



HAL
open science

Suivi et analyse d'un processus collaboratif de modélisation de connaissances spatiales

Florence Le Ber, Sylvie Lardon, Christian Brassac

► To cite this version:

Florence Le Ber, Sylvie Lardon, Christian Brassac. Suivi et analyse d'un processus collaboratif de modélisation de connaissances spatiales. Conférence Québéco-Française de Développement de la Géomatique (CQFD-Géo 2007), 2007, Clermont-Ferrand, France. pp.1-8. hal-00156976

HAL Id: hal-00156976

<https://hal.science/hal-00156976v1>

Submitted on 25 Jun 2007

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.

Suivi et analyse d'un processus collaboratif de modélisation de connaissances spatiales

Florence Le Ber^{1,2}, Sylvie Lardon³, Christian Brassac⁴

¹ CEVH (ENGEES – ULP), 1 quai Koch, B.P. 61039, F-67070 Strasbourg cedex

² LORIA, Campus universitaire, BP 239, F-54506 Vandœuvre-lès-Nancy cedex
florence.leber@engees.u-strasbg.fr

³ UMR Métafort, INRA & ENGREF POP'TER, Domaine universitaire des Cézèaux,
24 avenue des Landais, BP 90054, F-63171 Aubière cedex 9
lardon@engref.fr

⁴ Équipe Codisant – LabPsyLor, U. Nancy 2, BP 33-97, F-54015 Nancy Cedex
Christian.Brassac@univ-nancy2.fr

RÉSUMÉ. Nous rendons compte ici du développement d'un système à bases de connaissances, le système ROSA, qui a donné lieu à un processus conjoint de modélisation de connaissances spatiales. Ce processus a impliqué des agronomes, des informaticiens, des psychologues et des linguistes. Une partie des séances de travail entre agronomes et informaticiens a été filmée, ce qui a permis une analyse réflexive du processus de modélisation.

ABSTRACT. This paper focuses on the elaboration of a knowledge based system called ROSA, which has implied a conjoint modelling process of spatial knowledge. This process has included agronomists, computer scientists, psychologists and linguists. Three work sessions between agronomists and computer scientist were filmed, allowing a reflexive analysis of the modelling process.

MOTS-CLÉS : agronomie, territoire, système à bases de connaissances, raisonnement à partir de cas, modélisation de connaissances, interaction, travail collaboratif

KEYWORDS: agronomy, territory, knowledge base system, case based reasoning, knowledge modelling, interaction, collaborative work

1. Introduction

Pour une meilleure gestion de l'environnement et des territoires agricoles, des agronomes s'attachent à collecter l'expression des connaissances et des pratiques des agriculteurs, en s'appuyant sur des enquêtes, des documents, des témoignages et sur leur propre connaissance des territoires enquêtés. La variété et la complexité des éléments recueillis nous ont conduits à proposer de construire un système à bases de connaissances, apte à les mémoriser et les exploiter. Le développement du système, dénommé ROSA¹, s'est appuyé sur un processus de modélisation de connaissances, qui a été conduit par un ensemble conséquent de chercheur(e)s (en agronomie, en informatique, en linguistique, en psychologie)², dans un mouvement en trois étapes s'étalant sur une période de quatre à cinq années. La première étape a consisté à identifier les concepts que les agronomes utilisent pour décrire la structure spatiale et fonctionnelle d'une exploitation et, plus loin, à dresser une 'base de cas' sur laquelle repose le système informatique. La seconde étape a donné lieu à l'analyse et à la comparaison de 'cas' et a conduit à un prototype numérique. La troisième étape est une étape d'évaluation qui a fait évoluer le prototype vers une meilleure adéquation aux problèmes rencontrés.

L'élaboration du système à bases de connaissances est scandé par ces trois étapes, marquées par des séances de travail réunissant les agronomes et les informaticiens. Certaines séances ont été filmées, permettant la constitution de corpus qui rendent possible l'analyse de ces différents moments. Ce papier a pour objet d'en rendre compte succinctement. Nous commencerons par présenter la problématique agronomique et le modèle informatique utilisé. Nous poursuivrons en exposant successivement les contenus des trois étapes puis nous concluons.

2. Cadre

2.1. *Gestion de territoires agricoles*

Pour comprendre la contribution des exploitations à la transformation des paysages, les agronomes analysent les relations entre le fonctionnement des exploitations agricoles et leur organisation spatiale. Ils étudient en particulier les exploitations d'élevage ovin de la région des Grands Causses, au Sud de la France, soumises à des dynamiques d'embroussaillage (Cohen, 2003). Les enquêtes exhaustives des exploitations du Causse Méjan réalisées à plusieurs dates depuis 1974 permettent de suivre l'évolution des activités d'élevage ovin et du milieu caussenard (Lardon *et al.*, 1995). Pour chaque exploitation, l'enquête réalisée en trois passages porte sur les pratiques d'utilisation et de configuration du territoire de l'exploitation. Les pratiques d'utilisation de l'espace sont les façons dont les éleveurs, au cours d'une campagne,

1. Pour *Raisonnement sur des Organisations Spatiales Agricoles*.

2. Ce travail collectif a bénéficié du soutien du programme *GETM - Société de l'information* du CNRS - IGN - Cemagref (2002-2005).

mobilisent les ressources pour répondre à des objectifs de production ; ce sont les pratiques de culture, de pâturage et d'entretien du milieu. Les pratiques de configuration de l'espace sont les façons dont les éleveurs, sur plusieurs années, transforment l'espace pour une meilleure utilisation ; elles correspondent aux pratiques de constitution du territoire de l'exploitation, d'aménagement et de mise en valeur des surfaces.

L'analyse de ces exploitations a été faite en référence à la typologie des modes de conduite des exploitations et à l'évolution des systèmes techniques de production. La typologie des modes d'organisation spatiale est basée sur la méthode des chorèmes proposée par les géographes (Brunet, 1986). La construction progressive du schéma d'organisation spatiale de l'exploitation est faite à partir de la combinaison de chorèmes élémentaires auxquels sont associés des interprétations fonctionnelles agricoles (Lardon *et al.*, 2000).

2.2. *Raisonnement à partir de cas*

Le système ROSA est fondé sur le modèle du raisonnement à partir de cas (Kolodner, 1993) : il est composé d'une base de cas constituée d'informations sur les exploitations enquêtées, d'une base de connaissances sur le domaine et d'un module de raisonnement. Les objectifs de ce module sont de comparer et d'adapter les connaissances liées à un cas (par exemple une exploitation source, dont on connaît à la fois la structure spatiale et le fonctionnement) à un autre cas (une exploitation cible dont on ne connaît que la structure). La comparaison est établie sur des mesures de similarités entre les structures spatiales des deux cas et l'adaptation permet de proposer une logique de fonctionnement à l'exploitation cible à partir de la logique de fonctionnement de l'exploitation source en s'appuyant sur la similarité entre les structures spatiales. L'hypothèse sous-jacente peut s'exprimer ainsi : *le fonctionnement de source est au fonctionnement de cible ce que la structure de source est à la structure de cible* (Lieber, 1997).

Pour représenter les structures spatiales agricoles, les informaticiens ont fait le choix d'utiliser des graphes conceptuels (Sowa, 1984), qui sont un outil de modélisation facilement lisible et donc appropriable par les agronomes. Un graphe est composé de deux ensembles de sommets, l'un représente des entités, l'autre des relations. Les arêtes du graphe sont étiquetées par le rôle³ des entités dans la relation (voir figure 1). De plus, nous nous plaçons dans le cadre d'une représentation hiérarchique des connaissances (Ducournau *et al.*, 1998) : les concepts sont organisés selon une relation d'ordre, fondée sur une classification naturelle du domaine (par exemple, « *les parcs sont des surfaces en herbe* », « *la bergerie est un bâtiment d'exploitation* »). Cette relation d'ordre est utilisée par établir une mesure de ressemblance entre les concepts et entre les graphes, et c'est donc sur elle que reposent les opérations de comparaison et d'adaptation du système ROSA (Le Ber *et al.*, 2003).

3. Au sens des logiques de descriptions.

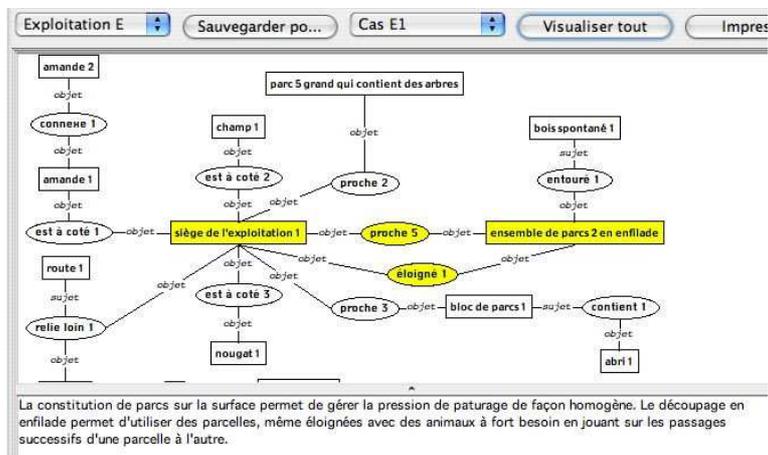


Figure 1. Interface du logiciel ROSA : vue d'un graphe décrivant la structure spatiale d'une exploitation ; le sous-graphe coloré en jaune correspond à un cas, l'explication associée est inscrite en bas de l'écran

3. Les étapes du processus de modélisation

3.1. Modélisation du domaine et définition des cas

Dans cette première étape, il s'agit de cerner les concepts que les agronomes utilisent pour décrire la structure spatiale et fonctionnelle d'une exploitation agricole, puis de formaliser les cas. La procédure de modélisation des connaissances s'appuie d'une part sur les chorèmes d'exploitations, utilisés par les agronomes, et d'autre part sur les graphes, utilisés par les informaticiens. Une interrogation systématique du sens porté par les schémas chorématiques conduit le groupe à deux choses :

- l'explicitation des concepts du domaine, qui peut, apparaître comme une découverte de concepts utilisés mais non nommés : *"alors là je ne sais pas comment tu vas appeler ça" ... "c'est un ensemble de parcs contenant champs"* (2001(7), 39'08) ;
- la définition de cas, qui sont des fragments d'espace dotés d'une organisation et d'une signification particulières : *"il y a une contrainte quand même, je présume que la bordure du ruisseau c'est plus humide" ... "c'est au travers d'une pratique de mise en herbe qu'il résout son problème"* (2001(1), 43'17).

Cette activité d'acquisition s'est déroulée au cours de plusieurs séances réunissant agronomes et informaticiens. L'une d'entre elles (en janvier 2001) a fait l'objet d'un filmage. Le filmage permet d'avoir accès à l'événement interactionnel qui s'est déployé dans un espace-temps donné. La documentation pré-existante au travail en commun et les formes graphiques et textuelles produites dans cet espace-temps sont ainsi disponibles pour analyse. Plus, du point de vue des psychologues et des lin-

guistes, on dispose de l'histoire de la production des multiples traces laissées par les acteurs. On peut par exemple constater les modifications concrètes (sur-traçages ou biffages) apportées à un chorème, ou suivre de façon très précise le processus d'élaboration d'un graphe, ce suivi mettant à jour hésitations, gommages, raturages ou autres altérations de tel ou tel tracé (pour une présentation et une analyse de ce corpus voir par exemple (Brassac *et al.*, 2005)).

Le travail a porté de façon exhaustive sur sept exploitations, donnant lieu à environ soixante-dix cas et à la construction d'une hiérarchie de concepts. À l'issue de cette première étape, le système ROSA se trouve donc doté d'une base de cas et d'une base de connaissances. La base de cas contient des graphes représentant des chorèmes d'exploitations. Chaque graphe peut être découpé en sous-graphes et donner ainsi lieu à plusieurs cas – c'est-à-dire des paires (sous-graphe, explication) – tels que celui présenté dans la figure 1. La base de connaissances contient les concepts (entités spatiales et relations) du domaine.

3.2. *Explicitation des connaissances d'adaptation*

La base de cas étant constituée à partir d'un ensemble d'exploitations connues, l'idée est de pouvoir utiliser (*adapter*) ces cas pour analyser d'autres exploitations en termes de rapport structure spatiale - modalité de fonctionnement. La deuxième étape consiste alors à examiner de nouvelles exploitations (notées EC pour *exploitations cibles*), et à les comparer aux fragments d'exploitations (notées ES, pour *exploitations sources*) contenus dans la base de cas. La procédure implique agronomes et informaticiens et consiste à mettre en regard des chorèmes et/ou des graphes de différentes exploitations, en évaluant les ressemblances et dissemblances entre leurs structures spatiales et fonctionnelles. Elle permet éventuellement de mettre à jour des connaissances ou règles permettant d'*adapter* le savoir issu d'une 'EC' à une 'ES'.

Là encore, ce travail s'est déroulé en plusieurs séances réunissant ces deux groupes de chercheurs. L'une d'elles, mise en place en mars 2003, a donné lieu à un filmage. La séance s'est déroulée en deux temps. Le premier temps fut consacré à la comparaison de chorèmes d'exploitations sur deux zones différentes (Causse Méjan, Causse Sauverre), le deuxième temps à l'appariement (manuel) de graphes : l'agronome produit un graphe d'exploitation à partir d'un chorème, puis l'informaticien(ne) cherche à appairer des parties du graphe créé avec les sous-graphes des cas (sources) déjà constitués. L'agronome valide ou non l'appariement, puis évalue et adapte l'explication associée au cas source considéré. Ce travail résulte en différents points concernant ROSA :

- une validation des cas décrits dans la base de cas : ils sont vérifiés et utilisés ;
- un affinage du modèle du domaine avec re-qualification des concepts et remise en cause partielle de la structure hiérarchique : “*pacage c'est le terme qu'on a utilisé pour parler de ces surfaces. En herbe. De type parcours. Ou parc*” ... “*non pour moi je pense qu'ils sont tous parcs oui*” (2003 (11), 41'13) ;

– la mise en évidence de structures spatiales génériques, qui peuvent servir à organiser les cas et mener à des « structures-concepts » : *“l'idée c'est qu'une surface donne accès à une autre surface commodément à partir d'une bergerie” ... “alors que jusqu'à présent on était sur des chemins ou sur des proximités”* (2003(22), 14'38).

À l'issue de cette seconde étape, le système ROSA a pu être doté d'une mesure de distance et d'une procédure d'appariement entre graphes, noyau du module de raisonnement à partir de cas. Le prototype ainsi constitué est validé au cours de la troisième étape du processus.

3.3. Évaluation du prototype

L'évaluation du système ROSA s'est réalisée en deux séances, la seconde (février 2005) seulement étant filmée. Celle-ci est lourdement instrumentée. Un des informaticiens dispose d'un ordinateur, dont l'écran est projeté en temps réel sur un tableau blanc (effaçable) *via* un vidéo-projecteur ; les acteurs de la réunion peuvent ainsi réagir directement, par écrit, sur le produit du logiciel (figure 2).

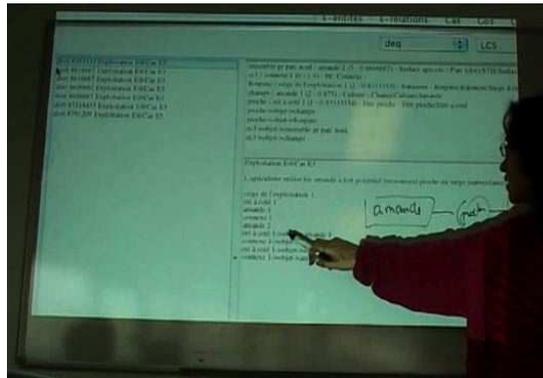


Figure 2. Séance 2005 - visualisation sur écran blanc des cas cible et source appariés par le logiciel ROSA

La séance se déroule de la manière suivante. Le groupe choisit une exploitation 'nouvelle', en examine le chorème, et élabore un graphe qui est soumis au logiciel. Le module de raisonnement fournit en réponse des cas (graphe plus explication), issus de la base de cas, qui présentent des similitudes hiérarchisées avec le graphe cible. Le groupe a alors pour tâche d'évaluer ces similitudes en référence à la hiérarchie du domaine. Ainsi, cette étape d'évaluation du prototype donne lieu à une ré-interrogation des connaissances construites lors des deux premières phases, à savoir :

– les graphes et les explications contenus dans la base de cas : certains graphes sont modifiés, certaines explications réécrites.

– le contenu et la structure de la base de connaissances : en particulier, au-delà de la simple représentation hiérarchique des concepts, on voit qu’il est nécessaire de représenter des continuums : “l’idée c’est que à la fois bloc nougat fertilisé petit proche fait que la surface en herbe bascule plutôt du côté des bonnes surfaces de type champ” (2005, V2, 20’50);

– le mode d’appariement entre graphes, les appariements autorisés entre concepts : “Les amandes on peut les associer aux champs. Et pas aux parcs. Même à des petits parcs” ... “les bons petits parcs soignés si” (2005, V1, 1h22).

Ces dernières observations n’ont pas donné lieu à une évolution concrète du prototype, qui est resté en l’état à l’issue du projet (Metzger, 2005).

4. Conclusion

Le projet s’est construit autour de trois questions-propositions émanant des agronomes, des informaticiens et des psychologues et linguistes. Les premiers s’interrogeaient sur les outils de représentation de l’espace agricole et sur leurs capacités d’explicitation et de généralisation des modes de gestion des agriculteurs. Les seconds s’intéressaient à la modélisation de raisonnements sur l’espace, proposant de développer des approches qualitatives aptes à rendre compte de la singularité et de la complexité de ces modes de gestions. Les troisièmes, quant à eux, étudiaient la mobilisation d’objets dans la construction et le développement d’une interaction sociale finalisée et se sont penchés - dans une vision réflexive - sur la confrontation des modèles agronomiques et informatiques.

Le processus de modélisation conjointe – entre agronomes et informaticiens – des connaissances mises en jeu dans le système ROSA a été suivi et analysé grâce à la mise en place de trois expérimentations, correspondant à trois phases-clé de ce processus : la première phase concernait la modélisation des cas et des connaissances du domaine, la deuxième la modélisation du raisonnement, et la troisième le test du logiciel. Les trois expérimentations ont donné lieu à trois corpus vidéos, supports de confrontations et d’apports concrets entre les disciplines concernées (voir (Le Ber *et al.*, 2005) pour d’autres textes sur ce projet).

5. Bibliographie

- Brassac C., Le Ber F., « Inscription spatiale d’une activité cognitive collective de représentation de l’espace », *Intellectica*, vol. 2-3, n° 41–42, p. 181-200, 2005.
- Brunet R., « La carte-modèle et les chorèmes », *Mappemonde*, vol. 86, n° 4, p. 3-6, 1986.
- Cohen M. (ed.), *La brousse et le berger. Une approche interdisciplinaire de l’embroussaillage des parcours*, Espaces et Milieux, CNRS Éditions, Paris, 2003. 354p.
- Ducournau R., Euzenat J., Masini G., Napoli A. (eds), *Langages et modèles à objets — État des recherches et perspectives*, Collection Didactique D-019, INRIA, Le Chesnay, 1998.
- Kolodner J., *Case Based Reasoning*, Morgan Kaufman Publishers, 1993.

- Lardon S., Capitaine M., Benoît M., « Les modèles graphiques pour représenter l'organisation spatiale des activités agricoles. », *Représentations graphiques dans les systèmes complexes naturels et artificiels. Journées de Rochebrune*, ENST, p. 127-150, 2000.
- Lardon S., Osty P. L., Triboulet P., « Élevage et éleveurs du Causse-Méjan (Lozère). Dynamique de mise en valeur et contrôle de l'espace », in J.-L. Bonniol, A. Saussol (eds), *Grands Causses. Nouveaux enjeux, nouveaux regards*, Fédération pour la Vie et la Sauvegarde du Pays des Grands Causses, 1995.
- Le Ber F., Benoît M., Capitaine M., Lardon S., Osty P.-L., Brassac C., Préau J.-M., Mondada L., Metzger J.-L., Napoli A., « Modélisation, comparaison et interprétation d'organisations spatiales agricoles », *Actes du colloque "Société de l'information"*, Lyon, p. 194-198, 2005.
- Le Ber F., Napoli A., Metzger J.-L., Lardon S., « Modeling and comparing farm maps using graphs and case-based reasoning », *Journal of Universal Computer Science*, vol. 9, n° 9, p. 1073-1095, septembre, 2003.
- Lieber J., « Raisonnement à partir de cas et classification hiérarchique. Application à la planification de synthèse en chimie organique », Thèse de l'Université Henri Poincaré – Nancy I, 1997.
- Metzger J.-L., « Contribution à l'élaboration d'un modèle de raisonnement à partir de cas pour l'aide à l'interprétation d'organisations spatiales agricoles », Thèse de l'Université Henri Poincaré – Nancy I, 2005.
- Sowa J., *Conceptual Structures : Information Processing in Mind and Machine*, Addison Wesley, Reading, Massachusetts, 1984.