographie

- [BARTH 87] Barth B.-M. (1987), *L'apprentissage de l'abstraction*, Retz, Paris.
- [BRUSILOVSKY & SU 02] Brusilovsky P. & Su H.D. (2002), "Adaptive visualization component of a distributed web-based adaptive educational system", in *Proceedings of 6th International Conference on Intelligent Tutoring Systems, ITS'2002*, p. 229-238.
- [DELESTRE 00] Delestre N. (2000), METADYNE : un Hypermédia Adaptatif et Dynamique pour l'Enseignement, Thèse de doctorat, Université de Rouen.
- [DELORME & al. 04a] Delorme F., Delestre N. & Pécuchet J.-P. (2004a), "Évaluer l'apprenant à l'aide de cartes conceptuelles", in *TICE'2004*, p. 25-31.
- [DELORME & al. 04b] Delorme F., Delestre N. & Pécuchet J.-P. (2004b), "Using concept maps for enhancing adaptation processes in declarative knowledge learning", in *AH 2004*, p. 397-400, Springer-Verlag.
- [DUDA & al. 01] Duda R., Hart P. & Stork D. (2001), *Pattern Classification*. Wiley Interscience - 2e édition.