Sous-ensembles flous définis sur une ontologie

Rallou Thomopoulos[1], Patrice Buche[1], Ollivier Haemmerlé[1,2]

[1]UMR INA P-G/INRA BIA, 16 rue Claude Bernard,
F-75231 Paris Cedex 05, France
{prénom.nom}@inapg.fr
http ://www.inapg.inra.fr/ens_rech/mathinfo/index.html
[2]LRI (UMR CNRS 8623 - Université Paris-Sud) / INRIA (Futurs), Bâtiment 490,
F-91405 Orsay Cedex, France

Résumé. Les sous-ensembles flous peuvent être utilisés pour représenter des valeurs imprécises, comme un intervalle aux limites mal définies. Ils peuvent également servir à l'expression de préférences dans les critères de sélection de requêtes en bases de données. En représentation des connaissances, l'utilisation de hiérarchies de types est largement répandue afin de modéliser les relations existant entre les types d'objets d'un domaine donné. Nous nous intéressons aux sous-ensembles flous dont le domaine de définition est une hiérarchie d'éléments partiellement ordonnés par la relation "sorte de", que nous appelons ontologie. Nous introduisons la notion de sous-ensemble flou défini sur une partie de l'ontologie, puis sa forme développée définie sur l'ensemble de l'ontologie, que nous appelons extension du sous-ensemble flou. Des classes d'équivalence de sous-ensembles flous définis sur une ontologie peuvent être caractérisées par un représentant unique que nous appelons sous-ensemble flou minimal. Nous concluons par un exemple d'application dans un système d'information relatif à la prévention du risque microbiologique en sécurité alimentaire.

1 Introduction

Le projet de recherche dans lequel ont été menés les travaux présentés dans cet article est un programme national, le projet Sym'Previus, associant des organismes de recherche (INRA, INA P-G), des industriels (Danone, Bongrain, ...) et des centres techniques (ADRIA, ...), avec pour objectif la construction d'un outil d'analyse des risques microbiologiques dans les produits alimentaires. Cet outil s'appuie sur une base de connaissances qui contient des informations issues de publications scientifiques et de données industrielles en microbiologie. Ces informations décrivent le comportement (croissance, décroissance ou survie) de germes pathogènes (par exemple Listeria Monocytogenes) au cours du cycle de vie des produits alimentaires.

Nous avons été amenés à proposer des solutions afin de prendre en compte plusieurs spécificités de ce projet :

– les données stockées peuvent être imprécises. En effet, (i) la complexité des processus biologiques rend difficile la reproduction à l'identique des conditions expérimentales : des résultats différents pour une même expérience seront synthétisés par un intervalle de valeurs, (ii) les données sont parfois vagues, comme par exemple “Sur des produits ayant une activité de l’eau (a_w) **faible**...”, (iii) les techniques de mesure sont souvent limitées par des seuils de détection ;

– une base de données contenant des résultats d’expériences en microbiologie est incomplète par nature : les études réalisées sont en nombre limité et de nombreuses expérimentations possibles restent inexplorées.

Nous nous sommes intéressés à la théorie des sous-ensembles flous et des possibilités, afin de représenter l’imprécision des données, mais également d’autoriser l’expression de préférences dans les critères de sélection des requêtes. L’utilisateur peut ainsi exprimer des requêtes “larges”, fournissant des résultats classés par ordre de préférence.

Les terminologies utilisées dans le domaine microbiologique sont souvent hiérarchisables par la relation “sorte-de” (par exemple les taxonomies de germes pathogènes, de produits alimentaires, ...). Nous avons donc été amenés à étudier l’expression de sous-ensembles flous définis sur des domaines de valeurs hiérarchisés.

Différents travaux antérieurs se sont penchés sur l’introduction du flou dans une hiérarchie, notamment par la prise en compte du flou dans le modèle objet [Cross, 1996, Lee *et al.*, 2001].

On peut distinguer plusieurs types de problématiques traitées. La distinction proposée ici est schématique, chacun des travaux cités abordant généralement plusieurs thématiques, dans le cadre du modèle objet pour la plupart : (i) les valeurs des attributs peuvent être floues [Vignard, 1985] ; (ii) les classes peuvent être décrites par des attributs dont le domaine de valeurs n’est pas strictement délimité. [Rossazza *et al.*, 1997] distingue la notion de domaine de valeurs autorisées et celle de domaine de valeurs typiques. Ces domaines sont graduels et peuvent être représentés par des sous-ensembles flous ; (iii) les attributs d’une classe peuvent être pondérés [Granger, 1988] [Torasso et Console, 1989]. Les poids associés aux attributs peuvent par exemple être utilisés dans un but de classification : pour décider à quelle classe appartient un objet donné, les attributs de plus grand poids seront les plus déterminants ; (iv) les liens d’héritage (entre une classe et une sous-classe) peuvent être flous [George *et al.*, 1993, Graham et Jones, 1988, Bordogna *et al.*, 1994, Itzkovich et Hawkes, 1994]. Différents degrés peuvent être utilisés pour représenter la notion d’héritage partiel ; (v) l’utilisation de règles permet de définir des relations supplémentaires entre classes. [Ralescu et Berenji, 1989, Lukasiewicz, 1995] introduisent des règles associées à des limites inférieures et supérieures de probabilité ; (vi) un objet peut appartenir partiellement à une classe [Gyseghem *et al.*, 1993, Gomez *et al.*, 1995, Baldwin et Martin, 1996]. Cette approche peut être liée à une problématique de classification, mais pas nécessairement : [Cao, 1999] définit l’appartenance simultanée d’un objet à plusieurs classes non comparables entre elles, avec des valeurs de vérité floues différentes, par le biais des types flous conjonctifs, dans le modèle des graphes conceptuels ; (vii) un langage d’interrogation d’un modèle orienté objet flou est proposé dans [Na et Park, 1996, Nepal *et al.*, 1999].

À notre connaissance, l’utilisation d’un domaine hiérarchisé dans l’expression d’un sous-ensemble flou n’a pas encore été étudiée. C’est ce que nous proposons de faire dans cet article. Dans la section 2 nous rappelons brièvement la théorie des sous-ensembles flous. La section 3 présente les sous-ensembles flous définis sur une ontologie (que nous notons SEFO). Nous introduisons la notion de SEFO défini sur une partie des éléments de l'ontologie ; nous proposons ensuite un algorithme permettant de construire, à partir d'un SEFO, sa forme développée définie sur tous les éléments de l'ontologie, que nous appelons extension du SEFO. Nous montrons ensuite qu'il existe des classes d'équivalence de SEFO relativement à leur extension, et nous mettons en évidence l'existence et l'unicité d'un représentant de chacune des classes d'équivalence, que nous appelons SEFO minimal.

2 Notions préliminaires : les sous-ensembles flous

Dans cet article, nous utilisons la théorie des sous-ensembles flous et des possibilités proposée dans [Zadeh, 1965, Zadeh, 1978]. Le formalisme des sous-ensemble flous est utilisé de deux manières différentes : (i) dans une requête, afin d'exprimer les préférences de l'utilisateur concernant un critère de sélection. Par exemple, le sous-ensemble flou *DuréeLongue* de la figure 1 permet à l'utilisateur d'indiquer qu'une réponse comprise entre 50 et 70 est totalement satisfaisante, les valeurs proches de cet intervalle étant également acceptables avec un moindre degré de satisfaction ; (ii) dans les données, pour représenter des valeurs imprécises. Un sous-ensemble flou peut être défini sur un domaine continu ou discret, comme l'illustre la figure 1.

Définition 1 *Un* sous-ensemble flou A *sur un domaine* X *est défini par une fonction d'appartenance* μ_A *de* X *dans [0, 1] qui associe, à chaque élément* x *de* X*, le degré avec lequel* x *appartient à* A*.*

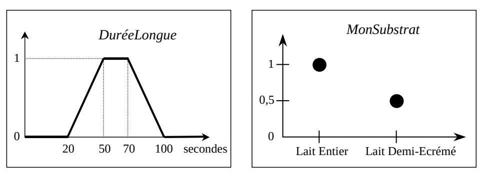

FIG. 1 – Deux sous-ensembles flous

Définition 2 *Soient* A *et* B *deux sous-ensembles flous de* X*. On dit que* B *est* inclus *dans* A *(noté* $B \subseteq A$*), si et seulement si pour tout* x *de* X*,* $\mu_B(x) \leq \mu_A(x)$*.*

Dans le but de mesurer l'adéquation d'une donnée imprécise à un critère flou, deux degrés sont classiquement utilisés : (i) le degré de possibilité [Zadeh, 1978] et (ii) le degré de nécessité [Dubois et Prade, 1988]. Ces degrés sont comparables respectivement à une limite optimiste et pessimiste du degré avec lequel la donnée satisfait le critère, ou à un degré d'intersection et d'inclusion de la donnée dans le critère.

Définition 3 *Le* degré de possibilité *d'adéquation entre* A *et* B*, noté* $\Pi(A\ ;B)$*, est défini par :* $\Pi(A\ ;B) = sup_{x \in X} min(\mu_A(x), \mu_B(x))$*.*
Le degré de nécessité *d'adéquation de* B *à* A*, noté* $N(A\ ;B)$*, est défini par :* $N(A\ ;B) = inf_{x \in X} max(\mu_A(x), 1 - \mu_B(x))$*.*

2.1 Ontologies

3 Sous-ensembles flous définis sur une ontologie

3.1 Présentation

Nous utilisons des ontologies comme domaine de valeur pour les SEFO que nous définissons ci-dessous. La figure 2 présente un exemple d'ontologie.

Définition 4 *Une ontologie Ω est un ensemble d'éléments partiellement ordonnés par la relation "sorte-de". Soient deux éléments v_1 et v_2 appartenant à Ω. Nous dirons que v_1 est une généralisation de v_2 et que v_2 est une spécialisation de v_1 (noté $v_2 \leq v_1$) si v_2 est un prédécesseur de v_1 dans l'ordre partiel induit par la relation "sorte de" de Ω.*

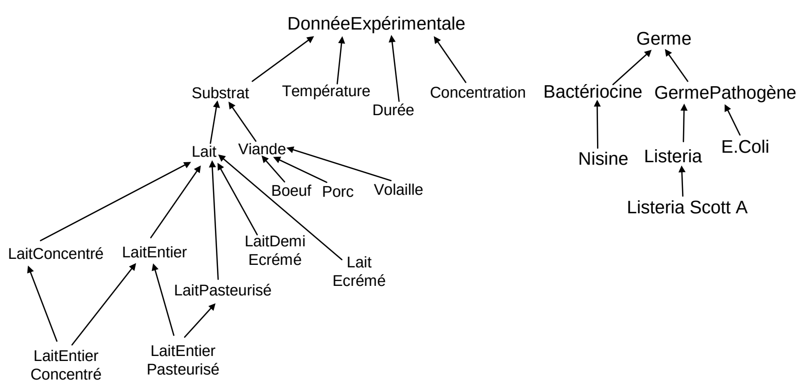

FIG. 2 – Exemple d'ontologie

Définition 5 *Un* sous-ensemble flou défini sur une ontologie (ou SEFO) *S est un sous-ensemble flou défini sur un domaine D constitué d'un sous-ensemble des éléments de l'ontologie Ω.*

Par exemple, le sous-ensemble flou *MonSubstrat* représenté sur la Figure 1 est un SEFO conforme à la Définition 5. Il est défini sur un sous-ensemble des éléments de l'ontologie présentée sur la Figure 2.

Notons qu'aucune restriction n'a été imposée concernant les éléments qui composent le domaine de définition d'un SEFO. En particulier, l'utilisateur est libre d'associer à deux éléments v_1 et v_2 tels que $v_2 \leq v_1$ des degrés d_1 et d_2 différents. $d_2 \leq d_1$ représente une sémantique de restriction pour v_2 par rapport à v_1, tandis que $d_2 \geq d_1$ représente une sémantique de renforcement pour v_2 par rapport à v_1.

Par exemple, si l'utilisateur s'intéresse aux laits pasteurisés, mais souhaite également obtenir des informations complémentaires sur les autres laits, il peut exprimer ses préférences en utilisant dans sa requête le SEFO suivant : {*1/LaitPasteurisé, 0,5/Lait*}. Dans cet exemple, l'élément *LaitPasteurisé* a un degré supérieur à celui de sa généralisation *Lait*, ce qui correspond à une sémantique de renforcement de *LaitPasteurisé* par rapport à *Lait*. À l'inverse, si l'utilisateur s'intéresse à tous les types de lait, mais plus faiblement au lait concentré, du fait de sa moindre teneur en eau, il peut représenter ses préférences par le SEFO suivant : {*1/Lait, 0,2/LaitConcentré*}. Ici, l'élément *LaitConcentré* a un degré inférieur à celui de sa généralisation *Lait*, ce qui correspond à une sémantique de restriction de *LaitConcentré* par rapport à *Lait*.

Les SEFO peuvent également être utilisés dans les données, pour l'expression de connaissances imprécises. Par exemple, si le substrat est désigné dans une publication comme "du lait allégé en matière grasse", il peut être représenté par un SEFO associant le degré 1 à l'élément *LaitEcrémé* et un degré inférieur à l'élément *LaitDemiEcrémé*.

3.2 Notion d'extension d'un SEFO

L'utilisation des SEFO amène deux types de remarques :

- la première remarque est de nature sémantique. Supposons que l'on utilise comme critère de sélection dans une requête le SEFO {*1/LaitEcrémé, 0,5/Lait*} défini sur un sous-ensemble de Ω ; on peut constater qu'il fournit implicitement des informations sur les autres éléments de Ω. On peut en déduire que l'utilisateur s'intéresse aux laits en général, et particulièrement aux laits écrémés, mais qu'il ne s'intéresse pas à la viande par exemple ;
- la seconde remarque est d'ordre opérationnel. Le problème posé par la définition 5 est que deux SEFO différents ne sont pas nécessairement définis sur le même domaine, ce qui empêche d'utiliser les comparaisons classiques entre sous-ensembles flous pour comparer les SEFO (Cf. Définitions 2,3).

Ces constats nous ont amenés à introduire la notion d'*extension d'un SEFO*, qui est une forme développée du SEFO de la définition 5 telle que son domaine de définition soit l'ensemble des éléments de l'ontologie Ω.

Intuitivement, l'extension d'un SEFO prend en compte la relation "sorte de" en propageant à ses spécialisations le degré associé à un élément. Si par exemple, dans une expression de préférences, l'utilisateur s'intéresse à l'élément *Lait*, on considère qu'il s'intéresse à toutes les sortes de *Lait* (*LaitEntier*, *LaitEcrémé*, *LaitPasteurisé*, etc.). À l'inverse, on considère qu'une généralisation d'un élément est trop générale pour être pertinente. Dans notre exemple, les autres substrats (*Substrat* étant une généralisation de *Lait*) n'intéressent pas l'utilisateur.

Définition 6 *Soit S un SEFO défini sur un domaine $dom(S)$ ($dom(S) \subseteq \Omega$) et de fonction d'appartenance μ_S. L'*extension de S*, notée $ext(S)$, est définie sur l'ensemble des éléments de l'ontologie Ω et sa fonction d'appartenance $\mu_{ext(S)}$ est définie comme suit :*

Pour tout élément v de Ω, soit $E_v = \{v_1, \ldots, v_n\}$ l'ensemble des plus petites[1] généralisations de v (au sens large, c'est-à-dire $v_i \geq v$) dans $dom(S)$:

- *si E est non vide, $\mu_{ext(S)}(v) = max_{1 \leq i \leq n}(\mu_S(v_i))$, avec $v_i \in E_v$,*
- *sinon $\mu_{ext(S)}(v) = 0$.*

L'extension d'un SEFO se construit donc selon des règles que l'on peut reformuler intuitivement de la façon suivante. Pour tout $v \in E_v$,

[1] au sens de la relation d'ordre partiel "sorte de" de Ω

1. si v figure dans le SEFO, alors v conserve le même degré dans son extension (cas où $E_v = \{v\}$) ;
2. si v a une unique plus petite généralisation v_1 qui figure dans le SEFO, alors son degré de v_1 est propagé à v dans l'extension (cas où $E_v = \{v1\}$ avec $v_1 > v$) ;
3. si v a plusieurs plus petites généralisations $\{v_1, \ldots, v_n\}$ qui figurent dans le SEFO avec des degrés différents, un choix doit être fait concernant le degré qui sera affecté à v dans l'extension. Le choix proposé dans la définition 6 est de prendre le maximum des degrés des $v_1, \ldots, v_n$;
4. tous les autres éléments, à savoir les généralisations et les éléments non comparables avec ceux figurant dans le SEFO, sont considérés comme non pertinents, on leur associe le degré 0 (cas où $E_v = \emptyset$).

Exemple 1 *La figure 3 montre l'extension du SEFO* {0,8/Lait, 0,3/LaitEntier, 1/LaitConcentré}.

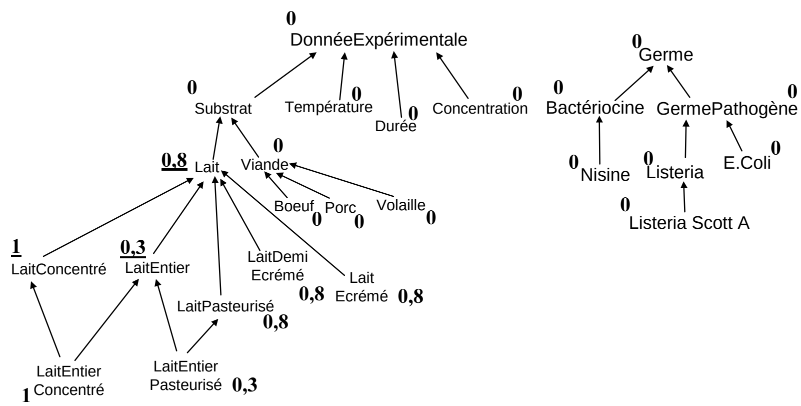

FIG. 3 – Un exemple d'extension de SEFO. Les degrés des éléments du SEFO avant calcul de l'extension sont soulignés.

Dans le SEFO de la figure 3, l'utilisateur a affecté le degré 1 à LaitConcentré *et seulement 0,3 à* LaitEntier. *C'est le maximum de ces deux degrés qui est affecté à leur spécialisation commune* LaitEntierConcentré *dans l'extension.*

En revanche, il n'y a pas d'ambiguïté pour LaitEntierPasteurisé *car rien n'a été spécifié sur* LaitPasteurisé *dans le SEFO. C'est donc le degré associé à* LaitEntier, *c'est-à-dire 0,3, qui est affecté à* LaitEntierPasteurisé *dans l'extension, et non 0,8 qui est associé à l'élément plus général* Lait.

Définition 7 *Soient S_1 et S_2 deux SEFO. Nous dirons que $S_1 \subseteq S_2$ si $ext(S_1) \subseteq ext(S_2)$.*

Définition 8 *Soient S_1 et S_2 deux SEFO.*

1. *le* degré de possibilité d'adéquation *entre S_1 et S_2 noté $\Pi(S_1; S_2)$ est défini par $\Pi(ext(S_1); ext(S_2))$;*

2. *le* degré de nécessité d'adéquation *de* S_2 *à* S_1 *noté* $N(S_1; S_2)$ *est défini par* $N(ext(S_1); ext(S_2))$ *;*

En d'autres termes, les comparaisons et opérations entre SEFO nécessitent un passage aux formes en extension, afin de travailler sur des sous-ensembles flous définis sur un même domaine de définition.

Dans le cas où plusieurs éléments $e_1, \ldots, e_n$ non comparables, figurant dans un SEFO S avec des degrés différents, ont une spécialisation commune e qui n'apparaît pas dans S, le choix du degré à affecter à e dans l'extension de S peut être discuté. En effet, les degrés affectés à $e_1, \ldots, e_n$ apportent des informations éventuellement contradictoires. Il faut pouvoir décider quel degré sera affecté à e. Dans notre application, étant donné que l'on se situe dans un cadre d'interrogation élargie où l'on ne cherche à écarter aucune réponse, nous choisissons $max(\mu_S(e_1), \ldots, \mu_S(e_n))$ comme indiqué dans la définition 6. Ainsi, dans le cas d'un critère de sélection, si l'utilisateur s'intéresse à l'élément *Produit de salaison* avec le degré 1 et à l'élément *Viande crue* avec le degré 0.7, mais ne précise rien sur l'élément *Viande crue salée*, spécialisation commune des deux éléments précédents, on affectera à l'élément *Viande crue salée* le degré 1, soit le plus grand des degrés affectés à ses généralisations.

Dans le cas d'une donnée, le choix de l'opérateur *max*, proposé dans la définition 6, n'écarte aucune valeur possible pour la donnée, mais la rend aussi de ce fait moins spécifique. Nous avons retenu cette solution pour deux raisons : (i) par souci d'homogénéité entre le traitement des requêtes et celui des données ; (ii) pour aller dans le sens d'une interrogation élargie, puisqu'une donnée moins spécifique a plus de chances de partager des valeurs communes avec celles spécifiées dans la requête. Le degré de possibilité d'adéquation de la donnée à la requête peut donc être accru (mais le degré de nécessité, au contraire, peut diminuer).

La complexité du calcul de l'extension d'un SEFO défini sur un domaine $dom(S) \subset \Omega$ est de l'ordre de $|\Omega|.|dom(S)|^2$, à condition que le codage choisi pour la représentation de l'ontologie autorise la comparaison de deux de ses éléments en temps constant. Remarquons que, dans le cas général, le domaine de définition d'un SEFO se limite à quelques éléments : le temps de calcul effectif de l'extension d'un SEFO est donc raisonnable.

3.3 Notion de SEFO minimal

Deux SEFO différents peuvent avoir la même extension, comme l'illustrent les exemples suivants.

Exemple 2 *Les SEFO* Substrat1 = {0,5/Lait} *et* Substrat2 = {0,5/Lait, 0,5/LaitEcrémé} *ont la même extension (0,5 affecté à* Lait *et toutes ses spécialisations, 0 pour tous les autres éléments de l'ontologie).*

Exemple 3 *Les SEFO* Substrat3 = {1/Lait, 0,8/LaitEntier, 1/LaitPasteurisé} *et* Substrat4 = {1/Lait, 0,8/LaitEntier, 1/LaitEntierPasteurisé} *ont la même extension* {*1/Lait, 0,8/LaitEntier, 1/LaitPasteurisé, 1/LaitEntierPasteurisé, 1/LaitEcrémé, 1/LaitDemiEcrémé, 1/LaitConcentré, 0,8/LaitEntierConcentré, (0 pour tous les autres éléments de l'ontologie)*}.

Définition 9 *Deux SEFO S_1 et S_2 sont dits* équivalents *(noté $S_1 \equiv S_2$) si et seulement si ils ont la même extension.*

On constate que *Substrat2* contient le même élément que *Substrat1* avec le même degré, et comporte un élément supplémentaire (*LaitEcrémé*, affecté du degré 0.5). La spécification de cet élément supplémentaire ne change en rien l'extension. On dira que l'élément *0,5/LaitEcrémé* est *déductible* dans *Substrat2*.

Définition 10 *Soit le SEFO S défini sur $dom(S) = \{v_1, v_2, ..., v_j, ..., v_n\}$, et soit S_{-j} le SEFO résultant de la restriction de S au domaine $dom(S)\backslash\{v_j\}$. v_j est dit* déductible *dans S si $\mu_{ext(S_{-j})}(v_j) = \mu_S(v_j)$.*

En revanche *Substrat3* et *Substrat4* ont deux éléments communs (avec les mêmes degrés) et un élément différent : *Substrat3* comporte l'élément *1/LaitPasteurisé*, que *Substrat4* ne comporte pas. À l'inverse, *Substrat4* comporte l'élément *1/LaitEntier-Pasteurisé*, que *Substrat3* ne comporte pas. La suppression de *1/LaitPasteurisé* dans *Substrat3* ou de *1/LaitEntierPasteurisé* dans *Substrat4* modifierait l'extension obtenue.

Dans le SEFO *Substrat3* de l'exemple 3, l'élément *LaitPasteurisé* est déductible. Dans *Substrat1* et *Substrat4*, aucun élément n'est déductible.

Définition 11 *Le SEFO S est dit* minimal *si aucun élément de son domaine n'est déductible (le terme "minimal" n'est pas entendu ici au sens de la cardinalité).*

Les SEFO *Substrat1* et *Substrat4* sont minimaux, au contraire de *Substrat2* et *Substrat3*.

Nous proposons en figure 4 un algorithme qui permet de toujours calculer un SEFO particulier qui est l'unique SEFO minimal.

Propriété 1 *La condition d'arrêt de l'algorithme est toujours atteinte.*

Propriété 2 *Le SEFO $mnl(S)$ obtenu par l'algorithme est minimal.*

Propriété 3 *$mnl(S)$ est unique.*

Ces trois propriétés sont prouvées dans [Thomopoulos, 2003]. La preuve de la propriété 1 repose sur une induction : on montre qu'à la fin de chaque itération, $\mu_{ext(mnl(S))}(v)$ est égal à $\mu_{ext(S)}(v)$ pour tous les éléments déjà atteints par l'algorithme. La preuve de la propriété 2 repose également sur une induction : à chaque itération, l'élément ajouté n'est pas déductible et n'a aucune influence sur le degré des éléments déjà ajoutés. Enfin, la propriété 3 est prouvée par l'absurde.

Exemple 4 *Soit le SEFO* S = {0,5/Lait, 0,5/LaitEcrémé, 1/LaitEntier, 1/LaitEntier-Concentré}*. L'extension de* S *est présentée en figure 5. Le SEFO minimal $mnl(S)$, est obtenu comme suit :*

Au départ, $mnl(S)$ est vide. Son extension est donc le sous-ensemble flou qui associe le degré 0 à tout élément de Ω. On teste si cette extension est égale à $ext(S)$. La réponse est non, tous les éléments n'ayant pas le degré 0 dans $ext(S)$. On effectue donc un parcours selon une extension linéaire $lin(\Omega)$ de Ω.

Calcul d'un SEFO minimal $mnl(S)$ ayant pour extension $ext(S)$
<u>début</u>
  $mnl(S) \leftarrow \emptyset$
  <u>si</u> $(ext(mnl(S)) = ext(S))$
  <u>alors</u>
    arrêt (cas où $ext(S)$ est le SEFO associant le degré 0 à tout élément de Ω)
  <u>sinon</u>
    soit $lin(\Omega)$ une extension linéaire de l'ordre inverse de l'ordre induit par la relation de spécialisation sur les éléments de Ω (chaque élément est examiné postérieurement à ses généralisations)
    <u>répéter</u>
      $v \leftarrow$ élément suivant selon $lin(\Omega)$
      <u>si</u>$(\mu_{ext(mnl(S))}(v) \neq \mu_{ext(S)}(v))$
      <u>alors</u>
        $mnl(S) \leftarrow mnl(S) \cup \{v\}$
        $\mu_{mnl(S)}(v) \leftarrow \mu_{ext(S)}(v)$
      <u>finsi</u>
    <u>jusqu'à</u> $(ext(mnl(S)) = ext(S))$
  <u>finsi</u>
<u>fin</u>

Fig. 4 – Algorithme de calcul d'un SEFO minimal

On commence par examiner un des éléments n'ayant aucune généralisation dans Ω *(ici l'élément* DonnéeExpérimentale *par exemple).*

Cet élément a le même degré 0 dans $ext(mnl(S))$ *et dans* $ext(S)$*. On continue donc le parcours en examinant l'élément* Substrat*. Lui aussi a le même degré dans* $ext(mnl(S))$ *et dans* $ext(S)$*. On examine ensuite l'élément* Lait*. Il a le degré 0 dans* $ext(mnl(S))$*, tandis que son degré est 0,5 dans* $ext(S)$*.* Lait *est donc ajouté à* $mnl(S)$*, avec le degré 0,5. L'extension de* $mnl(S)$ *est maintenant le sous-ensemble flou qui associe le degré 0,5 à* Lait *et aux spécialisations de* Lait*, et 0 à tous les autres éléments de* Ω*, ce qui est différent de* $ext(S)$*. On poursuit donc le parcours .*
On passe à l'élément LaitPasteurisé*. Il a le même degré 0,5, dans* $ext(mnl(S))$ *et dans* $ext(S)$*. On continue donc le parcours.*
L'élément LaitEntier *a le degré 0,5 dans* $ext(mnl(S))$ *mais le degré 1 dans* $ext(S)$*.* LaitEntier *est donc ajouté à* $mnl(S)$*, avec le degré 1. L'extension de* $mnl(S)$ *est maintenant le sous-ensemble flou qui associe le degré 1 à* LaitEntier *et aux spécialisations de* LaitEntier*, le degré 0,5 aux autres laits (*Lait, LaitPasteurisé, LaitDemiEcrémé*, etc.) et 0 aux autres éléments de* Ω*, ce qui est égal à* $ext(S)$ *: l'algorithme s'arrête.*

On obtient finalement $mnl(S)$ = {0,5/Lait, 1/Lait entier}.

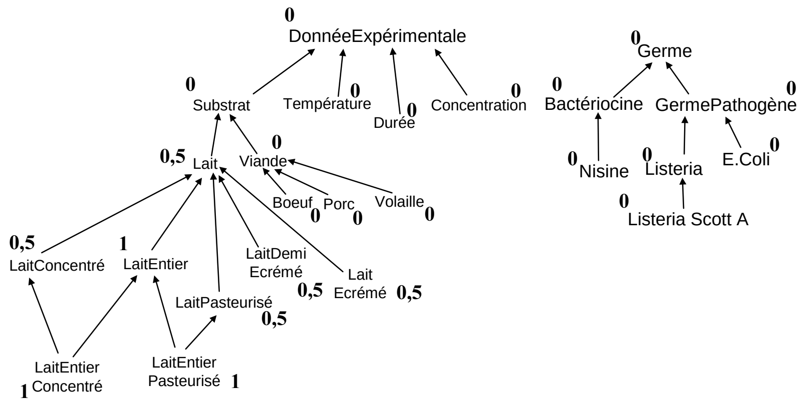

Fig. 5 – Extension du SEFO S.

4 Conclusion, perspectives

Le travail présenté dans cet article est une partie d'un projet appliqué visant à mettre en œuvre un système d'aide à la décision dans le cadre de la prévention du risque microbiologique dans les aliments. Le système MIEL (Moteur d'Interrogation ELargie) contient à l'heure actuelle des résultats provenant d'environ 700 publications scientifiques en microbiologie. Le système MIEL, présenté dans [Buche *et al.*, 2003], interroge simultanément deux sous-systèmes distincts : une base de données relationnelle autorisant la représentation de données floues (environ 90 tables) et une base de connaissances exprimées en termes de graphes conceptuels. L'interrogation se fait par le biais d'une interface utilisateur qui rend totalement transparente l'utilisation de deux bases distinctes.

Chacun des deux sous-systèmes utilise la notion de SEFO que nous avons présentée dans cet article, afin d'autoriser (i) la représentation de données imprécises (fréquentes dans le domaine de la microbiologie) mais également (ii) l'expression de préférences dans les critères de sélection d'une requête.

Plusieurs perspectives de développement de ce travail se présentent à nous à court terme. Nous travaillons à l'optimisation des algorithmes de comparaison de SEFO, qui sont fondées pour le moment sur l'extension des SEFO, et nécessitent donc autant de comparaisons que d'éléments présents dans Ω. Nous envisageons de ne stocker dans nos bases de connaissances que des SEFO minimaux (ce qui nécessitera un calcul effectué "une fois pour toutes") et de proposer autant que possible des algorithmes de comparaison uniquement fondés sur les SEFO minimaux ce qui nous permettra d'optimiser le calcul des comparaisons.

Références

[Baldwin et Martin, 1996] J.F. Baldwin et T.P. Martin. Fuzzy classes in object-oriented logic programming. In *Proceedings of the 5th IEEE International Conference* *on Fuzzy Systems (FUZZ-IEEE'96)*, pages 1358–1364, 1996.

[Bordogna *et al.*, 1994] G. Bordogna, D. Lucarella, et G. Pasi. A fuzzy object oriented data model. In *Proceedings of the 3rd IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'94)*, pages 313–318, Orlando, FL, USA, June 1994.

[Buche *et al.*, 2003] P. Buche, O. Haemmerlé, et R. Thomopoulos. Integration of heterogeneous, imprecise and incomplete data : an application to the microbiological risk assessment. In *Proceedings of the 14th International Symposium on Methodologies for Intelligent Systems, ISMIS'2003, Lecture Notes in Artificial Intelligence*, volume 2871, pages 98–107, Maebashi, Japan, October 2003. Springer.

[Cao, 1999] T.H. Cao. *Foundations of Order-Sorted Fuzzy Set Logic Programming in Predicate Logic and Conceptual Graphs*. PhD thesis, University of Queensland, Australia, 1999.

[Cross, 1996] V. Cross. Towards a unifying framework for a fuzzy object model. In *Proceedings of the 5th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'96)*, pages 85–92, 1996.

[Dubois et Prade, 1988] D. Dubois et H. Prade. *Possibility Theory - An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[George *et al.*, 1993] R. George, B. Buckles, et F. Petry. Modeling class hierarchies in the fuzzy object-oriented data model. *Fuzzy Sets and Systems*, 60 :259–272, 1993.

[Gomez *et al.*, 1995] M. Gonzalez Gomez, C. Faucher, et E. Chouraqui. Classification approchée d'instances dans un langage des schémas : Etude de l'étape d'appariement. In *Actes des Rencontres Francophones sur la Logique Floue et ses Applications (LFA'95)*, pages 200–207, Paris, France, November 1995.

[Graham et Jones, 1988] I. Graham et P.L. Jones. *Expert Systems : Knowledge, Uncertainty and Decision*. Chapmann and Hall, 1988.

[Granger, 1988] C. Granger. An application of possibility theory to object recognition. *Fuzzy Sets and Systems*, 28(3) :351–362, 1988.

[Gyseghem *et al.*, 1993] N. Van Gyseghem, R. de Caluwe, et R. Vandenberghe. Ufo : uncertainty and fuzziness in an object-oriented model. In *Proceedings of the 2nd IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'93)*, pages 773–778, San Francisco, USA, 1993.

[Itzkovich et Hawkes, 1994] I. Itzkovich et L.W. Hawkes. Fuzzy extension of inheritance hierarchies. *International Journal for Fuzzy Sets and Systems*, 62 :143–153, 1994.

[Lee *et al.*, 2001] J. Lee, J.Y. Kuo, et N.L. Xue. A note on current approaches to extending fuzzy logic to object-oriented modeling. *International Journal of Intelligent Systems*, 16(7) :807–820, July 2001.

[Lukasiewicz, 1995] T. Lukasiewicz. *Uncertain reasoning in concept lattices*, volume 946, pages 293–300. Springer-Verlag, Berlin, 1995.

[Na et Park, 1996] S. Na et S. Park. Management of fuzzy objects with fuzzy attribute values in a fuzzy object oriented data model. In H. Christiansen, H.L. Larsen, et T. Andreasen, editors, *Proceedings of the 1996 Workshop on Flexible Query-Answering Systems (FQAS'96)*, pages 19–40, Roskilde, Denmark, May 1996.

[Nepal *et al.*, 1999] S. Nepal, M.V. Ramakrishna, et J.A. Thom. A fuzzy object query language (FOQL) for image database. In *Proceedings of the 6th International Conference on Database Systems for Advanced Applications*, pages 117–124, 1999.

[Ralescu et Berenji, 1989] L.A. Ralescu et H. Berenji. Integrating structured knowledge and management of uncertainty in intelligence systems. In *Proceedings of the 4th International Workshop on Conceptual Structures, joint with IJCAI'89*, 1989.

[Rossazza *et al.*, 1997] J.P. Rossazza, D. Dubois, et H. Prade. A hierarchical model of fuzzy classes. In R. De Caluwe, editor, *Fuzzy and Uncertain Object-Oriented Databases - Concepts and Models*, volume 13 of *Advances in Fuzzy Systems - Applications and Theory*, pages 21–61. World Scientific, 1997.

[Thomopoulos, 2003] R. Thomopoulos. *Représentation et interrogation élargie de données imprécises et faiblement structurées*. PhD thesis, Institut National Agronomique Paris-Grignon, 2003.

[Torasso et Console, 1989] P. Torasso et L. Console. Approximate reasoning and prototypical knowledge. *International Journal of Approximate Reasoning*, 3(2) :157–178, 1989.

[Vignard, 1985] P. Vignard. *Un mécanisme d'exploitation à base de filtrage flou pour une représentation des connaissances centrées objets*. PhD thesis, Institut National Polytechnique de Grenoble, 1985.

[Zadeh, 1965] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8 :338–353, 1965.

[Zadeh, 1978] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1 :3–28, 1978.

Summary

Fuzzy sets can be used in order to represent imprecise values, such as an interval with fuzzy bounds. In databases, fuzzy sets can also be used to express the user's preferences in the selection criteria of a query. In knowledge representation, type hierarchies are widely used to modelize the relationship between the different objects of a given domain. We are interested in fuzzy sets whose definition domain is a hierarchy of elements, partially ordered by the "kind-of" relation. We call such a domain an ontology. We introduce the notion of fuzzy set defined on a subset of the ontology, then its developed form defined on the whole ontology. We exhibit equivalence classes of fuzzy sets defined on an ontology and we characterize them by a unique minimal fuzzy set. We conclude by an example of application in an information system dealing with the assessment of the microbiological risk in food products.