



HAL
open science

SIXIÈME ATELIER : Représentation et raisonnement sur le temps et l'espace (RTE 2011)

Sébastien Laborie, Florence Le Ber

► **To cite this version:**

Sébastien Laborie, Florence Le Ber. SIXIÈME ATELIER : Représentation et raisonnement sur le temps et l'espace (RTE 2011). Atelier Représentation et raisonnement sur le temps et l'espace (RTE 2011), 2011, Chambéry, France. 2011. hal-00640820

HAL Id: hal-00640820

<https://inria.hal.science/hal-00640820v1>

Submitted on 9 Feb 2012

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.

Plate-forme AFIA / Chambéry, du 16 au 20 mai 2011

SIXIÈME ATELIER :

Représentation et raisonnement sur le temps et l'espace (RTE 2011)



Sébastien Laborie
Florence Le Ber

Sixième Atelier "Représentation et Raisonement sur le temps et l'espace"

RTE 2011

Chambéry, 16 mai 2011

Avant propos

La représentation du temps et de l'espace ainsi que les modèles de raisonnements associés sont des thèmes largement étudiés en informatique, d'une manière générale, et en intelligence artificielle, en particulier. Ces thèmes sont de plus en plus importants dans de nombreux domaines de notre société, en particulier là où est disponible une très grande quantité d'informations et de services évoluant au cours du temps ou dans l'espace. Les techniques temporelles et/ou spatiales sont, par exemple, importantes dans : la gestion des grandes quantités de données, l'analyse et la fouille de ces données, la simulation et l'analyse de l'évolution temporelle de processus, l'évaluation de la sécurité et la sûreté, la gestion dynamique des connaissances, la gestion de l'espace, la prévention des risques naturels, la modélisation des systèmes dynamiques et complexes, etc. Elles offrent une alternative ou un complément aux méthodes statistiques et mathématiques de modélisation de l'espace et du temps.

En complément des contributions dans les conférences généralistes dans les domaines de l'intelligence artificielle, des systèmes d'informations, des bases de données, du traitement automatique de la langue naturelle, des systèmes temps réel et de l'informatique théorique, l'atelier RTE a émergé comme l'atelier national qui réunit les chercheurs français s'intéressant à la dimension temporelle et/ou spatiale pour discuter des travaux en cours, projets et idées. L'atelier RTE s'est tenu pour la première fois en 2005 à Nice dans le cadre de la plateforme AFIA, a eu lieu ensuite à Nantes (semaine de la connaissance, 2006), à Grenoble (plateforme AFIA, 2007), à Montpellier (conférence SAGEO, 2008), à Caen (conférence RFIA, 2010) et est cette année un atelier de nouveau associé à la plateforme de l'AFIA 2011 qui est accueillie à Chambéry. Il est ouvert à la présentation de travaux de chercheurs et doctorants portant sur l'un ou plusieurs des thèmes suivants : modélisation du temps, modélisation de l'espace, raisonnement spatial quantitatif et/ou qualitatif, raisonnement temporel quantitatif et/ou qualitatif, traitement du langage naturel, représentation des connaissances, fouille de données spatiales et temporelles, applications, ontologies, etc.

Pour cette sixième manifestation, nous avons retenu six soumissions qui présentent des travaux dans les différents champs cités. Ces présentations sont accompagnées d'un exposé de Christian Sallaberry sur l'indexation et la recherche d'information géographique dans des contenus textuels à connotation « territoriale ».

Nous souhaitons que les participants à cet atelier trouvent matière à idées et discussions dans l'ensemble des présentations et que ces discussions puissent se prolonger lors de futures manifestations. Nous remercions les organisateurs des précédents ateliers pour leur aide ainsi que leurs conseils, et en particulier Maroua Bouzid pour la mise en place du nouveau site Internet du groupe RTE à l'adresse <http://rte.greyc.fr>. Pour terminer, nous remercions chaleureusement les membres du comité de programme de l'atelier ainsi que les membres du comité d'organisation de la plateforme AFIA 2011, sans lesquels cet évènement n'aurait pu se tenir.

Sébastien Laborie

Florence Le Ber

Comité de programme

Isabelle Bloch (ENST, Paris)
Maroua Bouzid (GREYC, Caen)
Jean-François Condotta (CRIL, Lens)
Mauro Gaio (LIUPPA, Pau)
Jérôme Gensel (LIG, Grenoble)
Sébastien Laborie (LIUPPA, Bayonne)
Christine Largouët (IRISA, Rennes)
Florence Le Ber (LHYGES, Strasbourg / LORIA)
Gérard Ligozat (LIMSI, Orsay)
Philippe Muller (IRIT, Toulouse)
Odile Papini (LSIS, Marseille)
Pascal Poncelet (LIRMM, Montpellier)

Programme de l'atelier RTE 2011

- ◆ **9h00 – 10h15 : Session 1**
 - Abduction dans les logiques de description : apport de l'analyse formelle de concepts et de la morphologie mathématique.
(Jamal Atif, Céline Hudelot et Isabelle Bloch)
 - *Quels formalismes temporels pour représenter des connaissances extraites de textes de recettes de cuisine ?*
(Valmi Dufour-Lussier, Florence Le Ber et Jean Lieber)

- ◆ **10h15 – 10h45 : Pause-café**

- ◆ **10h45 – 12h : Session 2**
 - Exploration de données temporelles floues : application en archéologie.
(Cyril de Runz et Eric Desjardin)
 - *Localisation d'habitant dans un espace perceptif par réseau dynamique.*
(Pedro Chahuara, Michel Vacher et François Portet)

- ◆ **12h15 – 14h : Déjeuner**

- ◆ **14h – 15h30 : Conférence invitée**
 - ◆ Indexation et recherche d'information géographique dans des contenus textuels à connotation « territoriale ».
(Christian Sallaberry)

- ◆ **15h30 – 16h : Pause Café**

- ◆ **16h – 17h15 : Session 3**
 - Idée préliminaire pour un système de médiation fondé sur une ontologie spatiale.
(Lamia Belouaer)
 - *Méthode de modélisation et d'extraction des informations géographiques contenues dans des textes*
(Van Tien Nguyen et Mauro Gaio)

- ◆ **17h15 – 18h : Discussion**

Indexation et recherche d'information géographique dans des contenus textuels à connotation « territoriale »

C. Sallaberry (christian.sallaberry@univ-pau.fr), pour RTE'2011

Enseignant-chercheur, Laboratoire LIUPPA, Equipe T2i (<http://liuppa.univ-pau.fr/>)

Nos travaux de recherche s'inscrivent dans le domaine de la recherche d'information géographique (RIG). La RIG, nommée et définie pour la première fois par Ray Larson, est la tâche de recherche, au sein d'une collection documentaire, de documents satisfaisant un besoin d'information aux caractéristiques géographiques. La série de conférences GIR¹ débutée en 2004, a fortement contribué au développement de la RIG ; focalisée sur la dimension spatiale dans un premier temps, puis, pour les documents textuels, spatiale étendue par la dimension thématique véhiculée par les termes porteurs de sens autre que spatial. Nous retrouvons désormais la dimension temporelle associée à la dimension spatiale et/ou thématique dans de nombreux prototypes de RIG.

Nous avons pour partenaire la MIDR², qui nous apporte un domaine d'application composé de corpus textuels à connotation « territoriale ». Comme beaucoup de bibliothèques et de médiathèques, la MIDR numérise des documents aux formats divers (œuvres littéraires, récits de voyage, journaux, cartes géographiques anciennes, lithographies, cartes postales, etc.) qui ont pour dénominateur commun de traiter d'un territoire restreint (les Pyrénées), dans une période de l'Histoire déterminée (principalement du XVIII^e et du XIX^e siècle). Nous travaillons notamment sur les récits de voyage qui contiennent d'abondantes références à l'histoire, à la géographie, au patrimoine, en somme au territoire. L'objectif visé est d'offrir, au plus grand nombre, de nouveaux modes d'accès à des fonds documentaires désormais disponibles dans des formats numériques.

En comparaison avec les contenus du Web, nous considérons que de tels fonds documentaires sont stables (le contenu d'un livre ne change pas au cours du temps) et de taille raisonnable pour autoriser des travaux d'indexation approfondis relatifs à chacune des dimensions spatiale, temporelle et thématique.

Cette présentation s'inscrit dans les thématiques de l'atelier RTE. Elle se décompose en deux parties.

Dans la première partie nous décrivons les processus d'indexation et de recherche d'information spatiale et temporelle dans des corpus de documents textuels. Nous nous intéressons tout d'abord à la modélisation de l'information spatiale et temporelle dans un tel contexte de recherche d'information. Nous proposons un modèle spatial et un modèle temporel pivots intégrant les problématiques d'interprétation et de représentation des informations dans les index en vue de la mise en œuvre de calculs d'appariement en phase de recherche. Ensuite, nous proposons une première méthode d'extraction et d'indexation d'information spatiale, basée sur notre modèle pivot et sur un traitement sémantique spécifique. Nous adoptons une démarche similaire, pour proposer une méthode d'extraction et d'indexation d'information temporelle. Ensuite, nous nous intéressons aux approches de recherche d'information mises en œuvre dans les systèmes dédiés à l'information spatiale et temporelle. Nous proposons une méthode de recherche d'information spatiale utilisant les fonctions SIG pour calculer des représentations géo-référencées et pour implémenter un calcul de pertinence spatiale. Selon une démarche similaire, nous proposons une méthode de recherche d'information temporelle dédiée.

Dans la seconde partie, nous présentons le processus d'exploitation des index spatiaux et temporels ainsi produits dans un cadre de recherche d'information multicritère.

Nous avons fait le choix de traiter de manière spécifique chaque dimension de l'information géographique puis de les combiner à des fins de recherche d'information. Avant toute combinaison, nous uniformisons les représentations des données ainsi que les démarches de traitement des données relatives aux différentes dimensions. Nous proposons de soumettre chaque critère d'une requête au système de recherche d'information approprié (le thématique sera limité aux termes), puis, la combinaison des résultats. Nous proposons également de nouveaux opérateurs dont le but est d'associer une plus grande expressivité à chaque critère de la requête et, par conséquent, de retourner des résultats plus pertinents.

¹ <http://www.informatik.uni-trier.de/~ley/db/conf/gir/>

² Médiathèque Intercommunale à Dimension Régionale Pau Pyrénées

Abduction dans les logiques de description : apport de l'analyse formelle de concepts et de la morphologie mathématique

Jamal Atif¹, Céline Hudelot², Isabelle Bloch³

¹ Equipe TAO, Laboratoire de Recherche en Informatique,
Batiment 490, Université Paris-Sud, 91405 - Orsay Cedex jamal.atif@lri.fr

² Ecole Centrale de Paris - celine.hudelot@ecp.fr

³ Telecom ParisTech - CNRS LTCI, Paris -
isabelle.bloch@telecom-paristech.fr

Résumé : L'objet de cet article est d'enrichir les logiques de description par des services de raisonnement abductif à des fins d'interprétation d'images. Dans un cadre algébrique, nous mettons en synergie des ingrédients provenant de la morphologie mathématique, des logiques de description et de l'analyse formelle de concepts. Plus précisément, nous proposons de calculer la « meilleure » explication d'une observation donnée à partir d'une succession d'érosions algébriques sur le treillis de concepts associé à la théorie du domaine. Nous montrons que les opérateurs ainsi définis satisfont les postulats de rationalité du raisonnement abductif. Comme illustration, nous considérons le domaine de l'interprétation d'images à partir de modèles, où des connaissances *a priori* structurées bénéficient de représentations ontologiques et des logiques de description. Nous formulons donc le problème d'interprétation à base de modèle comme un problème d'abduction : l'image représente l'observation et la tâche d'interprétation revient à fournir la « meilleure » explication en se fondant sur la connaissance terminologique du domaine exprimée dans une logique de description.

Mots-clés : Abduction, logiques de description, morphologie mathématique, analyse formelle de concepts, interprétation d'images.

1 Introduction

Les logiques de description (LD) sont désormais un paradigme classique de la représentation des connaissances (Baader (2003)), avec des champs d'application dans des domaines aussi divers que le web sémantique, la robotique cognitive, la vision par ordinateur, ainsi que l'interprétation sémantique d'images. L'interprétation d'images peut bénéficier d'une connaissance experte exprimée sous forme d'une ontologie et des outils de raisonnement offerts par les LD, comme illustré par nos travaux (Hudelot *et al.* (2008); Hudelot *et al.* (2010)). Nous avons proposé des outils de raisonnement capables de traiter de façon unifiée l'information quantitative provenant de l'image et la connaissance experte, qualitative, décrite par une ontologie. La tâche d'interprétation est réalisée de manière séquentielle, en maintenant la cohérence entre les informations extraites

de l'image et la connaissance experte. L'interprétation est ainsi exprimée comme un processus de satisfaisabilité de la situation courante (une configuration spatiale) encodée dans la ABox de la LD et sa TBox. Cependant, ce mode de raisonnement ne permet pas de gérer les situations où la connaissance experte n'est pas entièrement cohérente avec l'observation, ce qui est commun en interprétation d'images. En outre, une configuration spatiale donnée peut être cohérente avec différents éléments de la connaissance *a priori* (ou cohérente à un certain degré).

Nous proposons donc d'adapter les outils de raisonnement des LD à de telles situations, dans le cadre de raisonnements abductifs. Ce type de raisonnement permet le calcul de la « meilleure » explication des phénomènes observés, ce qui en fait un outil adapté aux situations où la connaissance n'est pas entièrement cohérente avec les observations. Formellement, étant données une théorie \mathcal{K} , représentant la connaissance experte, et une formule C , représentant une observation dans le domaine d'étude, l'abduction consiste à calculer la formule explicative D tel que D soit satisfaisable relativement à \mathcal{K} , ce qui peut s'écrire : $\mathcal{K} \models D \rightarrow C$ (ou $\mathcal{K} \cup D \models C$).

A titre d'illustration, nous reproduisons, dans une forme simplifiée, l'exemple d'interprétation de scènes (figure 1) introduit par Möller & Neumann (2008) à des fins de raisonnement abductif. La tâche d'interprétation consiste à étendre la base de connaissances avec de nouvelles assertions sur les structures présentes dans l'image, sur leurs relations, ainsi que sur le contexte global de la scène. A l'instar de Möller & Neumann (2008), nous fondons notre analyse sur une ontologie décrivant le domaine sportif et une série d'algorithmes de traitement et d'analyse d'images permettant d'extraire des objets initiaux de l'image. La tâche d'interprétation dans ce contexte consiste à expliquer la présence dans l'image d'objets tels qu'un être humain, une barre horizontale, un poteau, en tenant compte de la théorie sur le domaine sportif et des premiers objets reconnus dans l'image. Une réponse plausible à cette question est que l'image représente un saut à la perche accompli par un athlète. Le but de cet article est d'introduire des approches computationnelles permettant une telle inférence.



FIG. 1 – Un exemple d'interprétation d'image. La présence dans l'image d'objets tels qu'un être humain, une barre horizontale, un poteau peut être expliquée par un saut à la perche.

Nous proposons d'enrichir les logiques de description avec des outils de raisonnement abductif, dans un cadre algébrique, en associant de manière originale des concepts de morphologie mathématique, de logiques de description et d'analyse formelle de concepts (AFC). L'inférence de la meilleure explication d'une observation est calcu-

lée à partir du treillis de concepts en effectuant des érosions successives sur ce dernier. Nous montrons que ces opérateurs satisfont les postulats de rationalité du raisonnement abductif. Les éléments utiles pour la suite de ces différentes théories sont rappelés dans la section 2. Puis nous introduisons dans la section 3 des opérateurs morphologiques en AFC, et proposons des relations d’explication construites à partir de ces opérateurs.

2 Préliminaires

2.1 Logiques de description

Dans cette section, nous considérons la logique de description classique \mathcal{ALC} . Soient N_C et N_R deux ensembles dénombrables désignant les noms de concepts et les noms de rôles. Nous utilisons les lettres A et B pour les noms de concepts, la lettre R pour les noms de rôles, et les lettres C et D pour les concepts. Les symboles \top et \perp désignent les concepts universel et vide, respectivement. L’ensemble des \mathcal{ALC} -concepts est le plus petit ensemble tel que : (1) tout nom de concept est un concept ; (2) si C et D sont deux concepts et R un nom de rôle, les expressions suivantes désignent des concepts : $\neg C$ (négation), $C \sqcap D$ (conjonction), $C \sqcup D$ (disjonction), $\forall R.C$ (restriction de valeurs sur des noms de rôles), et $\exists R.C$ (restriction existentielle sur des noms de rôles). Une interprétation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ se compose d’un ensemble $\Delta^{\mathcal{I}}$, le domaine de \mathcal{I} , et d’une fonction $\cdot^{\mathcal{I}}$ qui associe tout concept C à un sous-ensemble $C^{\mathcal{I}}$ de $\Delta^{\mathcal{I}}$ et tout rôle R à un sous-ensemble $R^{\mathcal{I}}$ de $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ tel que, pour tous les concepts C , D et les rôles R , les propriétés suivantes sont satisfaites : (1) $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$ et $\perp^{\mathcal{I}} = \emptyset$, $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$, (2) $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$, $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$, (3) $(\exists R.C)^{\mathcal{I}} = \{x \mid \exists y \text{ s.t. } (x, y) \in R^{\mathcal{I}} \text{ et } y \in C^{\mathcal{I}}\}$, et (4) $(\forall R.C)^{\mathcal{I}} = \{x \mid \forall y, (x, y) \in R^{\mathcal{I}} \text{ implique } y \in C^{\mathcal{I}}\}$.

Une base de connaissances \mathcal{K} , associée à une DL, comporte deux parties : une partie terminologique (TBox) et une partie assertionnelle (ABox). La TBox \mathcal{T} décrit la terminologie en listant les concepts, les rôles et leurs relations. C’est un ensemble d’axiomes ou schémas. La ABox contient des informations sur les individus. Une interprétation \mathcal{I} est un modèle d’un axiome d’une LD (TBox ou ABox) ssi elle satisfait cet axiome, et est un modèle de la base de connaissances \mathcal{K} ssi elle satisfait tous les axiomes de \mathcal{K} .

Le service de raisonnement fondamental en LD est le calcul de la relation de subsumption entre deux concepts. Etant donné deux concepts C et D , on dit que D *subsume* C (noté $C \sqsubseteq D$) ssi $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ pour toute interprétation \mathcal{I} .

Dans l’exemple introduit en section 1, la théorie, ici la TBox \mathcal{T} décrivant le domaine sportif, est donnée dans la figure 2.

Man	\sqsubseteq	$Human$			
$Woman$	\sqsubseteq	$Human$		$Jumping_Event$	\sqsubseteq $Event \sqcap$
$Athlete$	\equiv	$Human \sqcap$			$\exists hasPart.Jumper \sqcap$
		$\exists hasProfession.Sport$		$Pole_Vault$	\sqsubseteq $Jumping_Event \sqcap$
$Jumper$	\sqsubseteq	$Athlete \sqcap$			$\exists hasPart.Pole \sqcap$
		$\exists use.SportEquipment$			$\exists hasPart.Horizontal_Bar \sqcap$
$Foam_Mat$	\sqsubseteq	$SportEquipment$			$\exists hasPart.Foam_Mat$
$Pole$	\sqsubseteq	$SportEquipment$		$High_Jump$	\sqsubseteq $Jumping_Event \sqcap$
$Javelin$	\sqsubseteq	$SportEquipment$			$\exists hasPart.Horizontal_Bar \sqcap$
$Horizontal_bar$	\sqsubseteq	$SportEquipment$			$\exists hasPart.Foam_Mat$
					...

FIG. 2 – Ontologie du domaine athlétique.

Les assertions initiales construites à partir des outils d’analyse d’images sont en-

codées dans une ABox, appelée ici \mathcal{A}_1 , comme suit : $\{pole_1 : Pole; human_1 : Human; bar_1 : Horizontal_Bar; \dots\}$.

2.2 Abduction et logiques de description

Dans le cadre des LD, l'abduction peut être considérée selon différents points de vue (Elsenbroich *et al.* (2006); Klarman *et al.* (2011)) : abduction de concepts, abduction de TBox, abduction de ABox et abduction de base de connaissances. Ici, bien que le cadre introduit soit général, nous considérons le cas de l'abduction de Abox.

Définition 1 (Abduction de ABox)

Soit \mathcal{L} une LD quelconque, \mathcal{K} une base de connaissances et A un ensemble d'assertions dans la Abox tel que pour tout $a \in A$ on a $\mathcal{K} \not\models \neg a$. Un Problème d'Abduction de Abox (PAA) noté $\langle \mathcal{K}, A \rangle$ consiste à trouver l'ensemble d'assertions γ tel que $\mathcal{K} \cup \gamma \models A$.

L'ensemble γ (cohérent avec \mathcal{K}) est appelé une explication de A . Le raisonnement explicatif consiste à sélectionner un sous-ensemble de cet ensemble d'explications. Il est donc nécessaire de disposer de critères de préférence pour effectuer cette sélection.

Revenons à notre exemple. L'interprétation vue comme un PAA $\langle \mathcal{K}, A \rangle$ se formule ainsi : $\mathcal{T} \cup \mathcal{A}_1 \cup \gamma \models A$, où A désigne $\{bar_1 : Horizontal_Bar, fm_1 : Foam_Mat, pole_1 : Pole, je_1 : Jumping_Event\}$. Intuitivement, un exemple d'assertions γ qui représente une solution du PAA serait $\{pv_1 : Pole_Vault\}$.

2.3 Analyse formelle de concepts

L'AFC est une théorie d'analyse de données, de représentation des connaissances et de gestion de l'information qui a pour but d'identifier des structures conceptuelles à partir des données (Ganter *et al.* (1999)). Elle se fonde sur une formalisation dans la théorie des treillis des notions de concepts et de hiérarchie de concepts. Un concept est vu comme une unité de pensée composée de deux parties : l'ensemble de tous les objets appartenant au concept, appelé l'extension, et l'ensemble des attributs de ces objets, appelé l'intention. Un *contexte formel* est un triplet $\mathbb{K} = (G, M, I)$, où G désigne l'ensemble d'objets, M l'ensemble d'attributs et $I \subseteq G \times M$ une relation entre les objets et les attributs. La paire $(g, m) \in I$ signifie « l'objet g a comme attribut m ». Les *concepts formels* du contexte \mathbb{K} sont les paires (A, B) avec $A \subseteq G$ et $B \subseteq M$ telles que (A, B) soit maximale avec la propriété $A \times B \subseteq I$. L'ensemble A se dénomme l'extension et l'ensemble B l'intention du concept formel (A, B) . L'ensemble des concepts formels d'un contexte peut être ordonné par inclusion de leur extension : $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2 (\Leftrightarrow B_2 \subseteq B_1)$. Cet ordre, reflétant la relation sousconcept-superconcept, induit toujours un treillis complet appelé *treillis de concepts* du contexte (G, M, I) . Pour $A \subseteq G$ et $B \subseteq M$, les opérateurs de dérivation α et β sont définis par $\alpha(A) = \{m \in M \mid \forall g \in A, (g, m) \in I\}$, et $\beta(B) = \{g \in G \mid \forall m \in B, (g, m) \in I\}$. Pour $A_1 \subseteq A_2 \subseteq G$ (resp. $B_1 \subseteq B_2 \subseteq M$), on a : (i) $\alpha(A_2) \subseteq \alpha(A_1)$ (resp. $\beta(B_2) \subseteq \beta(B_1)$), (ii) $A_1 \subseteq \beta(\alpha(A_1))$ et $\alpha(A_1) = \alpha(\beta(\alpha(A_1)))$ (resp. $B_1 \subseteq \alpha(\beta(B_1))$ et $\beta(B_1) = \beta(\alpha(\beta(B_1)))$). La paire (A, B) avec $A \subseteq G$ et $B \subseteq M$ est un concept formel ssi $\alpha(A) = B$ et $\beta(B) = A$. Dans le treillis de concepts, l'infimum et le supremum sont calculés comme suit :

$$\bigwedge_{t \in T} (A_t, B_t) = \left(\bigcap_{t \in T} A_t, \alpha(\beta(\bigcup_{t \in T} B_t)) \right), \quad \bigvee_{t \in T} (A_t, B_t) = \left(\beta(\alpha(\bigcup_{t \in T} A_t)), \bigcap_{t \in T} B_t \right)$$

où T désigne l'ensemble des concepts formels. Tout treillis de concepts est un treillis complet. Un treillis complet (L, \leq) est isomorphe au treillis de concepts $\mathbb{C}(L, L, \leq)$.

2.4 AFC et logiques de description

L'idée initiale d'exploiter l'AFC dans le cadre des LD était motivée par la définition d'outils efficaces de construction d'une hiérarchie de subsomption étendue d'un ensemble de concepts d'une LD, dans le but de mettre en exergue des dépendances hiérarchiques que les algorithmes classiques de classification ne pouvaient exhiber (Baader (1995)). A cette fin, la notion de *contexte sémantique* a été introduite. Les attributs de ce contexte formel sont les concepts de la LD, ses objets sont définis comme étant l'ensemble des contre-exemples de la relation de subsomption $C \sqsubseteq_{\mathcal{T}} D$ (interprétations \mathcal{I} de \mathcal{T} et un élément du domaine d'interprétation $d \in \Delta^{\mathcal{I}}$ tel que $d \in C^{\mathcal{I}} \setminus D^{\mathcal{I}}$).

Définition 2 (Contexte sémantique (Baader (1995)))

Le contexte sémantique $\mathbb{K}_{\mathcal{T}} := (G, M, I)$ est défini comme suit :

$$G := \{(\mathcal{I}, d) \mid \mathcal{I} \text{ est un modèle de } \mathcal{T} \text{ et } d \in \Delta^{\mathcal{I}}\} \quad (1)$$

$$M := \{m_1, \dots, m_n\} \quad (2)$$

$$I := \{((\mathcal{I}, d), m) \mid d \in m^{\mathcal{I}}\} \quad (3)$$

Dans ce qui précède, \mathcal{T} est une TBox et $\{m_1, \dots, m_n\}$ sont les concepts définis dans \mathcal{T} . Un résultat important obtenu par Baader (1995) stipule que le treillis de concepts associé au contexte sémantique $\mathbb{K}_{\mathcal{T}}$ est isomorphe à la hiérarchie de subsomption de toutes les conjonctions des sous-ensembles M par rapport à \mathcal{T} . L'algorithme d'exploration d'attributs est utilisé pour calculer la hiérarchie de subsomption des conjonctions des concepts d'une LD. Plus récemment, un nouveau contexte formel, appelé « contexte syntaxique », a été introduit par Baader *et al.* (2007), plus approprié au calcul du *plus petit subsumant commun* d'une conjonction de concepts et préservant la propriété d'isomorphisme.

Afin de définir nos opérateurs d'abduction sur le treillis de conjonctions et de disjonctions, nous considérons le contexte formel sémantique ; la construction du treillis associé repose sur l'algorithme *exploration distributive de concepts* de Stumme (1996, 1998).

Dans notre exemple, les objets G du contexte sont alors des individus (interprétations) et l'ensemble des attributs M est l'ensemble des concepts (y compris des rôles) définis dans \mathcal{T} décrits dans la figure 2. Nous donnons un exemple de contexte dans le tableau 1, et un extrait du treillis de concepts associé dans la figure 3.

3 Opérateurs d'abduction par morphologie mathématique sur des treillis complets

Soient (L, \preceq) et (L', \preceq') deux treillis complets (qui ne sont pas nécessairement égaux). Tous les résultats et définitions suivants dérivent du cadre algébrique de la morphologie mathématique (MM) (Bloch *et al.* (2007); Heijmans (1994); Heijmans & Ronse (1990); Ronse & Heijmans (1991); Serra (Ed.) (1988)).

TAB. 1 – Contexte formel $\mathbb{K}_{athletic}$ correspondant à la TBox de la figure 2.

$\mathbb{K}_{athletic}$	Human	Man	Woman	Athlete	Jumper	SportEquipment	Foam_Mat	Pole	Javelin	Horizontal_bar	Event	Jumping_Event	Pole_Vault	High_Jump
m_1	X	X												
jumper_1	X	X		X	X	X		X						
pole_1						X								
bar_1						X				X				
fm_1						X	X							
je_1					X	X					X	X		
jav_1						X			X					
pv_1					X	X	X	X		X	X	X	X	
hj_1					X	X	X			X	X	X		X

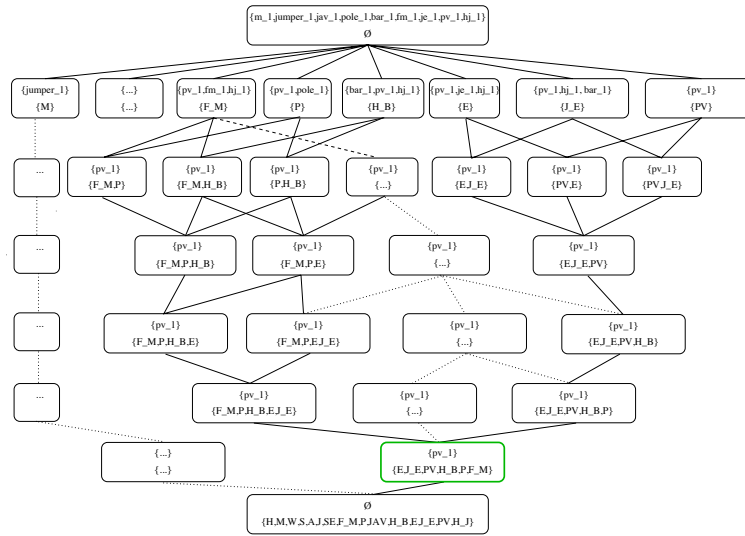


FIG. 3 – Extrait du treillis de concepts induit par le contexte formel $\mathbb{K}_{athletic}$.

Définition 3

Un opérateur $\delta : L \rightarrow L'$ est une dilatation s'il commute avec le supremum : $\forall(x_i) \in L, \delta(\bigvee_i x_i) = \bigvee'_i \delta(x_i)$, où \bigvee désigne le supremum associé à \preceq et \bigvee' celui associé à \preceq' . Un opérateur $\varepsilon : L' \rightarrow L$ est une érosion s'il commute avec l'infimum : $\forall(x_i) \in L', \varepsilon(\bigwedge'_i x_i) = \bigwedge_i \varepsilon(x_i)$, où \bigwedge et \bigwedge' désignent l'infimums associés à \preceq et \preceq' , respectivement.

Toutes les propriétés classiques de la morphologie mathématique sont vérifiées ici (Bloch et al. (2007); Heijmans (1994); Heijmans & Ronse (1990)), et ne sont donc pas rappelés.

3.1 MM sur des treillis de concepts

Dans le but de résoudre un PAA, il est opportun de raisonner sur les sous-ensembles de G . Ainsi, nous considérons le treillis complet $(\mathcal{P}(G), \subseteq)$ et définissons des opérateurs de $\mathcal{P}(G)$ vers $\mathcal{P}(G)$, où $\mathcal{P}(G)$ est l'ensemble des sous-ensembles de G (i.e les

interprétations). Puisque l'ordre (l'inclusion) défini sur G est équivalent (voir section 2.3) à celui dans M , raisonner dans G donne directement des résultats dans M .

Afin de définir des opérations dans $\mathcal{P}(G)$, nous introduisons des érosions et des dilatations particulières, dites morphologiques (Serra (1982)), qui reposent sur la notion d'élément structurant (une relation binaire b entre des éléments de G). Pour $g \in G$, nous notons par $b(g)$ l'ensemble des éléments de G en relation avec g .

Définition 4

La dilatation morphologique d'un sous-ensemble A de G par rapport à b s'écrit :

$$\delta_b(A) = \{g \in G \mid b(g) \cap A \neq \emptyset\}. \quad (4)$$

L'érosion morphologique de A s'écrit :

$$\varepsilon_b(A) = \{g \in G \mid b(g) \subseteq A\}. \quad (5)$$

La relation b peut représenter un système de voisinage des interprétations ou une relation de distance. Dans le cas d'une distance d entre les éléments de G , les éléments structurants peuvent être définis comme des boules de cette distance. Cet exemple est particulièrement intéressant pour le raisonnement abductif où la partie « la plus centrale » d'un modèle correspond intuitivement à l'explication préférée. L'érosion s'exprime alors par :

$$\varepsilon^n(A) = \{g \in G \mid d(g, G \setminus A) > n\}. \quad (6)$$

On note $\varepsilon(A) = \varepsilon^1(A)$, et on a $\varepsilon^0(A) = A$. Ici G est un espace discret fini, et par conséquent seules les valeurs entières de n sont considérées. La distance peut généralement être induite à partir du treillis.

Propriétés. Comme dit précédemment, toutes les propriétés classiques de la MM sont vraies dans ce contexte. Nous en donnons les plus importantes :

- L'érosion commute avec l'infimum : $\forall (A, A') \in \mathcal{P}(G)^2, \varepsilon(A \cap A') = \varepsilon(A) \cap \varepsilon(A')$.
- Pour le supremum on a seulement une inclusion : $\forall (A, A') \in \mathcal{P}(G)^2, \varepsilon(A) \cup \varepsilon(A') \subseteq \varepsilon(A \cup A')$.
- Si $g \in b(g)$, alors l'érosion est anti-extensive : $\forall A \in \mathcal{P}(G), \varepsilon_b(A) \subseteq A$. Cette propriété est en particulier vérifiée pour l'équation 6.
- Itérativité : $\varepsilon^n(\varepsilon^m(A)) = \varepsilon^{n+m}(A)$. Appliquer des érosions successives donne des résultats de plus en plus petits, équivalents à une application directe d'une érosion plus grande. Cette propriété sera exploitée pour définir des explications comme le résultat le plus petit possible d'une succession d'érosions.
- Une notion importante est celle d'adjonction : une paire d'opérateurs (ε, δ) forme une adjonction si $\forall x \in L, \forall y \in L', \delta(x) \preceq' y \Leftrightarrow x \preceq \varepsilon(y)$. Si (ε, δ) est une adjonction, alors ε est une érosion et δ une dilatation. En particulier, $(\varepsilon_b, \delta_b)$ est une adjonction. Cette notion est équivalente à celle de connexion de Galois, en inversant l'ordre sur le second treillis $(\mathcal{P}(M), \subseteq)$: pour un contexte formel (A, B) , $A \subseteq \beta(B) \Leftrightarrow B \subseteq \alpha(A)$. Ainsi les opérateurs de dérivation en AFC peuvent être interprétés en termes de morphologie mathématique (Bloch *et al.* (2007)).

3.2 Dernière érosion non vide

Comme montré par Bloch *et al.* (2001) dans le cadre de la logique propositionnelle, les érosions peuvent être utilisées pour trouver des explications. L'idée est de sélectionner la *partie la plus centrale* d'une formule comme étant la meilleure explication. Il

a été démontré que cette approche a de bonnes propriétés et satisfait les postulats de rationalité du raisonnement abductif (Pino-Pérez & Uzcátegui (1999)).

Ici nous suivons le même cheminement, mais adapté au contexte des treillis de concepts, en utilisant la définition des érosions introduite dans l'équation 6. Pour tout $A \subseteq G$, nous définissons son dernier érodé non vide par :

$$\varepsilon_\ell(A) = \varepsilon^n(A) \Leftrightarrow \begin{cases} \varepsilon^n(A) \neq \emptyset, \\ \text{et } \forall m > n, \varepsilon^m(A) = \emptyset. \end{cases} \quad (7)$$

Cette dernière érosion non vide définit le sous-ensemble de modèles et d'interprétations dans G qui sont les plus éloignés du complémentaire de A (en considérant la distance d), donc les plus centraux dans A .

Définition 5

Une explication préférée γ de A est définie à partir de la dernière érosion non vide par :

$$A \triangleright^{\ell ne} \gamma \stackrel{def}{\Leftrightarrow} \gamma \subseteq \varepsilon_\ell(A). \quad (8)$$

Lorsque l'on ajoute une contrainte définie par une théorie \mathcal{K} , alors cette définition est modifiée selon :

$$A \triangleright^{\ell ne} \gamma \stackrel{def}{\Leftrightarrow} \gamma \subseteq \varepsilon_\ell(\mathcal{K} \cap A). \quad (9)$$

Notons que cela définit un ensemble de meilleures explications possibles, qui n'est pas forcément un singleton. Cet ensemble est robuste dans le sens où il peut être modifié tout en restant dans A . Par exemple, dilater γ par une boule de d de taille plus petite que n conduit toujours à un sous-ensemble de A . La partie centrale peut être interprétée comme étant le sous-ensemble de modèles qui peut être le plus modifié tout en restant dans A . Dans le treillis de concepts, cette définition s'interprète de la manière suivante : partant d'un sous-ensemble à expliquer, les érosions successives reviennent à « descendre » dans le treillis autant que possible jusqu'à atteindre un ensemble non vide d'interprétations.

Dans l'exemple où $A = \{\text{bar}_1 : \text{Horizontal_Bar}, \text{fm}_1 : \text{Foam_Mat}, \text{pole}_1 : \text{Pole}, \text{je}_1 : \text{Jumping_Event}\}$, la partie du treillis de concepts correspondant à ce sous-ensemble (en rouge dans la table 1) est donnée figure 3. La dernière érosion non vide de ce sous-ensemble correspond à l'avant-dernière ligne, et $\gamma = \{\text{pv}_1\}$ (en vert sur la figure 3). Ce résultat correspond à l'intuition, et le saut à la perche est effectivement une explication raisonnable de la description donnée par A . Dans un contexte d'interprétation d'images, ce processus permet de déduire des descriptions de haut niveau à partir d'informations de plus bas niveau extraites des images.

3.3 Dernière érosion cohérente

Une autre façon d'introduire la contrainte liée à la théorie est d'éroder celle-ci jusqu'à ce qu'elle devienne cohérente avec A . Cela conduit à la définition d'une nouvelle relation d'explication.

Définition 6

Une explication préférée γ de A est définie à partir de la dernière érosion cohérente comme suit :

$$A \triangleright^{\ell c} \gamma \stackrel{def}{\Leftrightarrow} \gamma \subseteq \varepsilon_{\ell c}(\mathcal{K}, A) \cap A,$$

où ε_{lc} désigne la dernière érosion cohérente :

$$\varepsilon_{lc}(\mathcal{K}, A) = \varepsilon^n(\mathcal{K}) \text{ où } n = \max\{k \mid \varepsilon^k(\mathcal{K}) \cap A \neq \emptyset\}.$$

Cette définition a une interprétation différente : nous considérons cette fois l'érosion de \mathcal{K} , ce qui signifie que l'on cherche les modèles qui sont dans A et qui sont les plus centraux dans la théorie.

3.4 Propriétés et interprétations

Une première propriété intéressante est que raisonner sur G revient à raisonner sur l'ensemble de concepts formels. Ici, les explications résultant d'un raisonnement sur la ABox, conduisent à des érosions des sous-ensembles de G (modèles et interprétations). Soit (A, B) un concept formel, avec $A \subseteq G$. Les définitions des explications dans A induisent les concepts correspondants pour B en utilisant les opérateurs de dérivation $\alpha(\gamma) = \{m \in M \mid \forall g \in \gamma, (g, m) \in I\}$.

Dans notre exemple, $B = \alpha(A) = \{SE\}$ et $\alpha(\gamma) = \{Foam_Mat, Pole, Horizontal_Bar, Event, Jumping_Event, Pole_Vault\}$ est une explication de B .

Notons que éroder A revient à dilater B , ce qui est conforme à la correspondance entre la propriété de connexion de Galois entre les opérateurs de dérivation et les propriétés d'adjonction de la dilatation et l'érosion (section 3.1).

Considérons à présent les postulats de rationalité introduits par Pino-Pérez & Uzcátegui (1999) pour les relations d'explication. Il a été prouvé que la plupart d'entre eux sont satisfaits pour les opérateurs d'explication (dernière érosion non vide et dernière érosion cohérente) (Bloch *et al.* (2001)). Dans le cadre proposé ici, ces résultats sont également vérifiés et sont directement déduits des propriétés algébriques de l'érosion (section 3.1) :

- $\triangleright^{\ell ne}$ et $\triangleright^{\ell c}$ sont indépendants de la syntaxe (puisque'ils sont calculés dans l'espace des interprétations).
- Puisque les explications sont définies comme un sous-ensemble d'un résultat d'érosion, tout sous-ensemble non-vide d'une explication est une explication.
- Les définitions sont cohérentes dans le sens où $G \setminus A \not\subseteq \mathcal{K}$ ssi $\exists \gamma, A \triangleright \gamma$.
- La propriété de réflexivité est vraie pour les deux définitions : si $A \triangleright \gamma$, alors $\gamma \triangleright \gamma$.
- Disjonction d'explications : si $A \triangleright \gamma$ et $A \triangleright \delta$, alors $A \triangleright (\gamma \cup \delta)$, pour les deux définitions. Cela montre que s'il existe différentes explications possibles, leur disjonction est une explication.
- Disjonction à gauche : si $A \triangleright^{\ell c} \gamma$ et $A' \triangleright^{\ell c} \gamma$, alors $(A \cup A') \triangleright^{\ell c} \gamma$ (puisque l'érosion est appliquée sur \mathcal{K}). En revanche cette propriété n'est pas vraie pour $\triangleright^{\ell ne}$ puisque l'érosion ne commute pas avec le supremum.
- Pour les mêmes raisons, nous avons la propriété suivante pour $\triangleright^{\ell c}$: si $A \triangleright^{\ell c} \gamma$ et $A' \triangleright^{\ell c} \delta$, alors $(A \cup A') \triangleright^{\ell c} \gamma$ ou $(A \cup A') \triangleright^{\ell c} \delta$, mais cela n'est pas vrai pour $\triangleright^{\ell ne}$.
- Pour les conjonctions, une propriété de monotonie est vérifiée pour $\triangleright^{\ell c}$: si $A \triangleright^{\ell c} \gamma$ et $\gamma \subseteq A'$, alors $(A \cap A') \triangleright^{\ell c} \gamma$. Pour $\triangleright^{\ell ne}$, seule une forme faible est vérifiée : si $A \triangleright^{\ell ne} \gamma$ et $A' \triangleright^{\ell ne} \gamma$, alors $(A \cap A') \triangleright^{\ell ne} \gamma$. Notons néanmoins que cette forme faible est naturelle et intéressante.

Des propriétés équivalentes dans le formalisme des LD en sont directement héritées, grâce à la construction et aux liens entre les différentes théories décrits dans la section 2.

4 Conclusion

A des fins d'interprétation d'images, nous avons proposé des outils calculatoires pour considérer le raisonnement abductif dans le cadre des logiques de description. Cela a été possible grâce aux outils venant de l'analyse formelle de concepts, la morphologie mathématique, et le cadre général de la théorie des ensembles et des treillis. Les propriétés et interprétations des opérateurs d'explication ont été analysées, et les postulats de rationalité ont été prouvés. Un exemple d'application a accompagné le développement de ces méthodes pour montrer leur utilité en pratique. Il convient cependant d'étendre l'étude de cet exemple, en prenant en compte les aspects de complexité qui n'ont pas été abordés dans cet article.

Références

- BAADER F. (1995). Computing a minimal representation of the subsumption lattice of all conjunctions of concepts defined in a terminology. In *Knowledge Retrieval, Use and Storage for Efficiency : 1st International KRUSE Symposium*, p. 168–178.
- BAADER F. (2003). *The description logic handbook : theory, implementation, and applications*. Cambridge University Press.
- BAADER F., SERTKAYA B. & TURHAN A. (2007). Computing the least common subsumer wrt a background terminology. *Journal of Applied Logic*, **5**(3), 392–420.
- BLOCH I., HEIJMANS H. & RONSE C. (2007). Mathematical Morphology. In M. AIELLO, I. PRATT-HARTMAN & J. VAN BENTHEM, Eds., *Handbook of Spatial Logics*, chapter 13, p. 857–947. Springer.
- BLOCH I., PINO-PÉREZ R. & UZCÁTEGUI C. (2001). Explanatory Relations based on Mathematical Morphology. In *ECSQARU 2001*, p. 736–747, Toulouse, France.
- ELSENBRICH C., KUTZ O. & SATTLER U. (2006). A case for abductive reasoning over ontologies. In *OWL : Experiences and Directions*, Athens, Georgia, USA.
- GANTER B., WILLE R. & WILLE R. (1999). *Formal concept analysis*. Springer Berlin.
- HEIJMANS H. J. A. M. (1994). *Morphological Image Operators*. Academic Press, Boston.
- HEIJMANS H. J. A. M. & RONSE C. (1990). The Algebraic Basis of Mathematical Morphology – Part I : Dilations and Erosions. *Computer Vision, Graphics and Image Processing*, **50**, 245–295.
- HUDELLOT C., ATIF J. & BLOCH I. (2008). Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, **159**(15), 1929–1951.
- HUDELLOT C., ATIF J. & BLOCH I. (2010). Integrating bipolar fuzzy mathematical morphology in description logics for spatial reasoning. In *European Conference on Artificial Intelligence ECAI 2010*, p. 497–502, Lisbon, Portugal.
- KLARMAN S., ENDRIS U. & SCHLOBACH S. (2011). ABox Abduction in the Description Logic. *Journal of Automated Reasoning*, **46**(1), 43–80.
- MÖLLER R. & NEUMANN B. (2008). Ontology-based reasoning techniques for multimedia interpretation and retrieval. *Semantic Multimedia and Ontologies*, p. 55–98.
- PINO-PÉREZ R. & UZCÁTEGUI C. (1999). Jumping to Explanations versus jumping to Conclusions. *Artificial Intelligence*, **111**, 131–169.
- RONSE C. & HEIJMANS H. J. A. M. (1991). The Algebraic Basis of Mathematical Morphology – Part II : Openings and Closings. *Computer Vision, Graphics and Image Processing*, **54**, 74–97.
- SERRA J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, New-York.
- SERRA (ED.) J. (1988.). *Image Analysis and Mathematical Morphology, Part II : Theoretical Advances*. Academic Press, London.
- STUMME G. (1996). The concept classification of a terminology extended by conjunction and disjunction. In *PRICAI'96 : Topics in Artificial Intelligence*, volume LNCS 1114, p. 121–131.
- STUMME G. (1998). Distributive concept exploration—a knowledge acquisition tool in formal concept analysis. In *KI-98 : Advances in Artificial Intelligence*, p. 117–128 : Springer.

Quels formalismes temporels pour représenter des connaissances extraites de textes de recettes de cuisine ?

Valmi Dufour-Lussier^{1,2}, Florence Le Ber^{1,3}, and Jean Lieber^{1,4}

¹ LORIA (CNRS, INPL, INRIA, UN2, UHP)
615, rue du Jardin Botanique, 54602 Villers-lès-Nancy
valmi.dufour@loria.fr, jean.lieber@loria.fr

² Université Nancy 2

³ LHYGES (CNRS, ENGEES, Université de Strasbourg)
1, quai Koch, 67000 Strasbourg
florence.leber@engees.unistra.fr

⁴ Université Henri Poincaré, Nancy 1

Résumé : Le projet TAAABLE a pour objet de construire un système de raisonnement à partir de cas pour la recherche et l'adaptation de textes de recettes de cuisine. Dans ce cadre, nous nous intéressons aux aspects temporels des recettes et à la manière de les représenter afin d'en envisager l'adaptation textuelle.

Mots-clés : Adaptation, algèbres temporelles, traitement automatique des langues

1 Introduction

TAAABLE (Badra *et al.*, 2008) est un système informatique destiné à résoudre des problèmes de cuisine, développé dans le cadre du *Computer Cooking Contest*¹, dans le cadre duquel des requêtes du type « je désire la recette d'un plat de pâtes aux lentilles » doivent être résolues en cherchant dans un livre fourni par les organisateurs du concours une recette qui répond aux critères de l'utilisateur. Si une telle recette n'existe pas, un moteur de raisonnement à partir de cas sélectionne une recette similaire et suggère une adaptation. Initialement, l'adaptation ne portait que sur la liste d'ingrédients, mais la dernière version du système (Blanché *et al.*, 2010) tient compte des instructions textuelles de préparation dans le raisonnement et l'adaptation. Le raisonnement ne pouvant porter directement sur du texte brut, il est nécessaire de choisir un formalisme de représentation de cas et d'employer des méthodes de traitement automatique des langues pour créer ces représentations. TAAABLE utilise actuellement un formalisme *ad hoc*

1. <http://liris.cnrs.fr/ccc/ccc2010>

Lutheran Hotdish

- 1 lb Hamburger
- 1/2 lb Mild or spicy sausage
- 1 lg Onion (sliced and quartered) (up to)
- 5 Cloves garlic (minced) (up to)
- 2 cn Kidney beans (drained)
- 1 lb Uncooked pasta (i.e. elbow ; twisted ; wagon wheels, shells ; etc) (up to)

Brown hamburger and sausage with onion, garlic, and all seasonings. Meanwhile, prepare the pasta per pkg instructions. In a large pan, combine all ingredients. Add enough tomato sause until mixture is well coated, but no soupy. When well mixed, pour into greased and covered casserol dish and bake at 350F for 1hr. Remove cover for the last 15 minutes. Then enjoy.

FIGURE 1 – Une recette issue du livre fourni par le *Computer Cooking Contest*, illustrant certains phénomènes temporels intéressants.

fondé sur les combinaisons d'ingrédients pour ce faire, bien que Le Ber *et al.* (2009, 2010) aient initié une réflexion quant à un formalisme plus expressif et plus approprié.

Cet article poursuit cette réflexion. Dans la section 2, nous présentons les différents types de phénomènes temporels observés dans les textes de recette, des plus simples à représenter aux plus complexes. Puis, dans la section 3, nous définissons le type de raisonnement que nous entendons appliquer aux recettes avant de passer en revue, dans la section 4, quelques formalismes qui semblent appropriés à la représentation de ces phénomènes. La section 5 présente de façon concise l'adéquation que nous observons entre les phénomènes à modéliser et les formalismes étudiés. Finalement, dans la section 6, nous concluons et présentons succinctement les travaux futurs que nous comptons accomplir à partir de nos résultats.

2 Problématiques de représentation temporelle

Nous dressons d'abord une liste des phénomènes temporels que le formalisme retenu devrait être capable de représenter. La figure 1 présente une recette du livre employé par TAAABLE pour répondre aux requêtes des utilisateurs, reproduite telle quelle.

2.1 Durée

Pour les événements ayant une durée, celle-ci peut être définie explicitement de manière quantitative, comme c'est le cas des dernières instructions de la recette montrée en figure 1 (« bake at 350F for 1hr »), ou encore implicitement en référence à un changement d'état (« Add enough tomato sause until mixture is well coated »), ce que nous appellerons « durée qualitative ». Certaines recettes peuvent préciser les deux (« Bake at 350 degrees for about 25 minutes, or until lightly browned »), auquel cas une sémantique doit être déterminée : s'agit-il réellement d'une disjonction — l'action se conclut-elle dès que 25 minutes se sont écoulées *ou* que la couleur a changé ?

2.2 Ordre

Les actions nécessaires à l'exécution d'une recette de cuisine ne peuvent pas être effectuées dans n'importe quel ordre. La plupart des actions sont ordonnées implicitement, puisqu'il est entendu que, dans une recette de cuisine, les actions décrites à même la liste d'ingrédients sont à accomplir avant le début de la recette proprement dite, et que les actions décrites dans le texte doivent être effectuées dans l'ordre où elles sont énoncées, sauf mention contraire.

Dans la recette de la figure 1, les actions préliminaires (couper l'oignon, hacher l'ail et déguster les haricots) doivent être effectuées avant que la première action décrite dans le texte (« Brown hamburger and sausage... ») soit initiée. Cependant il semble que ces actions puissent être exécutées dans n'importe quel ordre, puisqu'elles ne dépendent pas l'une des autres : on peut couper l'oignon avant de hacher l'ail ou après, car l'aliment résultant de chacune de ces actions n'est pas utilisé pour l'autre.

Dans le texte même, certains adverbes ont pour effet de changer l'ordre normal d'exécution des actions. Par exemple, l'action décrite par la deuxième phrase (« Meanwhile, prepare pasta... ») doit être effectuée simultanément à la première action. La troisième phrase (« In a large pan, combine all ingredients ») ne précise pas de relation d'ordre explicite, de quoi on peut conclure que l'action doit être effectuée à la suite de l'action précédente.

Il est à noter que le moment d'exécution d'une action peut être défini en référence à la fin d'un intervalle, comme dans « Remove cover for the last 15 minutes. » Cet intervalle de référence peut être imprécis, comme dans « simmer 2–3 hours », ajoutant une difficulté additionnelle.

2.3 Actions répétées

Les actions répétées suggèrent une structure non linéaire du temps. Différentes situations peuvent poser différents problèmes de représentation. Une même action peut être répétée un nombre déterminé ou indéterminé (jusqu'à l'obtention d'un état donné) de fois, ou une série d'actions peuvent être répétées en alternance, possiblement avec des conditions d'ordre précises. Une action peut également être répétée de façon sporadique sur un intervalle donné, comme dans « simmer 2–3 hours, stirring occasionally ».

2.4 Disjonction

Les disjonctions d'actions constituent un dernier élément pouvant poser des difficultés de représentation. Celles-ci peuvent être liées à un état, ou encore représenter des alternatives dans une recette : « If you want this relish to be really hot then take 4 green chillis, seed them and chop finely, and put them in the frying pan at the same time as the onions. »

3 Raisonnement

Avant de discuter des formalismes que nous envisageons d'utiliser pour représenter les recettes, il convient de mieux définir le type de raisonnement que nous souhaitons appliquer à ces recettes. Si l'utilisateur lance la requête donnée en exemple dans l'introduction, c'est-à-dire un plat de pâtes aux lentilles, TAAABLE retournera comme résultat la recette de la figure 1 en suggérant de remplacer les haricots rouges en conserve par des lentilles.

Pour adapter le texte de préparation de manière adéquate, il ne suffira pas de remplacer toutes les instances d'« haricots » par « lentilles ». Il faut également introduire au bon endroit des informations issues des connaissances du domaine : par exemple, que les lentilles doivent cuire à l'eau *pendant* une demi-heure *puis* être égouttées *avant* d'intégrer la préparation. Pour représenter adéquatement les relations entre ces actions (pendant, puis, avant), il sera nécessaire d'employer une représentation temporelle.

Une manière de tenir compte de connaissances liées à un élément de substitution en raisonnement à partir de cas est l'adaptation par révision, proposée par Cojan & Lieber (2008) : il s'agit d'ajouter ces connaissances à celles issues du cas remémoré (ici, la recette du « Lutheran Hotdish ») en résolvant les contradictions éventuelles à l'aide d'un opérateur de révision des croyances, de manière à conserver toute l'information issue des connaissances du domaine et le maximum d'information issue du cas.

Un tel opérateur applicable aux algèbres qualitatives pourrait être dérivé des travaux de Condotta *et al.* (2009) sur la fusion des réseaux de contraintes qualitatives.

Une fois la représentation de la préparation de la recette adaptée par révision avec les connaissances du domaine, et à condition de pouvoir maintenir le lien entre le texte et cette représentation au cours de l'adaptation, il sera possible d'apporter les modifications correspondantes au texte.

4 Formalismes envisagés

Les formalismes de représentation temporelle sont généralement divisés en deux catégories par la communauté IA, en fonction du type de raisonnement envisagé. Les raisonnements sur les contraintes temporelles font généralement appel à des formalismes algébriques tels que les algèbres d'intervalle, que nous étudierons en particulier. Les raisonnements sur les actions et leurs effets sur des domaines dynamiques, quant à eux, donnent le plus souvent lieu à l'utilisation de formalismes inspirés du calcul des situations, qui sont moins appropriés à notre problème².

Nous nous pencherons brièvement sur des formalismes dont le but principal est de résoudre certains problèmes de représentation des informations temporelles dans les textes, ainsi que sur les flux opérationnels (*workflows*), représentant des processus et pouvant être utilisés pour raisonner sur ceux-ci.

2. Nous n'entendons pas que ces formalismes ne seraient pas utiles pour l'adaptation dans TAAABLE, et au contraire nous souhaitons étudier la contribution qu'ils pourraient apporter ultérieurement, mais le problème que nous cherchons à résoudre ici est relatif aux contraintes temporelles entre les actions, et non pas au raisonnement sur les actions et leurs effets.

4.1 Algèbres de relations

4.1.1 Contraintes qualitatives

Nous débutons notre exploration par l’algèbre des intervalles de Allen (1983), qui nous semble particulièrement adaptée à la représentation des phénomènes les plus fréquents. Cet algèbre permet d’exprimer des relations entre des intervalles temporels sous la forme de disjonctions entre treize relations de base : *before* (notée *b*, < ou *p*), *meets* (*m*), *overlaps* (*o*), *during* (*d*), *starts* (*s*), *finishes* (*f*), les inverses de ces six relations (respectivement *bi* — ou encore >, *pi* ou *a* —, *mi*, *oi*, *di*, *si* et *fi*), ainsi que *equals* (*e*, *eq* ou *=*).

En associant chaque action à un intervalle, il est facile d’en représenter l’ordre. Dans la recette de la figure 1, on aurait par exemple :

- ‘mince garlic’ {*b*} ‘brown hamburger’
- ‘prepare pasta’ {*d*,*f*} ‘brown hamburger’
- ‘combine all ingredients’ {*bi*,*m*} ‘prepare pasta’

En associant des états à des intervalles, on pourrait représenter les durées définies par un changement d’état, comme dans le cas de « ‘add tomato sauce’ {*m*} ‘mixture is well coated’ ».

Grâce à *IN²DU* (Pujari *et al.*, 1999), qui permet additionnellement d’exprimer des contraintes qualitatives sur la durée des intervalles (par exemple, un intervalle I_1 dont la fin coïncide avec le début d’un intervalle I_2 et dont la durée est inférieure à celle de ce dernier s’exprime « $I_1 \{m^{\leq}\} I_2$ »), il serait également possible de représenter jusqu’à un certain point des durées quantitatives en définissant des intervalles de durée fixe, par exemple : « ‘bake’ {*?=*} ‘une heure’ ». Cette même méthode nous permettrait de représenter aussi la dernière action de la recette : « ‘remove cover’ {*s?*} ‘15 minutes’ {*f?*} ‘bake’ ». Pour représenter la durée d’une qui est exprimée à la fois quantitativement et qualitativement, comme dans « Bake at 350 degrees for about 25 minutes, or until lightly brown », nous pouvons prendre le parti d’exprimer que la durée de 25 minutes constitue un maximum et écrire « ‘bake’ {*m?*} ‘is brown’, ‘bake’ {*?≤*} ‘25 minutes’ ».

Ladkin (1986) et Ligozat (1991), entre autres, ont proposé des extensions à l’algèbre de Allen afin qu’elle puisse exprimer des contraintes sur des intervalles non convexes. En définissant une propriété « cet intervalle est non convexe » applicable aux intervalles représentant une action qui se répète avec au moins deux occurrences disjointes, on serait capable d’exprimer ce que nous appellerons des « actions répétées sporadique », comme dans l’exemple ci-haut, « simmer 2–3 hours, stirring occasionally » : « ‘simmer’ {*contains*} ‘stir’ », où ‘stir’ est un intervalle non convexe. Il pourrait en revanche être plus simple pour le raisonnement de représenter une action répétée sporadiquement sur un intervalle convexe spécialement qualifié : « ‘simmer’ {*di*} ‘stir (sporadiquement)’ ».

Balbani & Osmani (1999) ont pour leur part proposé une extension permettant d’exprimer des contraintes sur des intervalles cycliques, qui sont représentés sur un cercle plutôt que sur une droite. Ces *c*_intervalles seraient utiles pour la représentation des actions répétées : par exemple, on pourrait représenter « add milk and flour in alternance a little at a time » par « ‘add milk’ {*mmi*} ‘add flour’ », où « *mmi* » représente la conjonction de « *m* » et « *mi* ». En revanche, les *c*_intervalles ne permettent pas de représenter les relations de précedence temporelle.

4.1.2 Contraintes métriques

Dechter *et al.* (1991) ont proposé un formalisme, l'algèbre de distance, permettant de représenter et de raisonner sur des contraintes métriques exprimées sur des points, qui sont des variables réelles. Une contrainte binaire dans cette algèbre est une disjonction d'expressions sous la forme « $\min \leq X_j - X_i \leq \max$ ». On pourrait donc représenter la durée d'une action comme « simmer 2–3 hours » par « $2 \leq I_{\text{simmer}}^+ - I_{\text{simmer}}^- \leq 3$ ». Si toutes les contraintes sur les intervalles ne peuvent pas être exprimées (par exemple, la disjonction d'intervalles $\{b, bi\}$), il semble que les contraintes utiles à la représentation des recettes de cuisine soient toutes exprimables, bien que pas toujours de façon très intuitive. Cependant, il existe un formalisme simple proposé par Kautz & Ladkin (1991) qui permet de représenter à la fois des contraintes métriques et des contraintes qualitatives sur des intervalles.

4.2 Représentations proposées par la linguistique informatique

Plusieurs formalismes temporels ont été proposés par des linguistes informaticiens spécifiquement pour représenter des informations temporelles extraites de textes, ou annoter des textes — c'est-à-dire ajouter une couche d'information formelle. Le langage d'annotation le plus expressif à notre connaissance est TimeML (Pustejovsky *et al.*, 2004). Celui-ci permet d'exprimer des relations entre des actions, des états, des intervalles, des points dans le temps et même des intervalles non convexes exprimés sous la forme d'« ensembles de repères temporels ». En termes de relations pouvant être encodées, TimeML est sans doute comparable aux algèbres combinant contraintes qualitatives et quantitatives, mais il n'est pas associé à une sémantique ni à des règles d'inférences, bien que certaines applications existantes proposent de telles règles, notamment le système de question-réponse de Harabagiu & Bejan (2005).

Voici par exemple les deux premières phrases de la recette de la figure 1 annotées en TimeML. Les balises EVENT soulignent ici les verbes, qui correspondent à des intervalles, alors que la balise SIGNAL souligne le mot exprimant la relation temporelle existant entre les événements correspondant à chacun des verbes, qui est formalisée par la balise TLINK — la relation IS_INCLUDED correspond exactement à la relation $\{di\}$ en algèbre des intervalles.

```
<EVENT eid="e1" class="OCCURENCE"> Brown </EVENT>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="INFINITIVE"
aspect="NONE" pos="VERB"/> hamburger and sausage with onion, garlic,
and all seasonings. <SIGNAL sid="s1"> Meanwhile </SIGNAL>, <EVENT
eid="e2" class="OCCURENCE"> prepare </EVENT> <MAKEINSTANCE
eiid="ei2" eventID="e2" tense="INFINITIVE" aspect="NONE"
pos="VERB"/> the pasta per pkg instructions. <TLINK
eventInstanceID="ei2" signalID="s1" relatedToEvent="ei1"
relType="IS_INCLUDED"/>
```

Les langages d'annotation permettent d'associer des contraintes directement à des mots. Pour raisonner sur des textes, et en particulier pour l'adaptation, cela constitue un avantage considérable : ainsi, toute modification dans le réseau de contrainte peut plus facilement être reportée sur le texte.

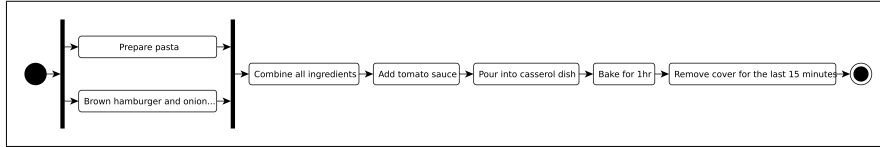


FIGURE 2 – La recette de la figure 1 formalisée sous forme de flux opérationnel. Les actions entre les deux bandes noires sont conjointes, définissant un ordre partiel.

4.3 Flux opérationnels (*workflows*)

Un flux opérationnel est une modélisation d'un ensemble d'actions qui doivent être accomplies, éventuellement par différents acteurs, pour réaliser une tâche. Si les possibilités expressives temporelles des flux sont assez restreintes, étant essentiellement limitées à spécifier l'ordre des actions ou éventuellement la possibilité de les accomplir en parallèle, ils présentent par ailleurs une plus grande richesse en terme de structures de contrôle. En particulier, il est facile d'exprimer des disjonctions et des boucles, qui sont les phénomènes qui nous posent le plus de problème pour la représentation algébrique temporelle des recettes de cuisine. Les flux opérationnels sont d'ailleurs déjà employés par Minor *et al.* (2010) pour l'adaptation de recettes. Cependant, tout comme les formalismes d'annotation de texte, les flux ne sont pas à proprement parler associés à une sémantique et à des processus d'inférence. Il faut noter par ailleurs qu'il existe de nombreux langages de flux proposant différents niveaux d'expressivité, et que les langages les plus expressifs peuvent s'apparenter à des langages algorithmiques et donc poser des problèmes de décidabilité. La figure 2 montre le texte de recette de la figure 1 formalisé sous forme de flux opérationnel.

5 Synthèse

La table 1 reprend les principaux phénomènes à représenter dans les recettes identifiés à la section 2 afin de noter ceux pour lesquels nous croyons avoir identifié une manière de les exprimer dans les différents formalismes présentés dans la section 4. Bien entendu, le fait que nous n'ayons pas identifié une manière de représenter un phénomène donné dans un formalisme ne constitue pas une preuve que cette manière n'existe pas.

6 Conclusion

Notre première constatation est que chaque phénomène que nous souhaitons représenter est exprimable dans au moins un formalisme étudié, mais qu'aucun formalisme ne semble capable d'exprimer tous les phénomènes. La question qui se pose donc est de savoir s'il faut créer un nouveau formalisme ou étendre un formalisme existant, ou encore s'il est possible de combiner certains formalismes que nous avons étudiés. En particulier, Condotta *et al.* (2009) se sont penchés sur le problème de la fusion de connaissances spatio-temporelles à travers l'utilisation de réseaux de contraintes

	Allen	$\mathcal{L}/\mathcal{N}/\mathcal{DU}$	Intervalles non convexes	Intervalles cycliques	Algèbre de distance	TimeML	Flux opérationnels
Durée qualitative	✓	✓	✓	✓	✓	✓	✓ ^a
Durée quantitative précise		✓			✓	✓	
Durée quantitative imprécise		✓			✓		
Ordre total	✓	✓	✓		✓	✓	✓
Ordre partiel	✓	✓	✓		✓	✓	✓
Simultanéité	✓	✓	✓		✓	✓	
Action répétée un nombre indéterminé de fois			✓	✓		✓	✓
Actions répétées en alternance			✓	✓			✓
Action répétée sporadiquement			✓			✓	✓ ^a
Disjonction exclusive							✓

TABLE 1 – Adéquation entre les phénomènes temporels à représenter et les formalismes. Les types de durée sont définis à la section 2.1, les types d'ordre et la simultanéité à la section 2.2, les types de répétition d'action à la section 2.3 et la disjonction à la section 2.4.

a. On peut représenter ces phénomènes en combinant une boucle et une instruction de type « no operation ».

qualitatives issues de différents formalismes. Ces travaux constituent une base intéressante sur laquelle nous comptons nous appuyer pour nos travaux futurs sur l'adaptation spatio-temporelle mais, plus spécifiquement, ils créent également un cadre permettant de combiner différentes algèbres qualitatives. À long terme, il serait intéressant d'étudier certains des formalismes que nous avons laissés de côté, par exemple ceux liés à la représentation des actions et des changements et ceux liés à la planification — en particulier, il existe une littérature assez étendue sur la planification à partir de cas, qui pourrait constituer une approche complémentaire à celle que nous tentons de développer. Par exemple, le calcul des situations ou un formalisme apparenté pourrait être employé pour vérifier que la cohérence d'une recette est maintenue après son adaptation.

Nous avons concentré notre attention sur les recettes de cuisine, mais notre étude aurait pu s'appliquer à plusieurs autres types de textes, voire de cas non textuels. En effet, plusieurs ensembles de cas de type procédural, que ce soient des instructions d'assemblage, de préparation pharmaceutique ou de réparation de locomotive, pourraient bénéficier d'une procédure d'adaptation automatique lorsque l'une des pré-conditions requises à l'exécution de la procédure n'est pas respectée. Les mêmes types de représentation et de raisonnement pourraient alors être employés. Par ailleurs, l'approche que nous souhaitons développer sera applicable également au raisonnement à partir de cas spaciaux ou spatio-temporels — par exemple pour l'adaptation d'itinéraires ou l'aide à la prise de décision sur des projets d'aménagement agricole —, soulevant des problèmes de représentation similaires.

Références

- ALLEN J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, **26**(11), 832–843.
- BADRA F., BENDAOU R., BENTEBITEL R., CHAMPIN P., COJAN J., CORDIER A., DESPRÉS S., JEAN-DAUBIAS S., LIEBER J., MEILENDER T., MILLE A., NAUER E., NAPOLI A. & TOUSSAINT Y. (2008). TAAABLE : Text Mining, Ontology Engineering, and Hierarchical Classification for Textual Case-Based Cooking. In *ECCBR Workshops, Workshop of the First Computer Cooking Contest*, p. 219–228, Heidelberg : Springer.
- BALBIANI P. & OSMANI A. (1999). Représentation et raisonnement sur les intervalles cycliques. In M.-C. ROUSSET, Ed., *Journées nationales sur les modèles de raisonnement (JNMR 1999)*.
- BLANSCHÉ A., COJAN J., DUFOUR-LUSSIER V., LIEBER J., MOLLI P., NAUER E., SKAF-MOLLI H. & TOUSSAINT Y. (2010). TAAABLE 3 : adaptation of ingredient quantities and of textual preparations. In C. MARLING, Ed., *Workshop Proceedings of the Eighteenth International Conference on Case-Based Reasoning (ICCBR 2010)*, p. 189–198, Alessandria, Italy.
- COJAN J. & LIEBER J. (2008). Conservative adaptation in metric spaces. In *Advances in Case-Based Reasoning (ICCBR 2008)*, p. 135–149 : Springer.
- CONDOTTA J.-F., KACI S., MARQUIS P. & SCHWIND N. (2009). Merging qualitative constraints networks defined on different qualitative formalisms. In *Conference on Spatial Information Theory (COSIT'09)*.
- DECHTER R., MEIRI I. & PEARL J. (1991). Temporal constraint networks. *Artificial Intelligence*, **49**, 61–95.
- HARABAGIU S. & BEJAN C. A. (2005). Question answering based on temporal inference. In *Proceedings of the AAI-2005 Workshop on Inference for Textual Question Answering*.
- KAUTZ H. & LADKIN P. (1991). Integrating metric and qualitative temporal reasoning. In *Proceedings of the 9th National Conference of the American Association for Artificial Intelligence (AAAI)*, p. 241–246, Anaheim, CA : AAAI Press/MIT Press.
- LADKIN P. (1986). Time representation : A taxonomy of interval relations. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI)*, p. 354–359.
- LE BER F., LIEBER J. & NAPOLI A. (2009). Utilisation d'une algèbre temporelle pour la représentation et l'adaptation de recettes de cuisine. In *Actes du 17e atelier de raisonnement à partir de cas*, Paris.
- LE BER F., LIEBER J. & NAPOLI A. (2010). Représentation temporelle qualitative de recettes de cuisine. In *Cinquième atelier « Représentation et raisonnement sur le temps et l'espace » (RTE 2010)*, p. 25–28, Caen, France.
- LIGOZAT G. (1991). On generalized interval calculi. In *Proceedings of the 9th National Conference of the American Association for Artificial Intelligence (AAAI)*, p. 234–240, Anaheim, CA : AAAI Press/MIT Press.
- MINOR M., BERGMANN R., GÖRG S. & WALTER K. (2010). Adaptation of cooking instructions following the workflow paradigm. In C. MARLING, Ed., *Workshop Proceedings of the Eighteenth International Conference on Case-Based Reasoning (ICCBR 2010)*, p. 199–208, Alessandria, Italy.

- PUJARI A. K., KUMARI G. V. & SATTAR A. (1999). INDU : An interval & duration network. *Advanced Topics in Artificial Intelligence*, p. 291–303.
- PUSTEJOVSKY J., INGRIA R., SAURÍ R., CASTAÑO J., LITTMAN J., GAIZAUSKAS R., SETZER A., KATZ G. & MANI I. (2004). The specification language TimeML. In I. MANI, J. PUSTEJOVSKY & R. GAIZAUSKAS, Eds., *The Language of Time : A Reader*, p. 545–557. Oxford : Oxford University Press.

Exploration de données temporelles floues : application en archéologie

Cyril de Runz, Eric Desjardin

CRéSTIC,
IUT de Reims Châlons Charleville,
Université de Reims Champagne-Ardenne,
Rue des Crayères, BP1035, 51687 Reims Cedex 2
cyril.de-runz@univ-reims.fr et eric.desjardin@univ-reims.fr

Résumé : Dans cet article, nous proposons une approche d'analyse temporelle de l'information basée sur l'exploration des périodes d'activités d'un ensemble d'objets archéologiques. En raison de l'imprécision de leurs limites, elles sont représentées par des nombres flous. Notre démarche exploite un indice d'antériorité entre deux nombres flous, afin de déterminer un indice temporel de chaque objet dans l'échantillon. Cet indice est alors utilisé pour extraire les positions temporelles de chaque objet ainsi que des représentants et des regroupements. L'information émergente met en lumière les rapports temporels entre objets. Nous avons appliqué cette démarche aux données portant sur les tronçons de rues de l'époque Romaine.

Mots-clés : Relation d'antériorité, rangement, groupement, représentant, intervalle flou, archéologie

1 Introduction

Dans un Système d'Information Géographique (SIG) dédié à l'archéologie, l'analyse exploratoire de données cherche à dégager des relations et des corrélations afin d'extraire de nouvelles connaissances entre les objets d'une base de données spatio-temporelles. Dans cet article, nous explorons la base de données afin de déterminer le positionnement temporel de chaque objet. Cette exploration est riche d'information pour l'archéologue.

Le comportement temporel des objets dans leur ensemble fournit de nouvelles connaissances sur l'information portée par ces objets. Ces nouvelles connaissances facilitent la compréhension des données archéologiques et donc l'expertise. Evaluer ce comportement lorsque les périodes étudiées sont imprécisément définies nécessite de mettre en oeuvre des méthodes non classiques. Pour ce faire, nous utilisons une représentation temporelle par des nombres flous et se pose alors le problème complexe de leur comparaison.

En effet, les comparaisons de deux nombres flous sont souvent non transitives (voir (Wang *et al.*, 1995)); celles-ci ne sont pas directement utilisées pour le classement de nombres flous dans un ensemble. Cependant, une alternative consiste à positionner chaque nombre flou individuellement dans l'ensemble. Celle-ci a conduit à la définition de nombreuses techniques permettant de positionner un nombre flou relativement à un ensemble de nombres flous et non vis-à-vis de chaque nombre.

Parmi ces méthodes, on trouve l'approche proposée par (Kerre, 1982). Selon lui, après avoir déterminé le maximum de l'ensemble selon le principe d'extension défini dans (Zadeh, 1965), le positionnement d'un élément correspond à sa distance au maximum des éléments. Une autre approche est celle de Jain qui considère les nombres selon leur intersection à un ensemble flou appelé « maximisant » construit sur l'union des supports des nombres flous à positionner (Jain, 1977). Dans la démarche proposée dans cet article, nous proposons d'étudier/explore le comportement temporel de chaque objet de la base de données via sa capacité à être antérieur ou/et postérieur à chaque objet de la base.

Pour cela, nous utilisons l'indice d'antériorité proposé dans (de Runz *et al.*, 2010). A l'aide de cet indice, nous définissons, au niveau de l'ensemble des objets, l'indice temporel de chaque objet par comparaison entre d'une part son aptitude à être antérieur aux autres et d'autre part son aptitude à avoir des objets qui lui sont antérieurs.

Comme l'exploration de données nécessite généralement une métrique, et que définir une métrique sur des informations floues est complexe, nous proposons d'exploiter l'indice temporel en entrée de cette étude. A partir de cet indice, nous pouvons définir le positionnement temporel (le rang) de chaque objet dans un ensemble, puis étudier de possibles groupes dans les objets; ce que nous ferons en exploitant une démarche proposée dans (Blanchard *et al.*, 2010) pour des données multidimensionnelles, étendue dans Blanchard *et al.* (2011) aux données relationnelles.

Dans le cadre du projet ArchéoChamp (Desjardin & de Runz, 2009), nous proposons de construire une telle approche sur les données portant sur les tronçons de rues romaines découverts lors de fouilles dont les périodes d'activité ont préalablement été modélisées par des nombres flous (afin de représenter les données et leur imprécision). Ces données sont stockées dans la base *BDFRues*.

Ainsi, après avoir présenté dans la section 2 les données étudiées, nous définirons l'indice temporel de chaque objet dans leur ensemble (section 3). Nous exposerons ensuite des premières structures qui apparaissent à l'aide de l'indice (section 4). Nous appliquerons ensuite une approche permettant l'extraction de représentants et de groupements temporels (section 5). Enfin, nous présenterons une conclusion.

2 Modélisation des données

Au-delà de l'élaboration de la cartographie archéologique de la ville de Reims et de ses environs à l'époque romaine, le projet ArchéoChamp soutenu par la

région Champagne Ardenne et cadre applicatif de ce travail, porte sur la mise en place d'un Système d'Information Géographique (SIG) pluridisciplinaire.

La base de données spatiotemporelles *BDFRues* est, dans ce cadre, dédiée aux éléments de rues romaines à Reims, *BDFRues*. Dans cette base de données, les tronçons de rues sont caractérisés notamment par une période d'activité.

La datation de la période d'activité des objets est généralement issue d'interprétations ou d'estimations dépendantes de l'environnement de la découverte (lieux de fouilles, stratigraphie, comparaison aux objets se situant dans la même pièce...). De plus, la codification linguistique de périodes temporelles n'a pas toujours la même représentation. Par exemple l'estimation du début du Bas Empire varie selon les experts entre 193 et 284 après J.C. Elle est donc largement imprécise.

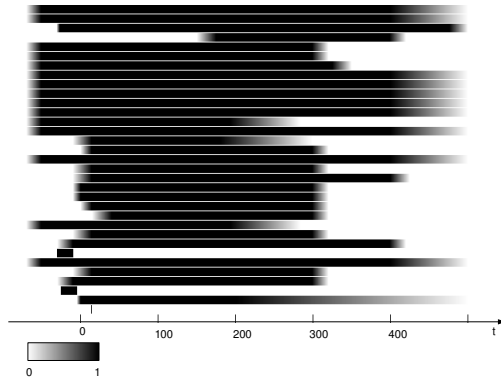


FIGURE 1 – Périodes floues d'activité des objets de *BDFRues* (Chaque nombre flou -chaque période- est représenté par une "bande". Le niveau de gris correspond au degré d'appartenance et l'abscisse au temps).

Nous représentons les périodes d'activité par des ensembles flous convexes et normalisés (généralement des intervalles flous). On peut ainsi prendre en compte cette imprécision. Une représentation visuelle de ces ensembles est proposée dans la figure 1. Les intervalles flous sont définis à l'aide d'intervalles de début et de fin donnés par les experts.

Dans cet article, chaque objet archéologique A_i aura pour composante temporelle un nombre flou $A_i.FDate$ ayant pour fonction d'appartenance $A_i.fdate$. L'ensemble de tous les objets sera appelé Ω .

3 Indice temporel

L'analyse exploratoire proposée se base sur la construction d'un indice temporel à partir de l'indice d'antériorité *Ant* défini dans (de Runz *et al.*, 2010).

Ce dernier quantifie une relation binaire, définie sur les nombres flous et interprétée comme antériorité, entre deux nombres flous $A_i.FDate$ et $A_j.FDate$ quelconques. Cet indice est construit de la manière suivante dans le cadre de deux objets archéologiques A_i et A_j :

$$Ant(A_i, A_j) = \begin{cases} \frac{K_\omega(A_i.FDate)}{K_\omega(A_i.FDate) + K_\omega(A_j.FDate)} & \text{si } K_\omega(A_i.FDate) + K_\omega(A_j.FDate) > 0, \\ 1 & \text{si } K_\omega(A_i.FDate) + K_\omega(A_j.FDate) = 0; \end{cases}$$

avec $\omega = \{A_i.FDate, A_j.FDate\}$. K correspond dans cette formule à l'indice de Kerre (Kerre, 1982) défini comme suit :

$$K_\omega(A_i.FDate) = D_H(A_i.FDate, \widetilde{max}(\omega))$$

avec

$$D_H(A_i.FDate, A_j.FDate) = \int |A_i.fdate(x) - A_j.fdate(x)| dx.$$

et \widetilde{max} est le maximum selon le principe d'extension de Zadeh.

L'indice temporel d'un objet A_i doit permettre de définir sa position temporelle. Son calcul doit donc prendre en considération à la fois son aptitude à être antérieur ainsi que celle à ce qu'on lui soit antérieur. C'est pourquoi, l'indice temporel $IndTemp_\Omega(A_i)$ de A_i est déterminé comme suit :

$$IndTemp_\Omega(A_i) = \sum_{A_j \in \Omega} Ant(A_j, A_i) - Ant(A_i, A_j).$$

Les valeurs de cet indice appartiennent à l'intervalle $[-card(\Omega), card(\Omega)]$.

4 Positions temporelles et premiers regroupements

4.1 Classes temporelles

D'après la définition de notre indice temporel, si l'appétence d'un objet à ce que l'on lui soit antérieur est supérieure (resp. inférieure) à son appétence à être antérieur aux autres, son indice temporel sera positif (resp. négatif). On peut donc considérer qu'un objet ayant un indice temporel négatif peut être qualifié, par rapport à l'ensemble des objets, comme un objet « plutôt antérieur » tandis qu'un objet ayant un indice temporel positif correspond à un objet « plutôt postérieur ». Nous proposons ainsi une classification temporelle simple des objets. La figure 2 présente les objets de *BDFRues* selon ces deux classes.

Pour les données de *BDFRues*, la classe des objets « plutôt postérieurs » regroupe deux fois plus d'objets que celle des objets « plutôt antérieurs ». On peut

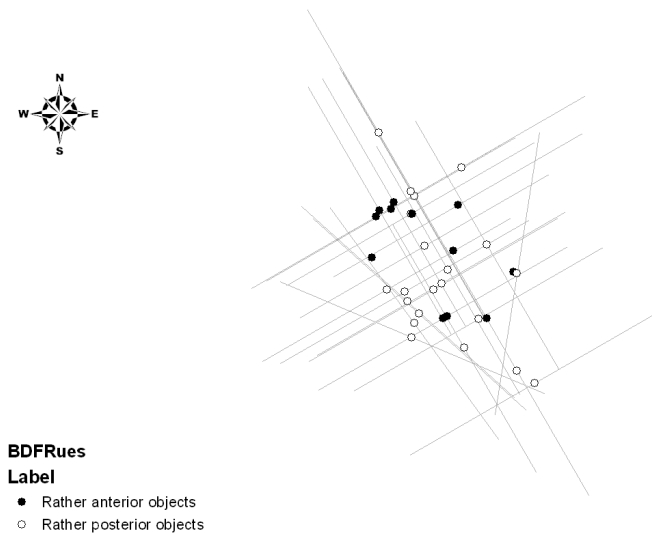


FIGURE 2 – Carte des objets de *BDFRues* classés en « plutôt postérieurs », « plutôt antérieurs »

supposer que les objets « plutôt antérieurs » ont des valeurs de l'indice d'antériorité avec les autres objets proches de 1. Cela expliquerait le fort différentiel entre le cardinal de la classe des objets « plutôt postérieurs » avec celui de la classe des objets « plutôt antérieurs ». Cette hypothèse est d'ailleurs vérifiée puisque, sachant qu'il y a trente trois objets :

- la valeur de l'indice temporel de l'élément le plus antérieur est de -29.8 ;
- la plus grande valeur d'indice temporel des éléments « plutôt antérieurs » est -5.4 ;
- plus de 72% des objets « plutôt postérieurs » ont une valeur d'indice temporel supérieure à 6.

La bipolarisation des valeurs est donc pertinente. Cependant, on peut remarquer que trois objets ont une valeur de l'indice temporel supérieur à 0 mais proche de 0. Ces trois éléments ont une période d'activité singulière débutant au début de la période d'étude mais finissant à un peu plus de sa moitié. Ils ont donc un comportement temporel très particulier en étant presque tout autant antérieur que postérieur.

4.2 Positions temporelles

À l'aide de l'indice temporel, nous pouvons ranger simplement les données. Si la valeur de l'indice d'un objet A_i est plus grande que celui de A_j alors son rang sera plus grand. Nous appelons ce rang la position temporelle. Pour l'objet A_i , il sera noté $PosTemp_{\Omega}(A_i)$. Le principe de rangement est le suivant :

Si $IndTemp_{\Omega}(A_i) > IndTemp_{\Omega}(A_j)$, alors $PosTemp_{\Omega}(A_i) > PosTemp_{\Omega}(A_j)$.

La position temporelle de A_i correspond donc au rang de A_i dans la liste des objets de Ω ordonnée selon les valeurs de l'indice temporel.

Bien que les positions temporelles obtenues par notre approche peuvent différer des rangs obtenus par les méthodes classiques (Kerre, Jain, etc.), les positions temporelles ont une interprétabilité plus forte que les rangs issus de Kerre ou de Jain. En effet, la position temporelle d'un objet archéologique est obtenue à l'aide de l'indice temporel de l'objet. Ainsi, elle reflète donc la relation temporelle de l'objet à l'ensemble des autres. La figure 3 fournit les positions temporelles des objets de BDFRues. On peut noter que les positions temporelles fournissent un pré-ordre total d'un ensemble de nombres flous.

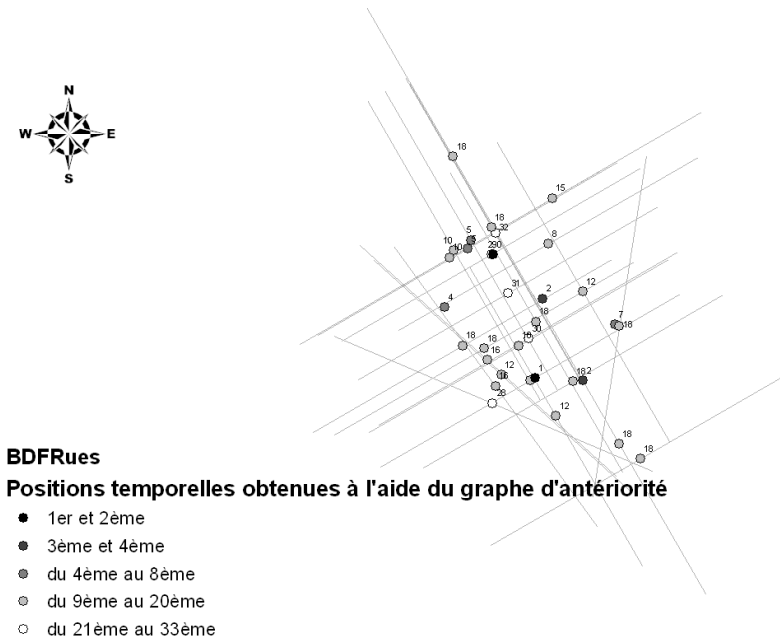


FIGURE 3 – Carte des positions temporelles des 33 objets de *BDFRues* selon leurs périodes d'activité

On peut de plus extraire trois objets particuliers : l'objet le plus antérieur, l'objet le plus postérieur et l'objet temporellement médian. L'objet le plus ancien dans le cadre applicatif, c'est-à-dire le plus antérieur, noté PA , est celui dont la valeur de l'indice temporel est la plus petite dans l'ensemble des valeurs de l'indice temporel des objets de Ω . La position temporelle de l'objet le plus antérieur est la position temporelle minimale des objets de Ω :

$$PosTemp(PA) = \min_{A_i \in \Omega} (PosTemp(A_i))$$

L'objet le plus récent dans le cadre applicatif, c'est-à-dire le plus postérieur, noté PP , est celui dont la valeur de l'indice temporel est la plus grande dans l'ensemble des valeurs de l'indice temporel des objets de Ω . La position temporelle de l'objet le plus antérieur est la position temporelle maximale des objets de Ω :

$$PosTemp(PP) = \max_{A_i \in \Omega} (PosTemp(A_i))$$

Grâce à l'approche par rang (position temporelle), il est trivial de définir l'objet temporellement médian. Il est celui dont la valeur de l'indice temporel est médiane à l'ensemble des valeurs de l'indice temporel pour les objets de Ω . Le médian étant en statistique un estimateur robuste de position dans un échantillon, l'objet temporellement médian est donc un estimateur de position de l'ensemble des objets étudiés. Ces objets particuliers sont identifiables dans la figure 4.

Les classes et les positions extraites à l'aide de l'indice temporel ne fournissent pas directement des informations sur de possibles groupes de données proches temporellement ni sur les représentants de ces groupes.

5 Représentants et groupes temporels

Basée sur la théorie du vote de Borda, l'approche exploratoire proposée dans (Blanchard *et al.*, 2010, 2011) a pour objectif de faire émerger des représentants et des groupes de l'ensemble étudié. Pour cela, elle se base sur la construction d'un graphe de représentativité à l'aide de l'établissement du degré de représentativité de chaque élément dans l'ensemble.

5.1 Degré et graphe de représentativité

Le calcul du degré de représentativité se décompose en trois étapes :

- Chaque objet de Ω évalue tous les objets de Ω afin de définir des relations valuées entre les objets, deux à deux.
- A partir de ces relations, chaque objet range les autres par ordre de préférence. Ces rangs sont ensuite transformés en scores.
- Le degré de représentativité d'un objet A_i résulte de l'agrégation des scores de rangs obtenus par A_i pour l'ensemble des objets de Ω .

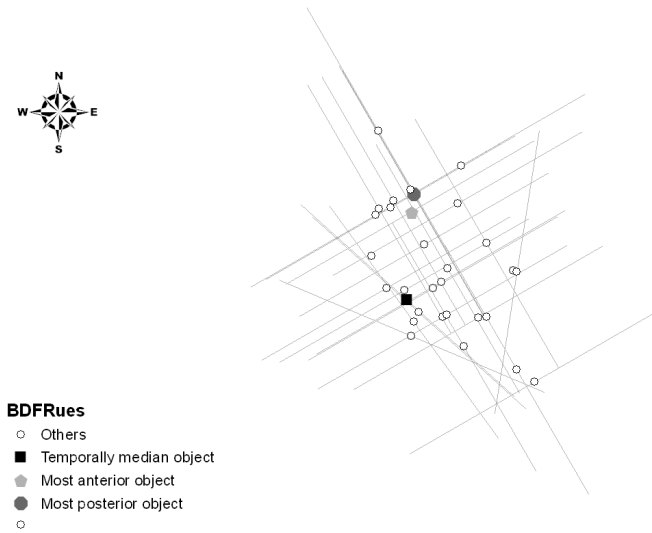


FIGURE 4 – Carte des objets de *BDFRues* surlignant l’objet le plus antérieur, le plus postérieur et le temporellement médian.

Le rang de A_i selon A_j noté $RG_{\Omega}^{A_j}(A_i)$ est, dans le cadre de notre travail, défini à l’aide des valeurs de l’indice temporel. Considérons trois objets A_i , A_j , A_k et la distance euclidienne D entre les valeurs de l’indice temporel. Si

$$D(IndTemp_{\Omega}(A_i), IndTemp_{\Omega}(A_j)) < D(IndTemp_{\Omega}(A_k), IndTemp_{\Omega}(A_j))$$

alors

$$RG_{\Omega}^{A_j}(A_i) < RG_{\Omega}^{A_j}(A_k).$$

$RG_{\Omega}^{A_j}$ prend ces valeurs dans l’intervalle $[1, card(\Omega)]$ et $RG_{\Omega}^{A_j}(A_j) = 1$.

Chaque objet classe les autres objets selon ce principe. Le degré de représentativité d’un objet A_i dans Ω est défini comme suit :

$$DR_{\Omega}(A_i) = \sum_{A_j \in \Omega} (card(\Omega) - RG_{\Omega}^{A_j}(A_i) + 1)$$

On le calcule pour l’ensemble des objets Ω .

5.2 Représentants et groupements

La deuxième partie de l’approche consiste à faire émerger une structuration des objets étudiés. Un graphe des représentants, appelé GR_{Ω} défini par $\langle X, U \rangle$

(X est l'ensemble des sommets, U l'ensemble des arcs), est alors construit à l'aide d'un voisinage au sens des k plus proches voisins (selon les proximités des valeurs des indices temporels). À chaque objet A_i de Ω est associé un sommet i et ce dernier est relié au sommet représentant l'objet dans son voisinage dont le degré de représentativité est le plus élevé, c'est à dire :

$$i \in X \Leftrightarrow A_i \in \Omega$$

$$\forall i, j \in X, (i, j) \in u \Leftrightarrow A_j = \underset{A_s \in V_k(A_i)}{\operatorname{argmax}} (DR_\Omega(A_s))$$

Dans ce graphe, chaque composante connexe forme un groupement de données et un arbre. Le nombre de groupements (de composantes connexes) est dépendant du voisinage choisi. Les éléments à la racine des arbres qui forment les composantes connexes sont les meilleurs représentants des groupes. Ils ont aussi le degré de représentativité le plus élevé dans le groupe.

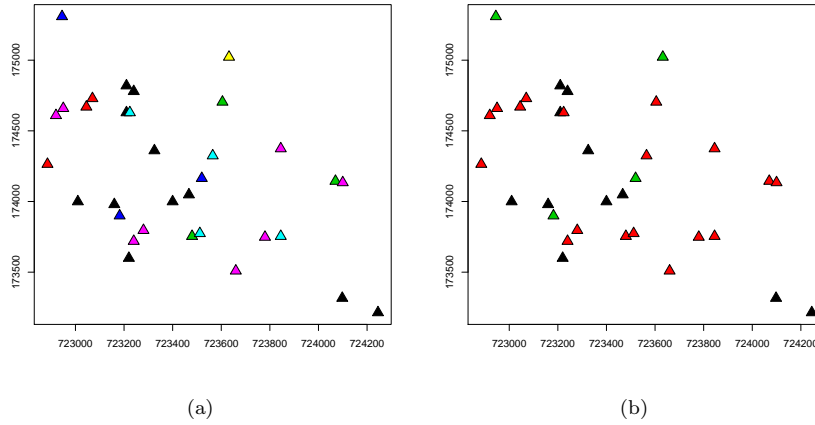


FIGURE 5 – Objets de *BDFRues* regroupés selon leur proximité d'indice temporel par le graphe de représentativité : (a) avec un *3-ppv* et (b) avec un *4-ppv*.

En utilisant cette approche sur les données de *BDFRues* nous obtenons, à l'aide d'un *3-ppv* et d'un *4-ppv*, les cartes présentées dans la figure 5. Nous passons de 8 groupements trouvés dans la figure 5a à 3 dans la figure 5b. Ces structures et leur évolution donnent une information complémentaire aux informations précédemment extraites. Ainsi, par exemple, dans l'approche *3-ppv* les groupements de données singulières peuvent être étudiés à l'instar des données d'indice temporel proche de 0, tandis que dans la recherche à l'aide du *4-ppv*, nous travaillerons plus sur les grandes tendances (*e.g.* objets présents sur toute la période Gallo-Romaine) et sur le ralliement des objets singuliers à ces dernières.

6 Conclusion

Dans cet article, nous avons proposé un processus exploratoire ayant pour but l'analyse des relations entre objets selon la temporalité de l'information archéologique disponible. Pour cela, nous avons construit un indice temporel de chaque objet dans l'ensemble des objets. A partir de cet indice, nous avons étudié des structures émergentes construites à partir d'une interprétation de l'indice, d'un rangement des objets, et d'une méthode de fouille de données.

Dans leurs expertises, les archéologues évaluent les objets qu'ils sont susceptibles de trouver dans un site tant d'un point de vue fonctionnel que temporel. Ils ont notamment besoin d'étudier les relations temporelles entre les objets stockés dans les bases de données afin de (i) regarder si la logique temporelle est respectée, et (ii) dégager une évolution temporelle de la cité. Dans cet objectif, ils peuvent utiliser les approches proposées dans cet article.

Références

- BLANCHARD F., DE RUNZ C., HERBIN M. & AKDAG H. (2011). Représentativité et graphe de représentants : une approche inspirée de la théorie du choix social pour la fouille de données relationnelles. In *Atelier Fouille de Données Complexes, Conférence Extraction et Gestion des Connaissances 2011*, p. 73–83, Brest, France.
- BLANCHARD F., VAUTROT P., AKDAG H. & HERBIN M. (2010). Data representativeness based on fuzzy set theory. *Journal of Uncertain Systems*, **4**(3), 216–228.
- DE RUNZ C., DESJARDIN É., PIANTONI F. & HERBIN M. (2010). Anteriority index for managing fuzzy dates in archaeological gis. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, **14**(4), 339–344.
- DESJARDIN É. & DE RUNZ C. (2009). Gissar : de la saisie de fouilles à l'analyse spatiotemporelle en archéologie. In *Spatial Analysis and GEomatics*, Paris, France.
- JAIN R. (1977). A procedure for multiple-aspect decision making using fuzzy set. *Internat. J. Systems Sci.*, **8**, 1–7.
- KERRE E. E. (1982). The use of fuzzy set theory in electrocardiological diagnostics. In M. GUPTA & E. SANCHEZ, Eds., *Approximate Reasoning in Decision-Analysis*, p. 277–282. North-Holland Publishing Company.
- WANG X., KERRE E. E., CAPPELLE B. & RUAN D. (1995). Transitivity of Fuzzy Orderings Based on Pairwise Comparis. *The Journal of Fuzzy Mathematics*, **3**(2), 455–463.
- ZADEH L. A. (1965). Fuzzy Sets. *Information Control*, **8**, 338–353.

Localisation d’habitant dans un espace perceptif par réseau dynamique [★]

Pedro Chahuara, Michel Vacher, François Portet

Laboratoire d’Informatique de Grenoble (UMR 5217 UJF/INPG/CNRS)
équipe GETALP, 385 avenue de la Bibliothèque
38041 Grenoble - France
{Pedro.Chahuara, Michel.Vacher, Francois.Portet}@imag.fr

Résumé : Cet article présente une approche de fusion de données temporelles pour localiser une personne dans un environnement perceptif à partir de sources non visuelles. Ces sources informent sur la localisation de manière incertaine et sont donc combinées en utilisant un réseau dynamique à deux niveaux d’hypothèses de localisation et en adaptant une méthode de propagation d’activation pour prendre en compte la validité éphémère et l’ambiguïté des sources. Les résultats sur des enregistrements réels montrent l’intérêt de l’approche.

Mots-clés : réseaux dynamiques, propagation d’activation, fusion de données temporelles, Intelligence Artificielle, informatique ubiquitaire.

1 Introduction

L’objectif de ce travail est de définir une méthode de fusion de données pour la localisation de l’habitant à l’intérieur de son domicile en utilisant des sources d’informations non visuelles (c.-à-d. sans caméra), événementielles et indirectes (c.-à-d. pas de capteur porté par la personne). Il se déroule dans le cadre du projet SWEET-HOME qui vise à concevoir un système de contrôle intelligent de la domotique à travers une interface de commande utilisant la voix pour un plus grand confort et une plus grande sécurité de la personne. La localisation de la personne est primordiale pour interpréter les commandes vocales (p.ex., si la personne prononce « allume la lumière », il est nécessaire de localiser la personne pour déduire la lampe à activer). Les approches présentées dans la littérature reposent souvent sur l’utilisation de la vidéo (Friedland *et al.*, 2010). Cependant, pour limiter les coûts et pour respecter la vie privée, seuls les capteurs domotiques sont considérés dans le projet. Ces sources (ici capteurs de présence, capteurs de contact, et microphones) sont donc ambiguës (p.ex., observation d’un même évènement par plusieurs capteurs éloignés), ont une dimension temporelle (p.ex., dans le couloir à 15 : 23 : 45) et dépendent de l’espace physique (p.ex., nombre de pièces, réverbération).

Parmi les capteurs considérés, le microphone est un capteur d’évènements et de localisation prometteur qui, par sa nature omnidirectionnelle ou unidirectionnelle, peut être

★. Ce travail a été réalisé dans le cadre du projet Sweet-Home qui bénéficie d’une aide de l’Agence Nationale de la Recherche portant la référence ANR-09-VERS-011

très sensible ou très spécifique. Cependant, très peu d'approches du domaine utilisent uniquement l'information sonore. On peut citer Bian *et al.* (2005) qui ont développé une grille de microphones dans une pièce pour déterminer l'origine d'un son dans un espace 3D avec une erreur moyenne de moins de 27 cm. Comme souligné par les auteurs, l'information audio nécessite beaucoup moins de bande passante que l'information vidéo et permet de détecter facilement certaines activités (conversations, sonneries de téléphone). Néanmoins, cette solution est trop complexe et trop coûteuse à mettre en place dans un appartement classique. De plus, si la vidéo est sensible aux changements de luminosité, le canal audio est sensible au bruit environnemental (Vacher *et al.*, 2011). Le canal audio, bien qu'étant une modalité pertinente et peu coûteuse, est donc une source bruitée et ambiguë. Il devient donc essentiel de mettre en place un système de localisation multisource qui tirerait partie des redondances et des complémentarités des sources choisies. Plusieurs travaux ont été dans ce sens, en utilisant principalement des méthodes de classification probabilistes telles que les réseaux bayésiens (Dalal *et al.*, 2005) ou les modèles de Markov (Kröse *et al.*, 2008). Cependant, l'information temporelle dans ces modèles est assez pauvre et rarement prise en considération. Duong *et al.* (2009) prennent en compte l'aspect temporel à travers d'un modèle semi-Markov caché hiérarchique pour la reconnaissance d'activités à partir de séquences d'événements en modélisant aussi le temps passé dans chaque état (représentant une sous-activité). Cependant, dans notre cas, il est difficile de prédire la localisation présente en évaluant une séquence de localisations passées car le nombre de chemins de déplacements possibles augmenterait l'incertitude. De plus, ces approches nécessitent un grand nombre de données étiquetées pour l'apprentissage. Certaines applications mettant en jeu des traitements sur la connaissance, telles que la recherche d'information (Crestani, 1997; Aswath *et al.*, 2005), la reconnaissance de forme (Niessen *et al.*, 2008) ou la navigation d'agents animés (Kortenkamp & Chown, 1993), emploient une approche de propagation d'activation à travers un réseau sémantique. Ces réseaux sont constitués de nœuds représentant des concepts et de liens représentant l'intensité de la relation entre concepts. L'activation d'un nœud se dilue de proche en proche en activant les nœuds voisins jusqu'à ce qu'elle s'épuise. Des travaux récents (Niessen *et al.*, 2008) ont montré l'intérêt de la propagation d'activation comme mémoire à court terme pour désambiguïser des événements sonores.

L'approche, que nous avons adoptée pour localiser une personne, se base sur un réseau dynamique où les événements des capteurs activent des hypothèses de localisation qui se périment au cours du temps. Ces hypothèses prennent en compte les activations précédentes, l'incertitude des événements et la durée de validité d'un événement. Dans notre approche, le réseau est dynamique dans le sens où il évolue en fonction des données mais aussi où il prend en compte des connaissances a priori sur l'environnement dans lequel évolue l'habitant. Cette approche est détaillée dans la section 2. La méthode proposée est ensuite évaluée en section 3 sur des données réelles. L'article se termine par une brève discussion.

2 Localisation d'habitant par propagation d'activations multisource

La méthode mise en œuvre pour la localisation d'une personne à partir de plusieurs sources d'information non-visuelles repose sur la modélisation des liens entre les événements et les hypothèses de localisation par un réseau dynamique à deux niveaux.

Chaque nouvelle information est propagée dans le réseau pour mettre à jour les hypothèses de localisation. Dans la plupart des approches de propagation d'activation, le réseau est bien établi, mais, dans notre cas, les informations ont une validité temporaire et les liens entre les nœuds évoluent en fonction du temps. La section 2.1 présente comment le réseau évolue au fil du temps en intégrant les données de plusieurs sources temporelles et comment cette information se « périmé » grâce à une fonction d'oubli. La section 2.2 détaille comment des connaissances a priori sur l'environnement peuvent être prise en compte avec l'information immédiate des capteurs pour calculer les relations entre les différents nœuds du réseau.

2.1 Réseaux dynamiques et propagation d'activation

La méthode que nous utilisons est basée sur celle de Niessen *et al.* (2008) qui ont présenté une approche basée sur les réseaux dynamiques permettant de désambiguïser la reconnaissance d'évènements sonores. Il s'agit d'un réseau à deux niveaux. Le niveau zéro est constitué des évènements sonores, le niveau 1 représente les hypothèses liées à un évènement (p.ex. : rebond de ballon ou claquement de mains), et le niveau deux représente le contexte de l'évènement (p.ex, match de basket, concert). Chaque évènement active des hypothèses selon l'évènement et les contextes auxquels les hypothèses sont liées. L'activation des hypothèses se propage ensuite aux contextes. Nous avons adapté cette approche pour localiser une personne à partir de différentes sources d'évènements. Le réseau dynamique, que nous avons conçu, est organisé en deux niveaux : le premier niveau correspond aux *hypothèses de localisation* générées à partir d'un évènement, et le deuxième niveau représente les *contextes d'occupation* de chaque pièce ; le poids d'activation des différents contextes indique la localisation la plus probable sachant les évènements précédents. La méthode utilise les définitions suivantes :

Définition 1 (observation)

Une observation o_{t_n} est une donnée structurée générée par un capteur ayant réagit à un évènement e_{t_n} au temps t_n avec $t, n \in N$. À chaque observation o est associée le type de capteur $o.type$, l'origine du capteur $o.captteur$ ainsi que le rapport signal-à-bruit (RSB) $o.rsb$ dans le cas de signaux numériques.

Définition 2 (observations simultanées)

Deux observations $o_{t_n}^i$ et $o_{t_k}^j$ sont dites simultanées si $t_n = t_k$ et $i \neq j$.

Définition 3 (hypothèse de localisation)

$h_{t_n}^i \in L$, où $L = \{Loc_1, \dots, Loc_R\}$ est le nœud hypothèse que l'habitant se trouve à la i^e localisation à l'instant t_n . Les hypothèses sont générées uniquement à partir des observations au temps t_n .

Définition 4 (contexte d'occupation)

$c^i \in P$ où $P = \{Piece_1, \dots, Piece_R\}$ est le nœud de contexte d'occupation de la i^e pièce. La valeur d'activation du contexte varie en fonction du temps et des hypothèses.

Définition 5 (poids de relation)

$w \in [0, 1]$ est la force de la relation entre deux nœuds du réseau. Ainsi, w_{o, h^i} est le poids du lien entre une observation o et la i^e hypothèse et w_{h^i, c^j} est le poids du lien entre la i^e hypothèse et le j^e contexte.

Définition 6 (fonction d'oubli)

La fonction d'oubli $f(t_n, t_{n-1}) = e^{-\frac{\Delta_t}{\tau}}$, avec $\Delta_t = t_n - t_{n-1}$, représente la décroissance d'un contexte au cours du temps. Elle permet de conserver une mémoire à court terme des contextes. Plus les observations seront espacées au cours du temps (p.ex., $\Delta_t > 3 \times \tau$), plus les contextes seront oubliés.

2.1.1 Évolution du réseau dynamique

L'algorithme fonctionne ainsi :

1. pour toute nouvelle observation $o_{t_n}^k$ un nouveau nœud est créé ;
2. des nœuds hypothèses $h_{t_n}^i$ sont ensuite créés et connectés aux $o_{t_n}^k$ avec les poids w_{o^k, h^i} ;
3. les hypothèses $h_{t_n}^i$ sont connectées aux contextes de localisation c^j avec les poids w_{h^i, c^j} ;
4. les activations sont propagées des $o_{t_n}^k$ aux $h_{t_n}^i$ et l'activation des $h_{t_n}^i$ est calculée ;
5. les activations sont propagées de $h_{t_n}^i$ à c^j et l'activation de c^j est recalculée ;
6. le c^j avec la plus forte activation devient la localisation courante ;
7. tous les nœud $h_{t_n}^i$ et $o_{t_n}^k$ sont supprimés du réseau.

Un exemple de réseau dynamique est donnée par la figure 1. À l'instant t_{n-2} , l'évènement $e_{t_{n-2}}$ est capté par un capteur qui génère une observation $o_{t_{n-2}}$ de laquelle 3 hypothèses sont déduites : $h_{t_{n-2}}^1$ avec une relation de 0.1 vers le contexte c^1 , $h_{t_{n-2}}^2$, avec 0.6 vers c^2 et $h_{t_{n-2}}^3$ avec 0.3 vers c^3 . S'il n'y a pas eu d'évènement antérieur, c^2 est la localisation la plus probable. À l'instant t_{n-1} , l'évènement $e_{t_{n-1}}$ est observé par deux capteurs. Les nœuds générés à t_{n-2} disparaissent, seul les contextes sont conservés. Si ces contextes sont encore actifs leur activation sera pondérée par $f(t_{n-1}, t_{n-2})$ auxquels vont s'ajouter les activations générées par les hypothèses $h_{t_{n-1}}^i$ avec une relation de 0.2 vers le contexte c^2 et $h_{t_{n-1}}^2$ avec 0.8 vers c^4 . La méthode s'appliquera ensuite de la même manière à l'instant t_n , etc.

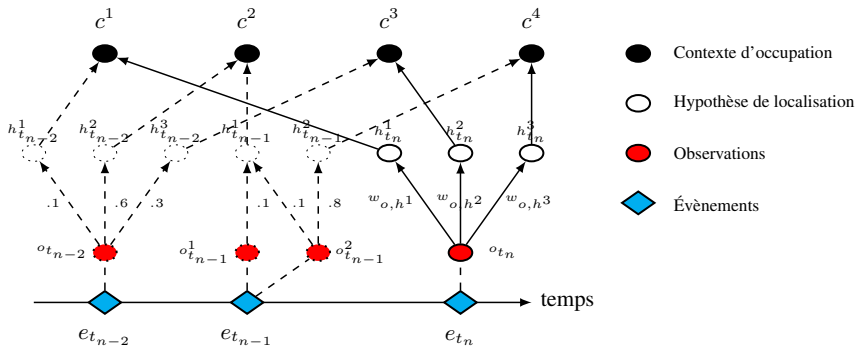


FIGURE 1: Exemple de réseau dynamique

2.1.2 Propagation de l'activation

L'activation recueillie par un nœud est classiquement définie (Crestani, 1997; Niessen *et al.*, 2008) par la formule $n_i(t) = \sum_{i \neq j} w_{i,j} \times A^j(t)$ où $w_{i,j}$ est le poids, j correspond à un voisin de i et $A^j(t)$ est l'activation de ce voisin au temps t . Bien entendu, un nœud qui a été activé par un voisin ne peut communiquer de nouveau cette activation à ce voisin. Dans notre cas, les activations sont toujours initiées par une observation o qui aura à chaque fois une activation maximale $A^o(t) = 1$ quelque soit son type et se propagera de bas en haut. Une fois l'activation $n(t)$ provenant des voisins accumulée, l'activation de sortie du nœud doit être calculée. Celle-ci diffère selon le niveau des nœuds. Pour les hypothèses de localisation, l'activation $A^{h^i}(t) \in [0, 1]$ sera calculée selon la formule 1 :

$$A^{h^i}(t_n) = n_i(t_n) = \sum_{o \in O_{t_n}} w_{o,h^i} A^o(t_n) \quad (1)$$

où O_{t_n} est l'ensemble des observations simultanées à t_n liées à h^i et $\sum_{o \in O_{t_n}} w_{o,h^i} \leq 1$. Pour les contextes d'occupation de pièce, l'activation de sortie résulte de l'accumulation des activations provenant des hypothèses et de l'activation précédente pondérée par un facteur d'oubli. L'équation 2 décrit la loi d'activation du contexte de localisation A^{c^i} suite à une activation externe au temps t_n .

$$A^{c^i}(t_n) = n_i(t_n) \times [M - e^{-\frac{\Delta t}{\tau}} A^{c^i}(t_n - \Delta t)] + e^{-\frac{\Delta t}{\tau}} A^{c^i}(t_n - \Delta t) \quad (2)$$

où $A^{c^i}(t_n - \Delta t)$ est l'activation précédente, $M = 1$ est l'activation maximale et $e^{-\frac{\Delta t}{\tau}}$ est la fonction d'oubli. Ainsi, si aucun événement n'apparaît pendant $5 \cdot \tau$ secondes, l'activation des contextes pourra être considérée comme nulle. L'introduction de M permet de maintenir toutes les activations entre 0 et 1.

2.2 Calcul des relations entre les différentes couches du réseau

Le réseau dynamique étant constitué de deux niveaux bien définis (cf. figure 1), deux types de relations existent, la relation *observation-hypothèse* et la relation *hypothèse-contexte*. Les liens entre les différentes couches dépendent fortement de l'application et de l'environnement considéré, c'est pourquoi il est nécessaire d'introduire tout d'abord l'environnement applicatif pour faciliter la compréhension de l'article.

2.2.1 Évènements considéré dans l'environnement perceptif

Un exemple d'environnement perceptif considéré dans cette étude est l'Habitat Intelligent pour la Santé (HIS) de la Faculté de Médecine de Grenoble (Le Bellego *et al.*, 2006). La figure 2 décrit l'appartement intelligent. Il est composé de 6 pièces contenant un ensemble de capteurs pour percevoir l'environnement : 7 microphones (*Mic*) placés dans le plafond (hauteur d'environ 3 mètres) ; 3 capteurs de contact sur des portes (*CP*) de meubles ; et 6 détecteurs de présence infrarouge (*PID*) placés sur les murs. Parmi les données recueillies par les capteurs (*Mic*, *CP* et *PID*), seuls les signaux sonores doivent subir un pré-traitement. En effet, les données *CP* et *PID* délivrent une information booléenne déjà représentative d'évènements. Les signaux audio ont été traités par le système AUDITHIS (Vacher *et al.*, 2010). Succinctement, les évènements audio

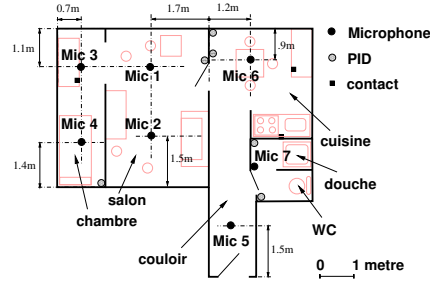


FIGURE 2: Plan de l'appartement et position des capteurs.

sont détectés, en temps réel, et sont ensuite classés comme parole ou comme son de la vie courante (p.ex. chute d'objet). Les microphones étant omnidirectionnels, un évènement sonore est généralement capté par plusieurs microphones en même temps. Pour chaque observation o de notre étude, $o.type \in \{CP, Mic, PID\}$.

2.2.2 Relation hypothèse-contexte

La relation *hypothèse-contexte* est dans notre cas une relation univoque car une hypothèse de localisation n'est liée qu'à une seule pièce. Il s'agit d'un choix expérimental car certaines hypothèses sur des pièces faiblement séparées (p.ex. salon/chambre) pourraient activer plusieurs contextes de pièce. Par ailleurs, l'application visée étant la localisation d'un habitant dans une pièce, tous les nœuds contextes de pièces sont créés à l'initialisation du réseau et en font partie en permanence.

2.2.3 Relation observation-hypothèse

La relation *observation-hypothèse* est unidirectionnelle et de type de un vers plusieurs. Le poids et les hypothèses générés varient en fonction de la source des observations. Pour représenter cette relation dans le réseau, deux types d'information sont considérées : celles qui sont dynamiques et dépendent des événements, et celle qui sont statiques provenant des connaissances *a priori* sur les capteurs.

Certaines connaissances sur l'environnement peuvent être utilisées pour définir les liens qui existent entre concepts. Pour une observation o , avec $o.type \in \{CP, PID\}$, un seul nœud hypothèse de localisation est créé. Les informations spatiales des *PID* et des *CP* sont univoques et certaines. Par exemple, l'ouverture du frigo ne peut se produire que si la personne se trouve dans la pièce. Par conséquent, un évènement est lié avec un poids $w_{o,h} = 1$ à une hypothèse unique de localisation, et, l'hypothèse est liée avec un poids de $w_{h,c} = 1$ au contexte. Il n'en va pas de même pour les informations des microphones. En effet, les microphones peuvent théoriquement capter toutes les ondes acoustiques générées dans les pièces d'une habitation. L'énergie du signal sonore détecté fournit bien une information dynamique sur la proximité de la source mais, il ne donne, par contre, aucune indication sur sa direction. Par exemple, un bruit capté à 1 mètre d'un microphone peut tout à fait avoir été généré dans une pièce différente de celle dans laquelle le microphone est fixé. Pour prendre en compte ce phénomène, nous avons modélisé l'ambiguïté d'un microphone en fonction de sa distance avec les

autres pièces par les poids w_{o,h^i} des relations dans le réseau dynamique. La valeur de ces poids a été calculée en estimant la probabilité $P(loc = i | Mic = j)$ que l'habitant soit à la i^e localisation sachant qu'un évènement a été détecté sur le j^e microphone. Le poids de la relation entre l'hypothèse h^i et l'observation o est donné par la formule 3 :

$$w_{o^k,h^i}(t_n) = \frac{P(loc = i | o_{t_n}^k \text{ .capteur}) \times o_{t_n}^k \text{ .rsb}}{\sum_{l \in L} \sum_{o \in O_{t_n}} P(loc = l | o \text{ .capteur}) \times o \text{ .rsb}} \quad (3)$$

où O_{t_n} est l'ensemble des observations simultanées provenant des microphones à l'instant t_n , $P(loc | Mic)$ est une connaissance *a priori* décrite en section 2.2.4, L est l'ensemble des localisations possibles et $o \text{ .rsb}$ le rapport signal sur bruit linéarisé de o . Ainsi w_{o^k,h^i} est un poids normalisé ($\sum_{i,k} w_{o^k,h^i} = 1$) qui représente le lien entre les observations sonores simultanées et les hypothèses de localisation.

2.2.4 Acquisition des informations statiques

Pour calculer la probabilité $P(loc = i | Mic = j)$, deux approches ont été testées : l'approche naïve et l'approche par analyse statistique sur corpus. Pour l'approche naïve, la référence sonore a été choisie à 1 mètre au dessous du microphone. À partir de cette référence, un cercle de 2 mètres de rayon a été tracé autour de chaque microphone, il correspond à une atténuation de $-6dB$ en considérant l'atténuation quadratique classique. Au delà de ce cercle, la perte d'énergie est supérieure à 75%. Le poids est calculé selon la surface de l'intersection entre le cercle et les pièces, avec une pénalité de 2 lorsqu'une cloison est traversée. L'autre approche est statistique, elle calcule directement les probabilités à partir du corpus annoté. Le tableau 1 indique les poids obtenus pour l'approche naïve et pour l'approche statistique.

Le mélange d'informations *a priori* et d'informations dynamiques permet une meilleure désambiguïsation. Par exemple, dans le cas de deux observations sonores simultanées captées par le micro 6 de la cuisine et le micro 7 de la salle de bain avec un RSB identique égal à $12dB$, les formules 1 et 3 et l'*a priori* naïf donnent pour l'activation $A^{h^{cuisine}} = rsb \times (.87 + .5) / [rsb \times (.87 + .5 + .18 + .13 + .03 + .18 + .1)] = 0,69$ qui sera supérieure à celle de la salle de bain $A^{h^{sdb}} = 0,09$ bien que le rapport signal à bruit soit le même.

3 Résultats

L'approche a été testée sur des données réelles acquises lors d'expériences (Fleury *et al.*, 2010) destinées à évaluer la reconnaissance automatique des activités de la vie

Mic	$P(Loc Mic)$ estimé par la méthode naïve							$P(Loc Mic)$ estimé à partir du corpus						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
chambre	.14	.07	.70	.85				.28	.29	.42	.43	.25	.18	.20
salon	.86	.93	.27	.14	.01	.13	.03	.59	.56	.47	.41	.07	.07	.09
cuisine			.03	.02	.10	.87	.50	.05	.08	.06	.09	.45	.63	.37
douche					.06		.18	.06	.05	.04	.04	.09	.04	.10
wc					.06	.18		.01	.01	.01	.01	.12	.05	.21
couloir					.77		.10		.02	.01	.02	.03	.03	.02

TABLE 1: Estimation de $P(Loc|Mic)$

Capteurs	Mic	PID	CP	Mic+CP	PID+Mic	PID+CP	PID+Mic+CP
sans a priori	25.7	88.9	26.5	32.8	87.7	89.4	88.2
a priori naïf	30.2	88.9	26.5	34.1	89.0	89.4	89.5
a priori naïf sujet 4	25.3	92.8	19.5	25.4	96.8	92.8	96.8
a priori statistique	30.9	88.9	26.5	34.8	89.7	89.4	90.1

TABLE 2: Exactitude (%) avec plusieurs combinaisons de sources

quotidienne d'une personne à son domicile dans le but de pouvoir détecter automatiquement une perte d'autonomie. Dans le HIS (cf. section 2.2.1), 15 volontaires ont effectué chacun au moins une fois 7 activités prédéfinies (p.ex., dormir, faire sa toilette, communiquer avec l'extérieur) sur une période d'environ 1 heure, sans consigne sur la manière ou l'ordre dans lequel elles devaient être effectuées. Les données ont ensuite été annotées grâce à des caméras placées uniquement à cet effet. Pour chaque enregistrement, les observations enregistrées par les capteurs *CP*, *PID* et *Mic* ont activé un réseau dynamique. Les traces d'activation de chaque contexte ont été conservées et l'emplacement de départ de l'expérimentation est supposé connu (ici la cuisine). Les performances de localisation ont été évaluées à l'échelle de la seconde. À chaque seconde, le contexte de localisation de plus haut poids est comparé à la vérité terrain. S'il y a correspondance, alors il s'agit d'un vrai positif (VP), sinon il s'agit d'une confusion. Ceci nous permet de générer une table de confusion à partir de laquelle le taux d'exactitude est calculé par $TE = \frac{nb(VP)}{nb(test)}$, où $nb(test)$ correspond à la durée de l'évaluation exprimée en secondes. Le tableau 2 montre les résultats sur les 47226 secondes (13h 7min 6s) d'enregistrement des 15 participants pour chaque combinaison de capteurs. Dans le cas sans a priori (c.-à-d., $P(loc = i | Mic = j) = 1$ si le capteur j est dans i^e pièce 0 sinon), l'ajout des contacts de porte *CP* améliore légèrement la précision obtenue avec seulement les capteurs infrarouges *PID* mais l'information sonore dégrade légèrement les performances (88,2% au lieu de 88,9%). Lorsqu'un a priori sur les relations événement sonore - hypothèse de localisation est ajouté et calculé par la méthode naïve, dans ce cas, comme dans la méthode statistique, nous observons une amélioration des performances. La connaissance a priori du lien entre un son et un capteur permet de corriger un grand nombre d'erreurs de localisation. Le gain reste faible néanmoins, lorsque le changement de pièce n'est pas détecté par les *PID* mais que les paroles sont bien identifiées, l'information sonore est utile pour améliorer la localisation (comme dans le cas du sujet 4) en compensant le défaut de sensibilité des infrarouges. Ceci illustre l'intérêt de la combinaison de plusieurs sources.

Pour le participant 4 et dans le cas a priori naïf, la valeur de l'activation des contextes de localisation pour chacune des occurrences d'événements est tracée sur la figure 3. Dans cet extrait, le participant est dans le séjour (événements 1 à 9), puis il se rend à la cuisine (événements 10 à 15) avant de revenir au séjour. Les événements donnant des informations sur cette transition sont sonores (moins précis que les *PID*). La méthode estime la localisation dans la douche (événements 16 à 18) avant de trouver la vraie localisation. Ensuite, le participant reste dans le salon (événements 16 à 60) pour effectuer un appel téléphonique. Au cours de la période dans le séjour, les deux contextes *séjour* et *chambre* s'affrontent car la conversation téléphonique se tient à proximité de la limite du salon et de la chambre à coucher. Le participant étant resté presque immobile pendant la conversation, il n'y a pas eu d'événement détecté par les *PID* et les microphones ont donc constitué la seule source d'information.

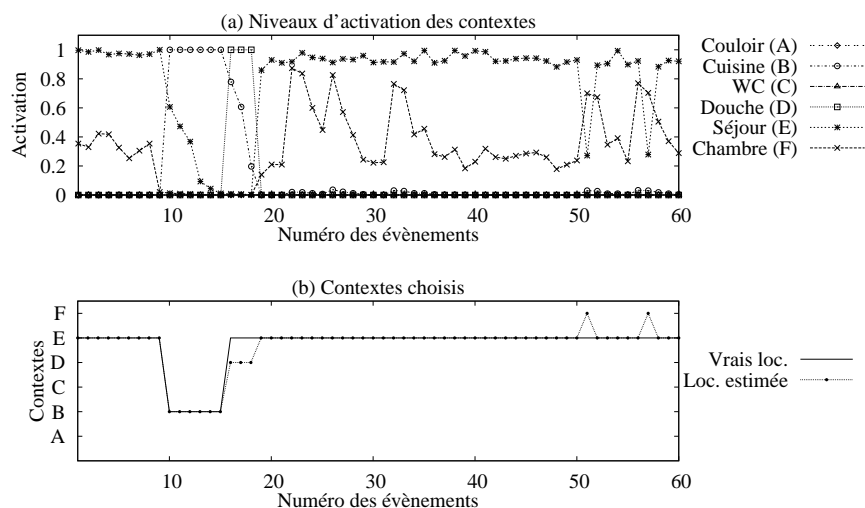


FIGURE 3: Extrait de l'évolution des niveaux d'activation des contextes (a) et contextes choisis pour les 6 localisations possibles (b).

4 Discussion

L'article présente une méthode de fusion d'informations multisources multimodales et temporelles par propagation d'activation dans un réseau dynamique qui permet de modéliser la connaissance statique (liée à l'organisation de l'espace) et dynamique (temporelle). Cette approche permet de fusionner des sources non visuelles avec une grande flexibilité et n'impose pas une perception continue de l'espace ni de patron d'évènements. De plus, cette méthode permet de prendre en compte facilement de l'a priori dans le réseau afin de tenir compte de connaissances expertes.

Les résultats montrent que la fusion d'information par propagation d'activation multi-source présente un intérêt même dans le cas où les sources ont une très bonne précision, c'est le cas des capteurs infrarouges (mais avec une sensibilité imparfaite). L'utilisation de sources de localisation plus incertaines, telles que la détection de parole, peut alors améliorer les performances. Malgré l'ambiguïté inhérente aux microphones omnidirectionnels, la prise en compte de la connaissance a priori du lien entre un son et un microphone permet de corriger un grand nombre d'erreurs de localisation. Cependant, l'utilisation unique des informations de parole conduit cette source à être très peu sensible. Dans notre étude, quasiment 60% des confusions sont dues à des périodes durant lesquelles le sujet ne parle pas, n'utilise pas d'objet et dont les mouvements ne sont pas captés par les PID. Nous prévoyons donc d'inclure la classification de différents types de sons afin de capter les mouvements générateurs de bruits qui ne peuvent être détectés par les PID.

Plusieurs autres pistes d'amélioration de cette méthode vont être suivies. Une première piste consiste à agir sur la constante de temps en fonction de la sémantique des événements et des lieux. Par exemple, une détection dans le couloir devrait avoir une

constante de temps plus faible que dans le cas des toilettes. À cet effet, nous prévoyons d'extraire les constantes de temps potentielles à partir des données. Sur le plus long terme, la méthode présentée peut être améliorée en utilisant la théorie de Dempster-Shafer. Une fusion d'information serait alors possible pour déterminer la localisation par consensus entre le niveau de croyance des sources d'information et les poids d'activation des hypothèses. Cette théorie de combinaison de preuves offre un cadre de travail plus formel que la combinaison de poids. Enfin, nous prévoyons d'appliquer cette méthode à la classification des sons de la vie courante en utilisant le contexte de la localisation pour désambiguïser les hypothèses de classification des sons.

Références

- ASWATH D., D'CUNHA J., AHMED S. T. & DAVULCU H. (2005). Boosting item keyword search with spreading activation. In *WI'05*, p. 704–707.
- BIAN X., ABOWD G. D. & REHG J. M. (2005). Using sound source localization in a home environment. In *Pervasive 2005*, p. 19–36.
- CRESTANI F. (1997). Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review*, **11**(6), 453–482.
- DALAL S., ALWAN M., SEIFRAFI R., KELL S. & BROWN D. (2005). A rule-based approach to the analysis of elders activity data : Detection of health and possible emergency conditions. In *AAAI Fall 2005 Symposium*.
- DUONG T., PHUNG D., BUI H. & VENKATESH S. (2009). Efficient duration and hierarchical modeling for human activity recognition. *Artificial Intelligence*, **173**(7-8), 830–856.
- FLEURY A., VACHER M., PORTET F., CHAHUARA P. & NOURY N. (2010). A multi-modal corpus recorded in a health smart home. In *LREC MMC Workshop*, p. 99–105, Matla.
- FRIEDLAND G., YEO C. & HUNG H. (2010). Dialocalization : Acoustic speaker diarization and visual localization as joint optimization problem. *ACM Transactions on Multimedia Computing, Communications, and Applications*, **6**(4), 27 :1–27 :18.
- KORTENKAMP D. & CHOWN E. (1993). A directional spreading activation network for mobile robot navigation. In *Second International Conference on Simulation of Adaptive Behavior*, p. 218–224.
- KRÖSE B., VAN KASTEREN T., GIBSON C. & VAN DEN DOOL T. (2008). Care : Context awareness in residences for elderly. In *The 6th International Conference of the International Society of Gerontology*, Pisa, Italy.
- LE BELLEGO G., NOURY N., VIRONE G., MOUSSEAU M. & DEMONGEOT J. (2006). A model for the measurement of patient activity in a hospital suite. *IEEE Transactions on Information Technology in Biomedicine*, **10**(1), 92 – 99.
- NIESSEN M. E., VAN MAANEN L. & ANDRINGA T. C. (2008). Disambiguating sounds through context. In *ICSC2008*, p. 88–95 : IEEE Computer Society.
- VACHER M., FLEURY A., PORTET F., SERIGNAT J.-F. & NOURY N. (2010). *Complete Sound and Speech Recognition System for Health Smart Homes : Application to the Recognition of Activities of Daily Living*, p. 645 – 673. Intech Book.
- VACHER M., PORTET F., FLEURY A. & NOURY N. (2011). Development of audio sensing technology for ambient assisted living : Applications and challenges. *International Journal of E-Health and Medical Communications*, **2**(1), 35–54.

Idée préliminaire pour un système de médiation fondé sur une ontologie spatiale

Lamia Belouaer¹

Université de Caen - UFR des Sciences
Département d'informatique
Campus Côte de Nacre
Bd Maréchal Juin
BP 5186, F-14032 Caen Cedex
prenom.nom@info.unicaen.fr

Résumé : L'objectif de cet article est d'utiliser une ontologie spatiale qui permet de faire communiquer des agents, en particulier dans le cas d'interaction homme robot (IHR). L'idée est d'exploiter la connaissance spatiale (les entités spatiales, les relations spatiales et l'information spatiale floue) pour une meilleure collaboration entre ces agents.

Mots-clés : Représentation spatiale, ontologie spatiale, relations spatiales, information floue, interaction homme-robot.

1 Introduction

Le travail collaboratif entre agents dans un environnement nécessite plusieurs outils, notamment des outils de définition et de partage de la connaissance spatiale.

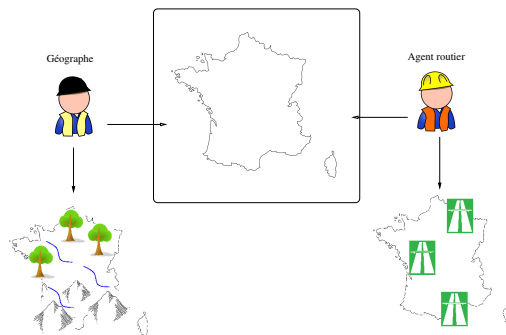


FIGURE 1 – Deux visions spatiales différentes de la carte de la France.

Ce type de travail fait intervenir des acteurs hétérogènes avec une représentation de la connaissance spatiale propre chacun d'entre eux. Dans la figure 1, deux agents n'ont pas la même vision d'un même espace (la France) : (1) l'agent routier a une vision métrique (la distance numérique pour la mesure de routes), (2) le géographe a une vision topologique (une rivière traverse une région). Pour que ces deux acteurs puissent communiquer leurs informations spatiales des fins d'exploitation, la mise en place d'un système de médiation spatiale est nécessaire.

Dans le cadre de la collaboration entre agents hétérogènes dans un environnement commun, nous considérons l'interaction homme-robot. Dans un espace commun, l'interaction homme-robot nécessite un système de médiation pour gérer l'hétérogénéité des données spatiales. Le système de médiation que nous définissons gère les aspects suivants de la dimension spatiale (Belouaer *et al.*, 2010) : (1) l'organisation hiérarchique de l'espace, (2) la représentation numérique, (3) la représentation symbolique qui exprime les relations topologiques et floues.

2 Représentation spatiale : *SpaceOntology*

SpaceOntology (fig. 2) décrit l'espace (entités spatiales, organisation hiérarchiques), les relations spatiales et l'information spatiale floue et imprécise.

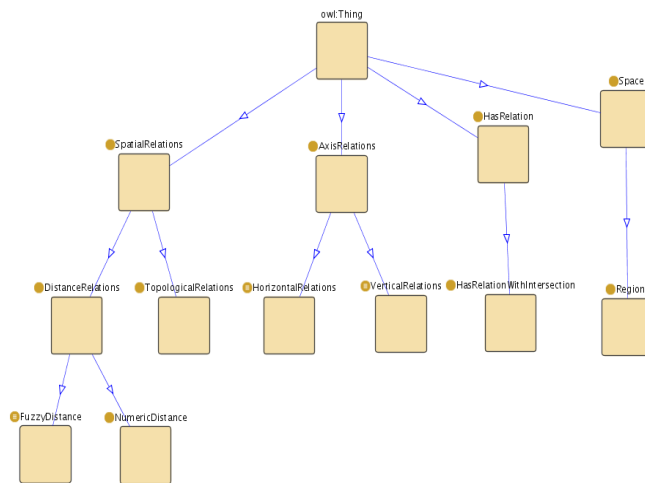


FIGURE 2 – Représentation graphique de *Spaceontology*.

C'est une ontologie (fig. 2) qui permet de :

- décrire les entités spatiales, les relations spatiales et le flou,
- représenter des informations spatiales complexes,
- gérer la sémantique de la connaissance spatiale concernant des entités spatiales ainsi que les liens entre ces entités,
- gérer facilement une grande quantité d'information,

- décrire une structure facilement extensible,
- d'inférer des nouvelles informations spatiales.

2.1 Les concepts *Space* et *Regions*

Le concept **Space** ($\text{Space} \sqsubseteq \mathbb{T}$) représente un environnement global (un pays, une ville, un bâtiment, ...). Le concept **Regions** représente un sous-espace compris dans l'espace global ($\text{Regions} \sqsubseteq \text{Space}$). Une région est elle-même composée en sous-régions. D'un point de vue géométrique, une région est définie par le plus petit rectangle englobant ces extrémités.

Naturellement, un espace est défini d'une manière hiérarchique. De cette organisation résulte une hiérarchie de concepts qui se termine lorsque toutes les informations sont modélisées par un concept unique (ou nous arrivons un niveau souhaité de socialisation). Le plus haut niveau représente l'espace avec le maximum de détails disponibles. Le niveau le plus bas représente l'espace par un seul concept. La relation hiérarchique entre régions est définie par les liens (fig. 3) :

- *consistsOf*, exprime que l'espace se compose d'une ou plusieurs régions
($\text{Space} \sqsubseteq \mathbb{T} \sqcap \exists \text{consistsOf}.\text{Regions} \sqcap \geq 1 \text{consistsOf}$),
- *isPartOf* exprime qu'une région peut appartenir un ou plusieurs espaces
($\text{Regions} \sqsubseteq \text{Space} \sqcap \exists \text{isPartOf}.\text{Space} \sqcap \geq 1 \text{isPartOf}$).

```
<Regions rdf:ID="E1_B6">
<consistsOf>
  <Regions rdf:ID="C11_B6">.....</Regions>
</consistsOf>
<isPartOf rdf:resource="#B6"/>
</Regions>
```

FIGURE 3 – Un extrait de *SpaceOntology* décrivant un étage d'immeuble.

La figure 3 illustre la définition hiérarchique entre deux régions : le premier étage *E1_B6* se compose de plusieurs régions (par exemple un couloir *C11_B6*). L'organisation hiérarchique a un impact sur l'évaluation de la distance. En effet, la distance de 10 mètres dans une ville est considérée comme *proche*, mais dans un couloir, cette distance est considérée comme *loin*. Ainsi, nous associons chaque niveau hiérarchique deux paramètres α et β . Ces deux paramètres sont nécessaires pour l'évaluation de la distance (voir la section 3.1.1).

Une région (*rg*) est associée un rectangle (*rect*) englobant ces extrémités. Notons par X_i la projection de $rect_i$ sur (xx') et par Y_i la projection de $rect_i$ sur (yy') (fig. 4).

SpaceOntology englobe un ensemble d'attributs. Parmi eux, des attributs qui définissent les concepts **Space** et **Regions** (fig. 5) : (1) l'attribut **level** indique le niveau hiérarchique, (2) l'attribut **scale** indique l'échelle de l'espace, (3) l'attribut **length** la longueur du rectangle, (4) l'attribut **width** la largeur du rectangle, (5) l'attribut **origine** l'origine du rectangle, (6) et les attributs **alpha** et **beta**.

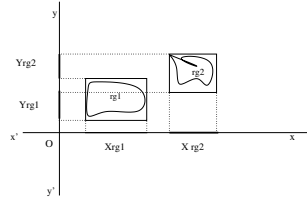


FIGURE 4 – Régions, rectangles et projections.

```
<Regions rdf:ID="E1_B6">
  <origine rdf:xsd:string>E1_B6</origine><level rdf:xsd:int">2</level>
  <scale rdf:xsd:int">3</scale><width rdf:xsd:int">2</width>
  <length rdf:xsd:int">4</length><alpha rdf:xsd:int">3</alpha>
  <beta rdf:xsd:int">5</beta>
</Regions>
```

FIGURE 5 – Un extrait de *SpaceOntology* décrivant un étage.

2.2 Le concept *SpatialRelations*

La sémantique d'*un banc en face de la machine café* n'est pas la même selon que le point de référence est la machine café elle-même ou un observateur extérieur. Afin d'avoir un sens unique, chaque relation spatiale est liée explicitement à un point de référence donné. Nous considérons une représentation en 2D de l'espace donné par (O, \vec{i}, \vec{j}) .

Une relation spatiale est un concept **SpatialRelations** ($\text{SpatialRelations} \sqsubseteq \text{T}$). Ce concept représente l'ensemble de toutes les relations spatiales entre deux régions. Il englobe **TopologicalRelations** ($\text{TopologicalRelations} \sqsubseteq \text{SpatialRelations}$) et **DistanceRelations** ($\text{DistanceRelations} \sqsubseteq \text{SpatialRelations}$).

2.2.1 Relations topologiques : Le concept **TopologicalRelations**

Nous utilisons l'algèbre **ABLR** (**A**bove **B**elow **L**eft **R**ight) (Laborie *et al.*, 2006). Une relation topologique est une relation **ABLR** : un couple $\langle r_X, r_Y \rangle$, o :

- $r_X \in \{L, O_L, C_x, I_x, O_R, R\}$ ($\sqsubseteq \text{HorizontalRelations}$),
- $r_Y \in \{A, O_A, C_y, I_y, O_B, B\}$ ($\sqsubseteq \text{VerticalRelations}$).

Relation horizontale	Relation verticale
X_1 Left $X_2 : L$	Y_1 Above $Y_2 : A$
X_1 Right $X_2 : R$	Y_1 Below $Y_2 : B$
X_1 Overlaps-Left $X_2 : O_L$	Y_1 Overlaps-Above $Y_2 : O_A$
X_1 Inside $X_2 : I_X$	Y_1 Inside $Y_2 : I_Y$
X_1 Contains $X_2 : C_X$	Y_1 Contains $Y_2 : C_Y$
X_1 Overlaps-Right $X_2 : O_R$	Y_1 Overlaps-Below $Y_2 : O_B$

TABLE 1 – La représentation spatiale ABLR (Laborie *et al.*, 2006).

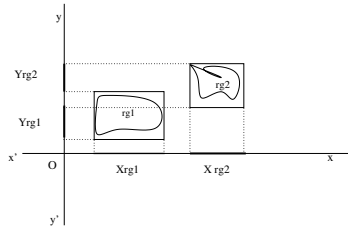


FIGURE 6 – La relation spatiale entre rg_2 et $rg_1 : rg_2 \langle R, O_A \rangle rg_1$.

En combinant les relations définies sur l'axe horizontal et celles définies sur l'axe vertical, cette algèbre (table 1) offre 36 relations au lieu des 169 relations.

2.2.2 Distance Relations : le concept DistanceRelations

Dans la suite, une région (ϵ) est identifiée par l'origine (ϵ) du rectangle qui lui est associé. Les distances peuvent être numériques (distance euclidienne) ou floues. Dans *SpaceOntology*, le concept **DistanceRelations** englobe les concepts :

- **NumericDistance** ($\text{NumericDistance} \sqsubseteq \text{DistanceRelations}$),
- **FuzzyDistance** ($\text{FuzzyDistance} \sqsubseteq \text{DistanceRelations}$).

Le concept **NumericDistance** décrit la distance euclidienne entre deux régions. Notons $d(\epsilon, r)$ la distance euclidienne entre l'origine des deux rectangles représentant deux régions (r région de référence et ϵ région cible). Le concept **FuzzyDistance** décrit la distance d'une manière floue en exploitant les variables linguistiques suivante : proche, assez proche, assez loin et loin .

2.3 Le concept HasRelation

Le concept **HasRelation** lie des régions cibles et une région de référence par une relation spatiale. Ce concept est utile pour décrire les configurations spatiales.

```
HasRelation  $\sqsubseteq$  T  $\sqcap$   $\exists$ 
concernsSpatialRelation.SpatialRelations  $\sqcap$  = 1
concernsSpatialRelation  $\sqcap$   $\exists$  hasReferent.Regions  $\sqcap$   $\geq$  1
hasReferent  $\sqcap$   $\exists$  hasTarget.Regions  $\sqcap$  = 1 hasTarget
```

Le concept **HasRelationWithIntersection** est un sous-concept de **HasRelation**. Dans le cas d'une intersection entre deux régions. Ce concept définit les points de contact.

```
HasRelationWithContact  $\sqsubseteq$  HasRelation  $\sqcap$  = 1
concernsSpatialRelation  $\sqcap$   $\exists$  intersectionresults.Regions  $\sqcap$   $\geq$  1
intersectionresults
```

2.4 Illustration

Nous illustrons l'aide d'un exemple simple mais concret, la modélisation de la connaissance spatiale l'aide de *SpaceOntology*

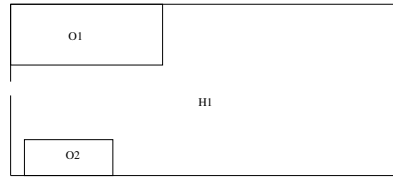


FIGURE 7 – Un plan partiel de l’environnement. Un étage dans un immeuble. Cet étage englobe deux bureaux (o1, o2) et un hall (h1).

La figure 8 modélise le l’environnement considéré dans la figure 7 en utilisant *SpaceOntology*. Nous considérons chaque entités spatiales comme une instantiation du concept **Regions**. Chaque relation spatiale entre les différentes entités est décrite en combinant les concepts **SpatialRelations** (dans la figure 8 il s’agit de la relation topologique et **HasRelation** (dans la figure 8 il s’agit de l’instance *relation1*).

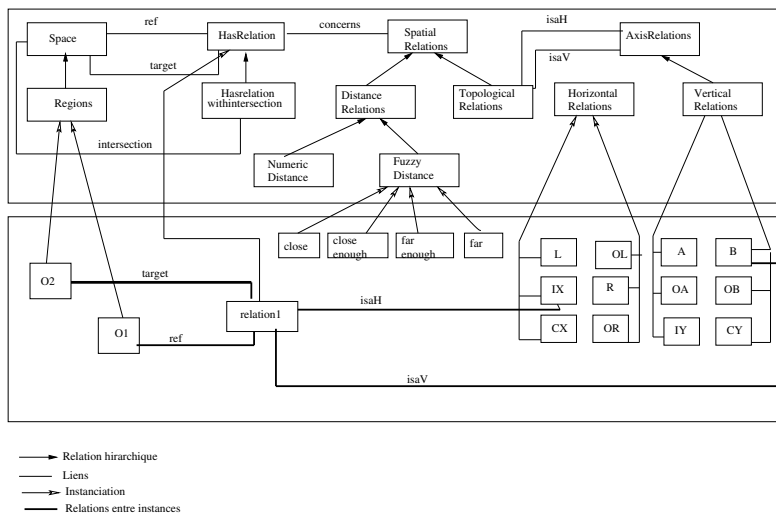


FIGURE 8 – Illustration de la modélisation avec *Spaceontology*.

3 SpaceOntology, un outil de médiation

Nous exploitons *SpaceOntology* pour la médiation entre un humain et un robot. Cette ontologie décrit un environnement donén. De plus, elle offre un ensemble de règles permettant de faire un raisonnement spatial (Belouaer *et al.*, 2010).

Nous considérons un agent humain (H) et un agent robot (R) et une description de l'espace (définie par *SpaceOntology*) dans lequel ils évoluent. L'humain et le robot s'changent la connaissance spatiale. Nous proposons de formaliser les changes entre un agent humain et un agent robot sous forme d'*expressions spatiales*. L'expression spatiale mise ou reçue par un agent humain (H_{SE}) est un triplet $H_{SE} = \langle rg_1, \varrho, Rg_c \rangle$ (table 2). L'expression spatiale mise ou reçue par un agent robot (R_{SE}) est un triplet $R_{SE} = \langle rg_1, \rho, Rg_c \rangle$ (table 2).

Expression	Attributs
$H_{SE} = \langle rg_1, \varrho, Rg_c \rangle$	rg_1 est la région de référence Rg_c est la région courante $\varrho = \begin{cases} \langle rg_2, r \rangle & \text{si } rg_2 \text{ est précise} \\ r & \text{sinon} \end{cases}$ $r \in \mathbf{SpatialRelations.Instance}$
$R_{SE} = \langle rg_1, \rho, Rg_c \rangle$	rg_1 est la région de référence Rg_c est la région courante $\rho = \begin{cases} \langle rg_2, r \rangle & \text{si } rg_2 \text{ est précise} \\ r & \text{si } rg_2 \text{ est imprécise} \end{cases}$ $r \in \mathbf{NumericDistance.Instance} \cup \{P, \theta\}$ $P = (x, y)$ tel que $(x, y) \in \mathbb{R}^+ * \mathbb{R}^+$ θ un angle, $\theta \in [0, 360]$

TABLE 2 – Définition des expressions spatiales possibles pour l'humain et le robot.

Notre objectif est de traduire une expression spatiale d'un agent humain vers une expression spatiale d'un agent robot et inversement. Formellement, nous définissons deux fonctions $T_{H \rightarrow R}$ et $T_{R \rightarrow H}$ utilisant *SpaceOntology* tel que $T_{H \rightarrow R}(H_{SE}) = R_{SE}$ et $T_{R \rightarrow H}(R_{SE}) = H_{SE}$. Dans le papier (Belouaer *et al.*, 2010), nous avons détaillé $T_{H \rightarrow R}(H_{SE}) = R_{SE}$. Ici, nous nous intéressons $T_{R \rightarrow H}(R_{SE}) = H_{SE}$.

3.1 Du robot l'humain : $T_{R \rightarrow H}(R_{SE})$

Nous exploitons *SpaceOntology* pour la médiation entre l'humain et le robot. Formellement, la fonction $T_{R \rightarrow H}$ est définie comme suit :

$$T_{R \rightarrow H} : \begin{array}{ccc} S_1 & \rightarrow & S_2 \\ R_{SE} : (rg_1, \rho, Rg_c) & \mapsto & H_{SE} : \langle rg_1, \varrho, Rg_c \rangle \end{array}$$

Avec,

- $S_1 = \mathbf{NumericDistance.Instance} \cup \{P, \theta\} \times \mathbf{Regions.Instance}$
- $S_2 = \mathbf{SpatialRelations.Instance} \times \mathbf{Regions.Instance}$

$$T_{R \rightarrow H}(R_{SE}) = T_{R \rightarrow H}(\langle rg_1, \rho, Rg_c \rangle) \quad (1)$$

$$= [T_{R \rightarrow H}^{Rg}(rg_1), T_{R \rightarrow H}^{Rl}(\rho), T_{R \rightarrow H}^{Rg}(Rg_c)] \quad (2)$$

La fonction $T_{R \rightarrow H}$ prend en entre une expression spatiale du robot (R_{SE}) et la traduit en expression spatiale pour l'humain. Cette transformation fait intervenir deux fonctions : $T_{R \rightarrow H}^{Rg}$ et $T_{R \rightarrow H}^{Rl}$. La fonction $T_{R \rightarrow H}^{Rg}$ est une identité.

$$\begin{array}{ccc} T_{R \rightarrow H}^{Rg} : Regions.Instance & \rightarrow & Regions.Instance \\ rg & \mapsto & rg \end{array}$$

Ainsi, la transformation d'une expression spatiale est réduite la traduction de une relation spatiale. Dans la suite, nous nous intéressons la transformation en utilisant la fonction $T_{R \rightarrow H}^{Rl}(\rho)$, o $\rho \in NumericDistance.Instance \cup P \cup \theta$.

$$\begin{array}{ccc} T_{R \rightarrow H}^{Rl} : NumericDistance.Instance \cup P \cup \theta & \rightarrow & SpatialRelations.Instance \\ \rho & \mapsto & \varrho \end{array}$$

$$T_{R \rightarrow H}(R_{SE}) \equiv T_{R \rightarrow H}^{Rl}(\rho) = \begin{cases} T_{R \rightarrow H}^{Rl}(\langle rg_2, r \rangle) & \text{si } rg_2 \text{ est précise} \\ T_{R \rightarrow H}^{Rl}(r) & \text{si } rg_2 \text{ est imprécise} \end{cases}$$

3.1.1 Transformation d'une relation de distance numérique

SpaceOntology, permet la traduction d'une relation de distance numérique en une relation de distance floue (proche des expressions utilisées par l'humain). Pour se faire, nous utilisons la fonction $N_{\alpha, \beta}$ et $F_{\alpha, \beta}$ (Schockaert, 2008).

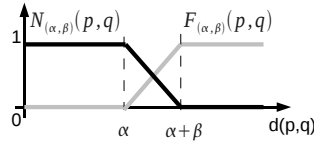


FIGURE 9 – Évaluation des distances. (1) si $d \in [0, \alpha]$, alors d est proche, (2) si $d \in]\alpha, \alpha + \frac{\beta}{2}]$ alors d est assez proche (3) si $d \in]\alpha + \frac{\beta}{2}, \alpha + \beta]$ alors d est assez loin, (4) si $d \in]\alpha + \beta, +\infty[$ alors d est loin.

La fonction $N_{(\alpha, \beta)}(p, q)$ (resp. $F_{(\alpha, \beta)}(p, q)$) mesure le degré avec lequel p est proche de q (resp. le degré avec lequel p est loin de q) ($\alpha, \beta > 0$). La figure 9 illustre la traduction d'une distance euclidienne en une distance symbolique. Les paramètres α et β permettent d'évaluer la distance entre deux spatiales entités. Ces deux paramètres dépendent de la région courante (fig. 10).

la fonction $T_{R \rightarrow H}^{Rl}$ exploite des règles définies dans *SpaceOntology*. Ces règles s'appuient sur le système d'évaluation défini dans la figure 9. La règle *Def_close_relation*

$$\begin{aligned}
 & \text{translate}(?link, ?t_link) \wedge \text{translate}(?relation, ?t_relation) \wedge \text{concerns}(?link, ?relation) \\
 & \wedge \text{hasTarget}(?link, ?tgt) \wedge \text{hasReferent}(?link, ?ref) \wedge \text{isANumericDistance}(?relation) \\
 & \quad \wedge \text{isPartOf}(?ref, ?region) \wedge \text{isPartOf}(?tgt, ?region) \wedge \\
 & \quad \text{swrlb : greaterThan}(\text{distanceValue}(?relation), 0) \wedge \\
 & \quad \text{swrlb : lessThanOrEqual}(\text{distanceValue}(?relation), \text{alpha}(?region)) \\
 & \quad \quad \quad \longrightarrow \\
 & \text{concerns}(?t_link, ?t_relation) \wedge \text{hasTarget}(?t_link, ?tgt) \wedge \text{hasTarget}(?t_link, ?tgt) \wedge \\
 & \text{hasReferent}(?t_link, ?ref) \wedge \text{isAFuzzyDistance}(?t_relation) \wedge \text{is}(?t_relation, "close")
 \end{aligned}$$

FIGURE 10 – Définition de la règle *Def_close_relation*.

(fig. 10) traduit une distance numérique (**NumericDistance**) (*distanceValue(?relation)*) en une valuation symbolique (**FuzzyDistance**) donnée par l'instance *close* pour proche. Cette règle s'applique toute distance (*distanceValue(?relation)*) entre 0 et α .

3.1.2 La relation est une coordonne ou un angle

Une position $P = (x, y)$ est la coordonne indiquant une position d'une région donnée, elle correspond l'origine du plus petit rectangle la définissant. On considère également que l'origine de la région de référence est l'origine du repère. En se basant sur l'ontologie, nous pouvons définir les rectangles correspondant chaque région. Ensuite, nous effectuons des projections de ces rectangles. Les projections nous permettent d'exploiter les relations topologiques définies dans **ABLR** (fig. 11).

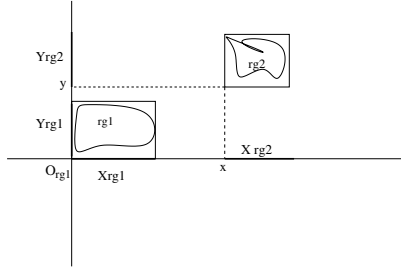


FIGURE 11 – Traduction de (x, y) .

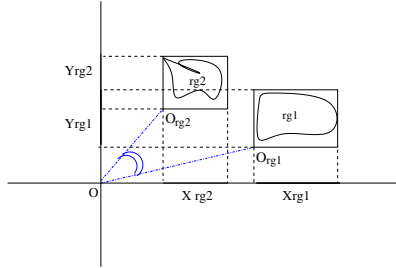


FIGURE 12 – Traduction d'un angle.

Une relation d'orientation est un angle (par exemple l'angle formé par la droite (OO_{rg1}) et la droite (OO_{rg2}))(fig. 11). L'origine correspond un point de référence qui correspond la position de l'observateur (ici le robot) si ce point n'est pas définie. Dans d'autres cas, le point O peut être définie. Cependant, la méthode de traduction reste la même. En effet, nous considérons l'origine de chaque rectangle présentant une région. Nous appliquons la projection des rectangles en se basant sur les points d'origine. Ce qui nous ramène au même cas que la traduction d'une coordonne (fig. 11).

4 Conclusion

Nous avons présenté un système de médiation basé de *SpaceOntology* pour résoudre les problèmes d'hétérogénéités des données spatiales entre agents, en particulier dans le cas d'interaction homme-robot. Cette ontologie permet de modéliser l'espace selon différents aspects, savoir l'organisation hiérarchique, la représentation numérique/symbolique et floue. Ce papier est une première idée sur la traduction d'une expression fournie par le robot vers l'humain. Nos travaux futurs concerneront l'application de cette fonction et de voir comment on peut l'améliorer.

Références

- BELOUAER L., BOUZID M. & MOUADDIB A. (2010). Ontology based spatial planning for human-robot interaction. In *2010 17th International Symposium on Temporal Representation and Reasoning*.
- LABORIE S., EUZENAT J. & LAYAIDA N. (2006). A spatial algebra for multimedia document adaptation. *SMAT*.
- SCHOCKAERT S. (2008). *Reasoning about fuzzy temporal and spatial information from the web*. PhD thesis, Ghent University.

Méthode de modélisation et d'extraction des informations géographiques contenues dans des textes

Van Tien NGUYEN, Mauro GAIO

Laboratoire LIUPPA, BP-1155, 64013 PAU Université Cedex
vantien.nguyen@univ-pau.fr, mauro.gαιο@univ-pau.fr

Résumé : Nous proposons une méthode permettant de modéliser et d'extraire des informations géographiques à partir de texte. Cette méthode est implémentée dans une chaîne de traitement informatique et testée sur un corpus relatif à des récits de voyage ; le but applicatif est d'enrichir une ontologie spécifique à un sous-domaine géographique, la topologie. Après une étape d'analyse morpho-syntaxique classique, un repérage et étiquetage des groupes de mots cibles (groupes nominaux, verbes de déplacement) est élaboré. Le coeur de la méthode consiste en la production d'une représentation symbolique d'un groupe de mot particulier appelé : Toponyme, celle-ci est implémentée grâce à des transducteurs spécifiquement conçus. Ce sont ces informations extraites à l'aide de cette démarche qui sont exploitées ici à des fins d'enrichissement de l'ontologie construite dans le projet ANR-GEONTO.

Mots-clés : information géographique, extraction des concepts topographiques, raisonnement spatial, modélisation spatiale

1 Introduction

L'enjeu scientifique des ontologies géographiques a été mis en avant dans plusieurs travaux tels que Uitermark (2001), Brodeur (2004). Ce type d'ontologie a été décrit et utilisé dans divers domaines tels que l'aménagement et l'urbanisme (Roussey *et al.* (2004)), le traitement d'images (Brisson *et al.* (2007)), ainsi que la recherche d'information (Fu *et al.* (2005)).

L'un des buts du projet GEONTO¹ est de créer une ontologie spécifique de domaine (ici il s'agira de la topographie) puis de l'enrichir de manière automatique. L'ontologie initiale a donc été créée en collaboration avec l'équipe de recherche COGIT de l'IGN impliquée dans le projet (Abadie & Mustière (2008, 2010)). Dans le cadre de cet article nous présentons l'extraction des termes ayant potentiellement une connotation géographiques à partir de textes à des fins d'enrichissement de cette première ontologie. Les textes grand public auxquels nous nous sommes intéressés dans ce travail sont des récits

1. GEONTO <http://geonto.lri.fr/>, est un projet ANR (ANR-07-MDCO-005-04).

de voyage qui comportent des évocations plutôt littéraires (et non expertes) relatives à des déplacements dans un territoire. Nous nous sommes exclusivement intéressés à des termes (comportant au moins un syntagme nominal), se retrouvant à proximité d'une expression de localisations géographiques nommées, et impliqués dans une relation de dépendance grammaticale avec un verbe de déplacement ou un verbe de perception et éventuellement avec une préposition. Notre méthode de modélisation et d'extraction de ces informations géographiques est détaillée dans la section 3.

2 Etat de l'art

La détection des entités nommées est un problème connu qui joue un rôle important dans le traitement automatique de la langue naturelle (TAL) et dans l'extraction d'information. Dans les travaux de (Poibeau (2003), Sagot & Boullier (2008)), les noms de personnes, d'entreprises, de lieux géographiques, etc. sont considérés comme renvoyant à des entités nommées. Nous nous intéressons exclusivement aux entités nommées géographiques.

Selon Borillo (Borillo (1998)), un lieu est une portion de l'espace matériel dans lequel nous nous situons et nous évoluons. Un lieu (ex : ville de Pau) se compose de plusieurs éléments dont les deux principaux sont une entité concrète (ex : ville), et un repère spatial (ex : Pau). Le couple (entité concrète, repère spatial) de Borillo vient corroborer la proposition, plus ancienne, de Vandeloise (1986) impliquant les concepts de cible et de site.

La mise en relation du couple (entité concrète, repère spatial) peut être réalisées par une simple préposition (la ville de Pau) ou via des expressions plus complexes que nous appellerons : indirection. Cette mise en relation, qu'elle soit simple ou complexe, exprime une relation spatiale. Il existe deux types d'approches de modélisation et de raisonnement à travers des relations spatiales : l'approche quantitative (Balbiani & Muller (2000), Vieu (1997), etc.) et l'approche qualitative (Allen (1991), Freksa (1992), Frank (1996), etc.). L'approche quantitative prend en compte les aspects mesurables relatifs aux lieux tels que la longitude et la latitude, tandis que l'approche qualitative raisonne sur les évocations spatiales sans utiliser de représentation numérique. Selon ces approches, les relations spatiales peuvent être classifiées en trois catégories principales : topologiques (Egenhofer & R.D. (1991)) (ex : de, à côté de, etc.), directionnelles (Ligozat (1998)) (ex : au sud de, etc.), et métriques (ex : à 10 km de, etc.).

Selon Talmy (Talmy (2000)), dans les langues latines comme le français, le mouvement est caractérisé par le verbe. Plusieurs travaux linguistiques (Boons (1987), Laur (1991), Sarda (2000)) ont été réalisés afin d'étudier le rôle des verbes de déplacement dans la langue. Ces auteurs ont proposé une catégorisation des polarités des verbes de déplacement. Les polarités sont : initial (ex : quitter, etc.), médian (ex : visiter, etc), ou final (ex : arriver, etc.). Ces polarités permettent d'exprimer la direction du déplacement. D'après une étude réalisée par P. Loustau (Loustau (2008)), l'expression du déplacement est essentielle dans le récit de voyage, il a donc utilisé avec succès le principe de polarité des verbes de déplacement pour extraire automatiquement des itinéraires décrits dans des récits de voyage. D'après une seconde étude sur le même corpus de récits de voyage, les verbes de perception sont apparus comme ayant également une importance

particulière, notamment lorsque l'auteur souhaite rendre compte de certaines actions ou de sensations.

Notre modélisation s'appuie, d'une part, sur des noms de lieux, un lexique et une structure locale permettant d'exprimer des relations spatiales et, d'autre part, sur les verbes de mouvement et de perception. Ainsi, cette modélisation opérationnalisée dans une chaîne de traitement automatique permet d'extraire un lexique pouvant avoir une connotation géographique. Dans le cadre du projet Geonto, la finalité applicative étant d'utiliser chaque terme de ce lexique comme entrée candidate à l'enrichissement de l'ontologie géographique.

3 Méthode & implémentation

3.1 La problématique

Le lexique à constituer est composé de syntagmes nominaux (ex : bassins lacustres, nappes d'eau, territoire aride, vallée, etc.) qu'il faut sélectionner et extraire dans un corpus de textes. Afin d'opérer automatiquement cette extraction, il faut avoir un modèle qui permette de différencier ces syntagmes parmi tous ceux contenus dans les textes.

En reprenant le principe du couple (entité concrète, repère spatial) de A. Borillo (op. cit.), une première proposition serait de considérer que les syntagmes nominaux (ou entité concrète) associés à des noms de lieux (ou repère spatial) sont autant de candidats potentiels à l'enrichissement de l'ontologie géographique (ex : la région d'Aragon). Toutefois, cette proposition si elle est intéressante, reste incomplète car des expressions non géographiques peuvent également être associées à une entité nommée de lieu (ex : maire d'Artouste, jambon de Bayonne). Afin de restreindre le champ des possibles, nous associons à cette première proposition les contraintes supplémentaires suivantes :

- l'expression extraite s'avère être une indirection détectable par le couple (forme, lexique) (ex : la région d'Aragon)
- l'expression extraite est décrite avant une indirection (ex : territoire au sud de l'Aragon) ou après (ex : au sud de la région d'Aragon)
- l'expression extraite est un élément d'une structure lexico-syntaxique (ex : descendre sur le territoire au sud de la région d'Aragon) qui évoque ainsi l'expression du déplacement ou la description de lieu.

Cela nous amène à proposer le modèle VPT. Ce modèle est détaillé par la suite.

3.2 Le modèle VPT

Le modèle proposé (figure 1) réuni de manière parcimonieuse les différents principes et concepts des travaux précités relatifs à l'expression spatiale, aux relations spatiales, et aux déplacements et intègre les verbes de perception. Le modèle global est ainsi décrit par un triplet (**VPT**) qui se compose d'un verbe de déplacement ou de perception (**V**), d'une préposition (**P**), et d'un *Toponyme* (**T**). Le *Toponyme* est défini de façon récursive à partir des noms toponymiques (le nom de lieu connu), des indirections (ou relations spatiales), et des termes associés. Le caractère étoile (*) dans la figure 1 signifie que le

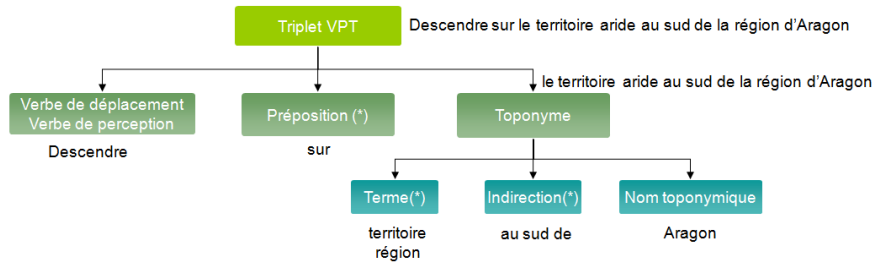


FIGURE 1 – VPT : Modèle spatial

composant correspondant pourra figurer de zéro à plusieurs fois dans le modèle VPT. Les autres composants doivent y apparaître au moins une fois.

3.3 La chaîne de marquage des triplets VPT

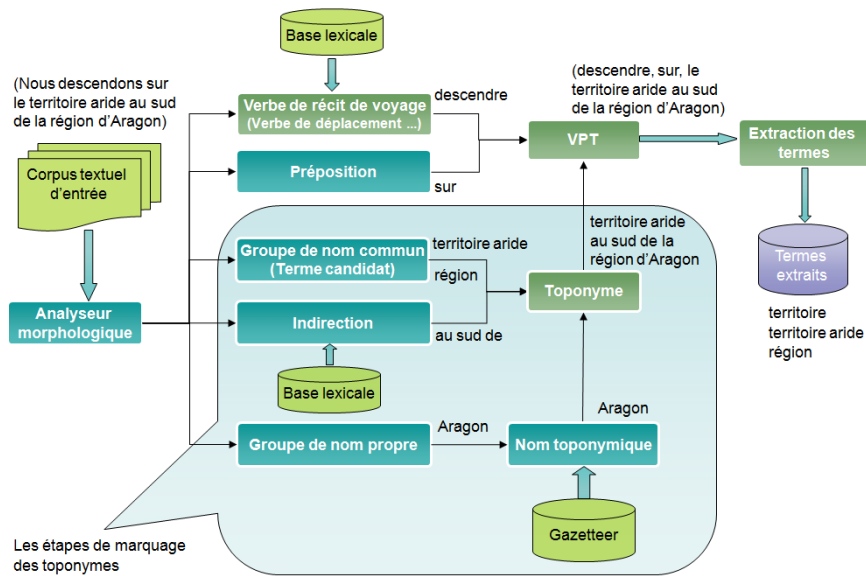


FIGURE 2 – Illustration de la chaîne de marquage des triplets VPT

L'objectif de notre chaîne de traitement (figure 2) est de marquer les triplets VPT, puis d'en extraire des termes ayant potentiellement une connotation géographique. Les éléments du modèle VPT y sont repérés au fur et à mesure par des patterns linguistiques

Hearst (1992)² exprimé par des grammaires en combinant avec divers ressources lexicales.

Tout d'abord, le texte est traité par un analyseur morpho-syntaxique (i.e. TreeTagger Schmidt (1994)). L'objectif est de tokeniser le texte et ensuite d'associer à chaque token une étiquette qui indique sa catégorie grammaticale (nom, verbe, adjectif, etc.). A partir de la sortie de l'analyseur, nous déterminons les étiquettes associées aux groupes de mots : le groupe de nom commun, le groupe de noms propres, l'indirection, le verbe de déplacement. Cela se fait en combinant des règles de grammaire hors-contexte (qui permettent de remplacer plusieurs entités consécutives par une seule de plus haut niveau) et des ressources lexicales. Le paragraphe suivant présente un exemple d'un ensemble de règles qui marque les groupes de noms communs (futurs termes candidats à l'enrichissement). La ligne 3 signifie que s'il y a une séquence de tokens composée d'un nom commun suivi d'un adjectif, cette séquence sera représentée par un nouveau groupe de type nom commun.

```
1 root(commonNoun:X) -> group(X).
2 group(adjectif:A..nom:N) -> adjectif(A), group(N). %case 1 : ex : belle ville
3 group(nom:N..adjectif:A) -> commonNoun(N), adjectif(A). %case2:territoire aride
4 %case 3 :hotel de ville (recursively)
5 group(nom1:N1..nom2:N2) -> commonNoun(N1), %hotel
6 (ls_token('de');ls_token('d\ ');ls_token('des')), %de
7 group(N2). %ville
8 group(X) -> commonNoun(X).%cas4 : territoire
9 commonNoun(adjectif:' '..nom:lemma:X) -> ls_token(_, lemma:X..stag:com, token).
10 adjectif(A) -> A@tag:adj.
```

Nous avons construit une base lexicale des verbes qui comporte 47 verbes de déplacement en nous appuyant sur les travaux précités et 50 verbes de perception. D'autre part, la base lexicale des indirections est construite afin de correspondre à au moins une des relations spatiales spécifiées dans les travaux de Egenhofer & R.D. (1991) et de Ligozat (1998). Afin de valider les groupes de noms propres comme les noms toponymiques, nous utilisons deux gazetteers (des bases de données de noms de lieux) : BDNyme de l'IGN³, et Geonames⁴. Ainsi, les Toponymes sont marqués par des règles de la grammaire hors-contexte à partir des groupes de nom communs, des indirections, et des noms toponymiques préalablement détectés.

Enfin, les triplets VPT sont marqués à partir des verbes de déplacement | perception, des prépositions, et des Toponymes précédemment marqués. En continuité des travaux de Loustau (2008), nous marquons les triplets VPT à l'aide d'un transducteur (figure 3). Initialement, le lecteur est à l'état 0. Le lecteur du transducteur parcourt la phrase de token à token. C'est la sémantique du token qui détermine la transformation de l'état du transducteur. Par exemple, la polarité du verbe de déplacement "descendre" fait passer le transducteur de l'état 0 à l'état 7. Ensuite c'est la polarité de la préposition associée au verbe (selon Laur (1991)) qui détermine le passage à l'état 8. Enfin, lorsque le parcours du Toponyme "territoire aride au sud de la région d'Aragon" est terminé, le transducteur

2. Hearst a proposés des patterns lexico-syntaxiques pour l'extraction des termes hyponymiques

3. Institut géographique national : <http://www.ign.fr>

4. <http://geonames.org>

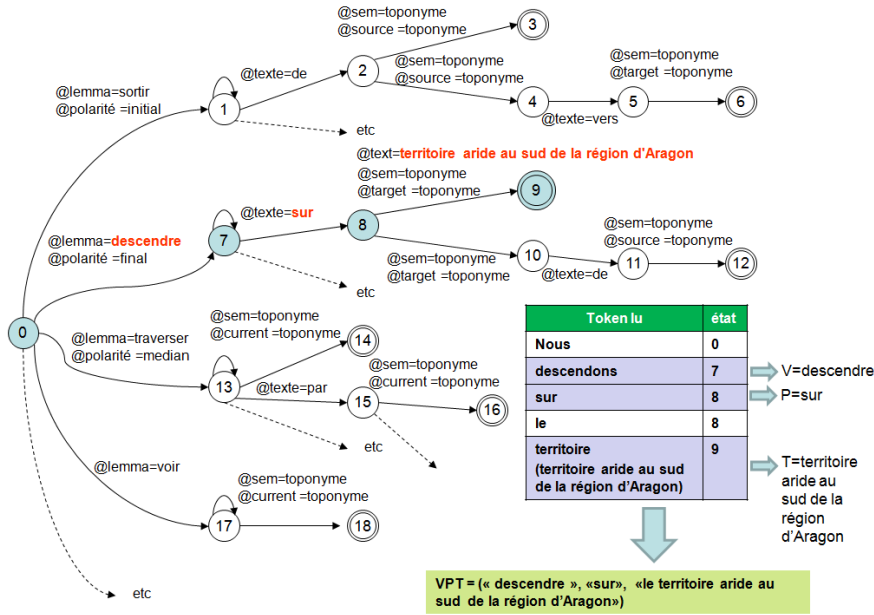


FIGURE 3 – Transducteur de marquage de triplets VPT

passé à l'état final (9). Ainsi, le triplet VPT (descendre, sur, territoire aride au sud de la région d'Aragon) est marqué.

3.4 Comment extraire des informations géographiques à partir des triplets VPT marqués

Comme mentionné dans l'introduction l'objectif premier de ce travail consiste à utiliser un corpus de textes littéraires de type récits de voyages.

Selon notre approche, tous les concepts à extraire sont englobés dans les triplets VPT. Autrement dit, ils se trouvent dans les groupes de nom communs du Toponyme composant du triplet VPT marqué. Formellement, les triplets VPT repérés sont décrits par une structure XML⁵. La tâche principale de cette étape est d'exploiter cette structure pour récupérer les termes. Dans cette structure, tous les éléments du triplet VPT sont organisés de manière arborescente. Cette structuration facilite l'extraction des termes à partir des triplets VPT préalablement marqué (qui sont des groupes des noms communs dans les triplets VPT). De plus cette structure rend l'étape d'extraction des termes moins coûteuse que les étapes de marquages des triplets VPT. Dans notre exemple, à partir du triplet VPT ("descendre", "sur", "territoire aride au sud de la région de l'Aragon"), nous avons extrait les termes "région", "territoire", et "territoire aride".

5. <http://www.w3.org/XML>

4 Expérimentation & évaluation

4.1 Avantages du modèle VPT

L'intérêt du marquage du verbe de déplacement | perception est de mieux cibler le contexte du terme à extraire. Par exemple, pour la phrase "nous mangeons du jambon de Bayonne", le terme "jambon" associé au nom de lieu "Bayonne" ne sera pas extrait car ne se trouvant pas dans un triplet VPT. De plus, l'avantage du marquage des Toponymes est de permettre d'extraire non seulement des termes (ex : région) directement associés à un lieu nommé, mais également des termes indirectement associés (territoire, territoire aride). En effet, la récursivité du Toponyme et du groupe de nom commun permet de traiter des expressions complexes de l'espace à travers des : noms toponymiques, des indirections (ou relations spatiales), des termes candidats.

En fait, dans la chaîne de traitement, la couverture des règles de notre grammaire semble suffisante, elle est actuellement capable de détecter 4 cas de groupes de noms communs, 14 cas de groupes de noms propres, 10 cas de Toponymes, et 15 cas de triplets VPT. La combinaison de ces règles permet de traiter des phrases très complexes. Par exemple : "**Je partis en conséquence pour Bagnères de Luchon une seconde fois et, passant par le Tourmalet, un beau village près d'Hourquette d'Arreau et un autre au sud de la belle vallée de Louron**". Dans ce cas, les règles étiquetant les textes de manière incrémentale permettent de marquer deux triplets VPT, le second étant construit par une liste de Toponyme, comme montré par la chaîne sur la figure 2.

4.2 Evaluation quantitative

Nous avons implémenté notre chaîne et expérimenté notre méthode sur 12 livres (récits de voyage) numérisés et fournis par la médiathèque de Pau. 10555 d'occurrences de verbes de déplacement ont été marqués. En se basant sur ces verbes, nous avons marqué 1390 triplets VPT. Ces triplets ont permis d'extraire 283 termes dont 178 validés par des experts comme termes à connotation géographique voire topographique. La précision est donc de $178/283 = 0,63$. Parmi ces 178 termes topographiques extraits et validés, 134 termes n'existent pas dans l'ontologie de l'IGN (Abadie & Mustière (2008, 2010)). L'ensemble de ces termes (voir tableau 1) a donc été proposé pour l'enrichissement de cette ontologie.

Précisons que le calcul du rappel demande une annotation manuelle des corpus, tâche reconnue comme coûteuse, que nous n'avons pas encore réalisées. Néanmoins afin de réaliser une première évaluation, nous avons fait une estimation du rappel en calculant sa valeur pour chaque terme extrait validé (les tableaux 1(a), 1(b)) grâce à un module logiciel. Ces valeurs sont donc approximatives car si le module compte bien les occurrences d'un terme celui-ci ne peut pas distinguer le contexte dans lequel celui-ci est évoqué dans sa connotation géographique. Par exemple, pour le terme "pied", s'il faut considérer son occurrence dans un contexte comme : "Je me trouve au pied de Pyrénées", il ne le faut pas dans un contexte comme : "Je vais à pied à Pau".

L'expérimentation a d'autre part permis de faire une première évaluation du gain apporté par l'ajout des verbes de perception (voir les tableaux 2). La deuxième colonne des tableaux indique le nombre de fois où le verbe est associé au triplet VPT. La dernière

Terme	Extrait	Total	Rappel	Terme	Extrait	Total	Rappel	Terme	Extrait
conriche	1	1	1.000	col	35	354	0.099	camarades	1
rameau	2	3	0.667	grotte	5	59	0.085	direction	1
logements	1	2	0.500	route	23	369	0.062	jour	1
pont de pierre	1	2	0.500	cirque	12	201	0.060	minutes	1
mer de glace	2	5	0.400	pic	34	578	0.059	malades	1
crête rocheuse	2	5	0.400	sommet	19	468	0.041	septembre	1
auberge	4	11	0.364	chemin	9	252	0.036	vie	2
étang	1	3	0.333	brèche	9	256	0.035	maire	2
coteau	1	3	0.333	pays	7	227	0.031	profondeur	2
chaumière	2	7	0.286	crête	8	269	0.030	suite	2
ourquette	2	8	0.250	glacier	11	389	0.028	Marbrière	1
hôtel	4	24	0.167	sommets	4	216	0.019	Enfance	1
château	12	74	0.162	montagne	4	490	0.008	Arête	6
lac	83	533	0.156	chaîne	1	205	0.005	Fonds	6
vallée	117	790	0.148	pied	1	358	0.003	Examen	1

TABLE 1 – Des termes extraits : (a, b) validés ; (c) non validés

est le pourcentage de cette fréquence par rapport au nombre total d'occurrences du verbe.

Verbe	VPT	Taux	Verbe	VPT	Taux
arriver	127	19.4%	voir	121	8.5%
passer	114	14.4%	apercevoir	26	12.6%
aller	109	10.0%	reconnaître	20	8.1%
venir	80	9.9%	paraître	17	3.5%
descendre	79	14.7%	jeter	16	11.1%
partir	79	21.5%	découvrir	16	8.7%
monter	78	12.8%	aimer	13	8.5%
conduire	77	25.9%	explorer	11	15.5%
quitter	51	23.1%	coucher	11	8.3%
suivre	46	8.1%	entendre	10	5.7%

(a) Verbe de déplacement

(b) Verbe de perception

TABLE 2 – Classification et distribution des verbes plus fréquents dans les corpus

4.3 Bruit et silence

L'expérimentation a montré plusieurs cas de bruits et de silences. Nous citons ici quelques raisons pour lesquelles les termes extraits n'ont pas une connotation géographique.

- La polysémie des verbes : Par exemple, "...j'ai **quitté le confortable de la vie de Bordeaux...**" A cause de la polysémie du verbe "quitter", le terme "vie" est mal extrait.
- Des termes exprimant du temps : Voyons les termes "bourg" et "journée" dans deux phrases suivantes *Nous devons la parcourir dans toute son étendue, visiter en passant le bourg de Vieille*

J'ai pris un cheval pour aller à Gèdre et passer la journée dans la vallée de Héas

- L'inexactitude de la sortie de l'analyseur morphologique : Par exemple *Il serait devenu un peu fier vis-à-vis de ses camarades d'Arrens...* , vis dans "vis-à-vis" a mal considéré comme le verbe de perception "voir".

En analysant le résultat de traitement, nous détectons ainsi les cas dont les termes ne sont pas bien extraits :

- Les gazetteers incomplets : pour le texte "...nous grimpons sur le plateau d'Esquierry", Esquierry n'existe pas dans les gazetteers.
- La structure verbale passive : "...le pic de Néthou , n'a été encore gravi par personne..."
- La faute de l'analyseur morphologique : "...Arrivés au fond de la vallée du Plan-des-Etangs, ...", "Etangs" est traité par TreeTagger comme "nom commun" donc n'est pas attaché à "Plan" pour constituer un nom toponymique "Plan-des-Etangs".
- Aucune indication permet de filtrer le syntagme à extraire : "Pour ce qui est des variations du niveau du gouffre, il y a, en effet, une crue et une baisse..."
- Des triplets VPT implicites : Dans certains cas, les termes à extraire ne sont pas directement attachés aux triplets VPT, comme le verbe "ville" dans cette phrase : "Jusqu'à cette ville, nous avons longé la Garonne".

5 Conclusion et perspectives

Dans cet article, nous avons proposé un modèle dédié à la représentation d'informations spatiales et à l'extraction d'un lexique à connotation géographique. Nous avons développé une chaîne de traitement complète basée sur l'approche lexico-syntaxique pour implémenter et expérimenter ce modèle. Cette chaîne de traitement est utilisée pour l'analyse d'un fonds documentaires de textes littéraires de type récits de voyage. Ces termes extraits sont ensuite proposés comme candidats à l'enrichissement d'une ontologie géographique. Un premier traitement a été réalisé sur un ensemble de 12 livres numérisés ; 283 termes topographiques ont été extraits automatiquement avec une précision de 0,63. Le patron VPT que nous avons proposé dans cet article peut également être utilisé dans un but de détection des noms toponymiques dans le texte ou encore de désambiguïsation des entités nommées. Nous avons l'ambition d'étendre notre méthode pour pouvoir extraire également des termes non directement attachés aux triplets VPT, par exemple, le terme "ville" dans la phrase "Jusqu'à cette ville, nous avons longé la Garonne". Pour le traitement de tel cas, nous travaillons actuellement sur un modèle basé sur les relations n-aires dont VPT est un cas particulier.

Références

- ABADIE N. & MUSTIÈRE S. (2008). Création d'une taxonomie géographique à partir des spécifications de bases de données. In *SAGEO*, Montpellier.
- ABADIE N. & MUSTIÈRE S. (2010). Constitution et exploitation d'une taxonomie géographique à partir des spécifications de bases de données. *Revue Internationale de Géomatique*, **20**(2), 145–174.

- ALLEN J. F. (1991). Planning as temporal reasoning. *KR*, p. 3–14.
- BALBIANI P. & MULLER P. (2000). Le raisonnement spatial. In C. G. H. PRADE, R. JEANSOULIN, Ed., *Le temps, l'espace et l'évolutif en sciences du traitement de l'information*, p. 33–53. Cepadues Editions.
- BOONS J.-P. (1987). La notion sémantique de déplacement dans une classification syntaxique des verbes locatifs. *LANGUE FRANÇAISE*, **76**(76), 5–40.
- BORILLO A. (1998). *L'espace et son expression en français, L'essentiel*. Orphrys.
- BRISSON R., BOUSSAÏD O., GAŃCARSKI P., PUISSANT A. & N. D. (2007). Navigation et appariement d'objets géographiques dans une ontologie. In *EGC '07*, p. 391–396.
- BRODEUR J. (2004). *Interopérabilité des données géospatiales : Élaboration du concept de proximité géosémantique*. PhD thesis, Université Laval, Québec.
- EGENHOFER M. & R.D. F. (1991). Point-set topological spatial relations. *International Journal for Geographical Information Systems*, **5**(2), 161–174.
- FRANK A. U. (1996). Qualitative spatial reasoning : Cardinal directions as an example. *International Journal of Geographical Information Science*, **10**(3), 269–290.
- FREKSA C. (1992). Using orientation information for qualitative spatial reasoning.
- FU G., JONES C. & ABDELMOTY A. (2005). Building a geographical ontology for intelligent spatial search on the web. In *IASTED International Conference on Databases and Applications*.
- HEARST M. A. (1992). Automatic acquisition of hyponyms from large text corpora. In *In Proceedings of the 14th International Conference on Computational Linguistics*, p. 539–545.
- LAUR D. (1991). *Sémantique du déplacement et de la localisation en français : une étude des verbes, des prépositions et de leur relation dans la phrase simple*. PhD thesis, Université de Toulouse II.
- LIGOZAT G. (1998). Reasoning about cardinal directions. *Visual Languages and Computing*, **9**(1), 23–44.
- LOUSTAU P. (2008). *Interprétation automatique d'itinéraires dans des recits de voyage*. PhD thesis, Université de Pau et des Pays de l'Adour.
- POIBEAU T. (2003). *Extraction automatique d'information*. Hermès Lavoisier.
- ROUSSEY C., LAURINI R., BEAULIEU C., TARDY Y. & ZIMMERMANN M. (2004). Un retour d'expérience pour la construction d'une ontologie urbaine. *Revue internationale de Géomatique*, **14**(2), 217–237.
- SAGOT B. & BOULLIER P. (2008). Sxpipe2 : architecture pour le traitement présyntaxique de corpus bruts. *TAL*, **49**(2), 155–188.
- SARDA L. (2000). L'expression du déplacement dans la construction transitive directe. *Syntaxe et Sémantique*, p. 121–137.
- SCHMIDT H. (1994). Probabilistic part-of-speech tagging using decision trees. In *International Conference on New Methods in Language Processing*, Manchester, UK.
- TALMY L. (2000). *Toward a Cognitive Semantics*, chapter How language structures space. The MIT Press.
- UITERMARK H. (2001). *Ontology-Based Geographic Data Set Integration*. PhD thesis, Universiteit Twente, the Netherlands.
- VANDELOISE C. (1986). *L'espace en français*. Paris, France : Seuil.
- VIEU L. (1997). Spatial representation and reasoning in artificial intelligence. *Spatial and Temporal Reasoning*, p. 5–41.