Inférence des relations sémantiques dans un réseau lexico-sémantique multilingue

Nadia Bebeshina-Clairet, Mathieu Lafourcade

Inférence des relations sémantiques dans un réseau lexico-sémantique multilingue

Nadia Bebeshina-Clairet Mathieu Lafourcade
Université de Montpellier, LIRMM, France
clairet@lirmm.fr, lafourcade@lirmm.fr

RÉSUMÉ

Les méthodes endogènes se trouvent au coeur de la construction des ressources de connaissance telles que les réseaux lexico-sémantiques. Dans le cadre de l'expérience décrite dans le présent article, nous nous focalisons sur les méthodes d'inférence des relations. Nous considérons, en particulier, les cas d'inférence des relations sémantiques et des raffinements de sens. Les différents mécanismes d'inférence des relations sémantiques y compris dans le contexte de polysémie de termes ont été décrits par Zarrouk (2015) pour le contexte monolingue. À notre connaissance, il n'existe pas de travaux concernant l'inférence des relations sémantiques et des raffinements dans le contexte d'amélioration d'une ressource multilingue.

ABSTRACT

Inferring semantic relations in a multilingual lexical semantic network.

Endogeneous methods are important for the graph based knowledge resource building. In the framework of the experiment we describe in the present paper, we focus on inferring new relations. In particular, we consider semantic relation and sense refinement inference. Various mecanisms of semantic relation inference including those with term polysemy have been detailed by Zarrouk (2015) in the monolingual context. To our knowledge, no experimental work have been done in order to elaborate such methods for a multilingual resource enhancement.

MOTS-CLÉS : réseau lexico-sémantique, ressource multilingue, inférence de relations.

KEYWORDS: lexical semantic network, inference, multilinguality.

Introduction

L'inférence endogène des relations sémantiques représente une moyen intéressant d'enrichissement des ressources lexico-sémantiques multilingues. En effet, la construction de ces-dernières est susceptible de s'appuyer sur des ressources structurées pré-existantes plus facilement disponibles pour les langues dites « riches ». Le contexte endogène offre ainsi la possibilité d'enrichir les partitions consacrées aux langues dites « peu dotées » à partir des relations sémantiques disponibles dans les partitions de langues « riches ». Cet aspect a également été souligné dans (Huang *et al.*, 2002) où l'expérience d'amorçage du WordNet du Chinois grâce à Princenton WordNet (PWN, (Fellbaum, 1998)) est décrite. Après avoir conçu un réseau lexico-sémantique multilingue pour un domaine de spécialité (alimentation), nous nous sommes tournés vers ces méthodes afin d'optimiser la construction de notre ressource.

1 État de l'Art, contexte de l'expérience

Pour les bases de connaissance factuelles telles que NELL (Carlson *et al.*, 2010), plusieurs approches d'inférence centrée sur l'équivalence des entités et des relations ont été proposées (par exemple, l'expérience de fusion de plusieurs éditions monolingues de NELL décrite dans (Hernández-González *et al.*, 2017)). Le processus d'inférence endogène a été étudié par Zarrouk (2015) et Ramadier (2016) dans le cadre d'un réseau lexico-sémantique pour le français RezoJDM (Lafourcade, 2007). Leurs méthodes reposent sur l'exploration des relations présentes dans ce réseau afin d'en proposer des nouvelles en suivant des schémas d'inférence par déduction/induction (exploitation des relations taxonomiques), par abduction (exploitation des termes jugés similaires), par raffinement. Gelbukh (2018) introduit un mécanisme d'inférence similaire à celui proposé par ces auteurs pour l'enrichissement d'une base collocationnelle et, notamment, l'inférence par abduction (inférence fondée sur la similarité sémantique acquise à partir de PWN) de nouvelles collocations. L'ancrage socioculturel de l'alimentation détermine les particularités de sa langue (présence de nombreux concepts implicites). Les méthodes d'analyse des textes de cuisine, un des domaines d'application de la ressource construite nécessitent un contexte riche qui se présente notamment comme méta-langage spécifique (approches décrites dans (Tasse & Smith, 2008), (Jermsurawong & Habash, 2015)), structures dynamiques (vecteurs d'état latents), vocabulaires ou ontologies construits à la volée, exploitation des ressources de connaissance sous forme de graphe comme celle construite dans le cadre de nos expériences.

Le réseau lexico-sémantique multilingue avec pivot interlingue ($RLSM_{PI}$) constitue le contexte de nos expériences. Le réseau contient 821 781 termes et 2 231 197 relations à l'heure où nous écrivons. Le modèle de $RLSM_{PI}$ (figure 1) s'inspire de celui du réseau lexico-sémantique RezoJDM, (Lafourcade 2007 & 2011). Il s'agit d'un graphe *orienté* (où chaque relation possède un terme source et un terme cible), *typé* et *valué* [1]. Ce graphe comporte k sous-graphes correspondant à chacune des k langues couvertes par la ressource (l'anglais, le français, le russe et l'espagnol) et un *pivot interlingue*. Pour éviter les difficultés propres à la construction d'un pivot artificiel dont la nécessité d'aligner *N* sens simultanément et dans l'impossibilité d'obtenir facilement un modèle d'alignement global sur la base d'un plongement complet du pivot interlingue, le pivot est amorcé comme un pivot naturel en utilisant l'édition anglaise de DBNary (Sérasset, 2012). Il évolue vers un pivot interlingue de façon incrémentale ce qui permet de réduire progressivement le phénomène contrastif artificiel défini par (Sérasset, 2012) comme « une perte d'information discriminatoire liée à une conceptualisation et lexicalisation qui divergent entre les langues » propre à l'utilisation d'un pivot naturel.

2 Inférence translingue des relations sémantiques

2.1 Principe et déroulement

Dans le cadre du $RLSM_{PI}$, il s'agit d'inférer les relations sémantiques dans les partitions qui en contiennent peu grâce aux relations présentes dans les partitions riches de cette ressource. L'exemple simplifié du terme russe пряник pour lequel on souhaite proposer des relations typées *r_has_part* grâce à la partition « fr » du $RLSM_{PI}$ permet d'illustrer la mise en oeuvre du processus d'inférence dans le contexte multilingue. Il existe une distinction de sens du terme *pain d'épices* en français qui peut être modélisée au niveau interlingue sous forme de deux raffinement du terme générique `in:`*gingerbread*

1. Ses arcs sont caractérisées par des poids et des annotations

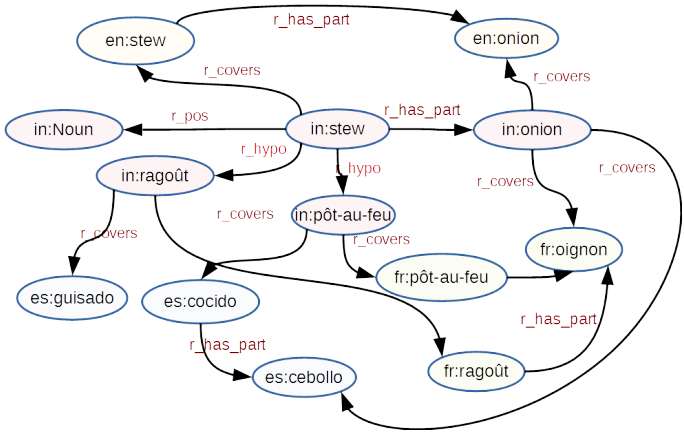

FIGURE 1 – Architecture du réseau lexico-sémantique multilingue avec pivot interlingue. Les éléments des partitions sont reliés uniquement via le pivot interlingue via la relation typée *r_covers*. Le terme interlingue est appelé *terme couvrant* et le terme lexicalisé correspondant est défini comme le *terme couvert*. Les sens d'usage des termes sont appelés *raffinements*.

qui sont `in:`*gingerbread>cake* et `in:`*gingerbread>biscuit*. L'inférence des relations sémantiques se déroule en deux temps. D'abord, elles sont inférées dans le pivot (dans notre exemple, à partir de `fr:`*pain d'épices*vers `in:`*gingerbread*) à l'aide des termes couvrants les termes-voisins de *pain d'épices* et de ses raffinements tels que `in:`*sugar* $\xrightarrow{r\_covers}$ `fr:`*sucre*, `in:`*ginger* $\xrightarrow{r\_covers}$ `fr:`*gingembre*, etc. Ensuite, les relations du pivot sont inférées dans les partitions à enrichir (dans notre exemple, la partition « ru » : `ru:` пряник $\xrightarrow{r\_has\_part}$ `ru:` сахар, etc.).

Le système d'exploitation du RLSM_{PI} se rapproche du système défini par (Ferber, 1995) comme une architecture fondée sur les *systèmes de production* [2]. L'inférence n'est pas conçue comme un processus séquentiel. Les relations exploitées dans le cadre d'inférence translingue pour identifier les termes dont les relations sémantiques génèrent les prémisses des règles d'inférence sont des relations typées *r_covers*. Elles relient le terme interlingue aux termes qu'il couvre. Par conséquent, aussi bien dans la phase ascendante (langue → pivot) que dans la phase descendante (pivot → langue), nous pouvons supposer qu'il s'agit des termes équivalents. Or, un seul et même terme couvert peut avoir plusieurs termes interlingues couvrants qui peuvent correspondre à plusieurs sens. Le cas inverse est également fréquent. Ainsi, nous considérons la relation typée *r_covers* comme une variante translingue de synonymie potentiellement incomplète. La conséquence de ce positionnement se traduit par le fait que ***en présence de plusieurs termes interlingues couvrants, le cas d'inférence est traité comme s'il s'agissait d'une inférence avec raffinements.*** Ce processus consiste à vérifier la présence des relations sémantiques entre les raffinements du terme et l'extrémité opposée de la relation à inférer. Ainsi, dans le cas du terme `en:`*cake* qui a plusieurs termes couvrants dans le pivot interlingue dont `in:`*cake>sweet food*, `in:`*cake>galette*, `in:`*cake>block* et pour lequel nous souhaitons proposer une relation typée *r_has_part* vers `in:`*ginger* nous pouvons choisir de nous assurer qu'il existe un voisin du terme interlingue (par exemple, un générique tel que `in:`*sweet food*) qui a une lien sémantique (relation ou chemin typé de longueur 2) avec « `in:`*ginger*.

2. *Un système de production est défini par la combinaison d'une base de faits (BF), d'une base de règles de production (BR) et d'un interprète, le moteur d'inférence (MI). Op.cit.* p. 137. Les règles de production ont la forme générale : « si *liste de conditions* alors *liste d'actions* ».

Dans le cadre monolingue, le mécanisme d'inférence adapté au cas des termes similaires est l'inférence par abduction. Il s'agit de sélectionner un ensemble de termes similaires à un terme T et de proposer les relations détenues par ces termes à T. On considère ainsi les demi-relations partagées par une paire de termes (relations typées de/vers un terme-voisin).

Dans le cadre multilingue et ascendant la relation à inférer est considérée comme une instance de règle d'inférence par abduction. On transforme ses termes source et cible en ensembles de termes qui peuvent contenir aussi bien des termes interlingues que lexicalisés. On recherche des « faisceaux » de relations existantes entre ces ensembles de termes. Autrement dit, on récupère tous les termes interlingues et lexicalisés pour les deux termes de la relation considérée. Puis, on explore le voisinage de l'intersection des ensembles de termes ainsi obtenus. Si la cardinalité de l'intersection entre les voisinages typés est suffisante (définie par un seuil [3]), la relation issue de la partition lexicalisée est proposée pour les termes du pivot interlingue. *Dans le cadre multilingue et descendant*, le processus d'inférence s'appuie principalement sur le filtrage logique par triangulation car de multiples termes couverts pour un seul terme couvrant sont possibles. Il s'agit de vérifier la présence d'un chemin typé entre les termes lexicalisés couverts respectivement par le terme interlingue source et par le terme interlingue cible de la relation à inférer. À terme de l'évolution du pivot, l'algorithme d'inférence des relations deviendra celui d'une simple « remontée-descente » et ne nécessitera plus de recours systématique à un mécanisme d'inférence.

2.2 Filtrage

Les inférences fondées sur les exemples génèrent un nombre important de relations-candidates qui nécessitent une procédure de filtrage afin de ne pas introduire de bruit (dû à la polysémie des termes) dans la ressource. Nous avons appliqué un *pré-filtrage par parties du discours* car celui-ci, étant une simple vérification de la présence et du poids suffisant [4] de la relation typée *r_pos* permet de réduire le temps de calcul des relations candidates. La plupart des relations sémantiques considérées dans les expériences que nous décrivons relient deux termes dont la partie de discours est « nom ». D'autres relations telles que *r_carac* (caractéristiques typique) mais aussi les relations actantielles [5] dans le cadre des langues flexionnelles telles que le russe imposent des contraintes morpho-syntaxiques qui peuvent être exploitées. Nous avons introduit un *filtrage statistique* car les relations du RLSM_{PI} peuvent être analysées en considérant leur *nombre*, leur *poids* et leur *origine*. Le *poids* w correspond à la *force d'association* et s'applique aux relations issues des ressources construites par peuplonomie comme, par exemple, RezoJDM où il est proportionnel à la fréquence avec laquelle les termes source et cible de la relation sont associés par les joueurs [6]. Le cas échéant, il est fixé par défaut. L'*origine* est une liste de chaînes de caractères qui désignent les processus ou les ressource qui ont fourni la relation. Dans le cadre du filtrage, nous avons introduit l'information sur la confiance accordée à l'origine d'une relation (ressources de connaissance, processus endogènes) sous forme d'un ensemble d'indices de confiance $\psi = \{i_1, i_2, ...i_n\}$ où $i_j \in [0; 1]$ ainsi que la cardinalité de l'ensemble des demi-relations partagées par les termes ϕ. La fonction de filtrage se calcule pour un poids positif ou négatif de la relation $w \in Z$ et $|\psi| > 0$ comme suit :

$$f(r) = \phi \times \frac{w}{Max(\psi) \times \log(|\psi|)}$$

3. Ce seuil a été empiriquement fixé à 3.
4. Le poids suffisant fixé empiriquement est un poids $w \geq 25$.
5. Relations dites « prédicat - arguments » : *r_object* (patient typique), *r_instr* (instrument typique), etc.
6. Dans le contexte des *GWAP* (*Games With a Purpose*), « jeux avec un but ».

Un score combinant la cardinalité et le maximum de ψ ainsi que le poids est utilisé lorsqu'il ne s'agit pas des termes supposés similaires. Outre le filtrage statistique, le filtrage logique par triangulation peut s'appliquer.

2.3 Expérimentation

Les expérimentations ont été conduites sur l'ensemble des relations sémantiques et sur l'ensemble des langues du RLSM_{PI}. Le chiffrage des expérimentations s'appuie sur les informations telles que le nombre de relations dans la partition d'origine (**#orig**), le nombre de relations candidates (**#cand**) soit le nombre de relations qui vérifient les prémisses d'une règle d'inférence, la productivité de l'algorithme (**prod**) qui correspond au nombre de candidats par rapport à celui des relations présentes dans la partition d'origine, le nombre de relations acceptées (**#acc**) soit le nombre de relations candidates qui permettent de fournir la conclusion et qui subsistent après la procédure de filtrage, le pourcentage des relations acceptées (**%acc**) et la précision (**pr**) évaluée manuellement sur un échantillon de 500 relations par type de relation. **rang** a été introduit pour exprimer le rapprochement de la productivité « idéale » qui se situerait autour de 100% ». La partition **fr** est la plus riche en termes du nombre de types de relations. L'inférence de certaines relations taxonomiques et méronymiques affiche une productivité assez faible due aux intersections entre les ressources déjà intégrées dans le RLSM_{PI} dont WordNet (Fellbaum, 1998)) ainsi que la nature non encore interlingue du pivot.

type	*lang*	**#orig**	**#cand**	**prod**	**#acc**	**%acc**	**rang**	**pr**
r_isa	fr	148 409	24 379	16 %	13 886	**57 %**	1	72 %
	en	314 452	50 118	16 %	20 534	**41 %**	1	93 %
r_has_part	fr	178 286	40 855	23 %	26 555	**65 %**	2	69 %
	en	39 086	10 628	27 %	3978	**37 %**	1	78 %
r_matter	fr	16 419	4 547	28 %	3728	**82 %**	1	87 %
	en	1 709	575	34 %	490	**85 %**	1	94 %
r_object	fr	14 655	54 517	372 %	10 592	**19 %**	1	94 %
	en	7 088	9 190	129 %	7 566	**91 %**	2	89 %
r_carac	fr	32 585	58 809	180 %	7 057	**12 %**	2	75 %
	en	5 474	2 436	45 %	2 094	**86 %**	1	94 %
r_manner	fr	1 873	1 423	76 %	1 109	**78 %**	2	77 %
	en	1 751	5 938	339%	1603	**27 %**	1	89 %
r_location	fr	2 499	1 821	73 %	1 071	**59 %**	1	89 %
total	-	764 286	265 236	-	100 263	-	-	-
moyenne (arith.)	-	-	-	104 %	-	**57 %**	-	80 %

TABLE 1 – Inférence ascendante des relations sémantiques. Inférence ascendante des relations sémantiques **en→pivot** reflète l'intégration massive des relations taxonomiques depuis les ressources expertes telles que WordNet, RWN (Loukachevitch *et al.*, 2016).

L'expérience d'inférence descendante a concerné les partitions les moins peuplées, **es** et **ru**. La table 2 liste les résultats pour les relations principales de ces partitions. Nous faisons état du nombre de relations d'un type donné dans la partition lexicalisée avant inférence descendante (**avant_inf**), le nombre de nouvelles relations obtenues par inférence (**inf**) et évolution. L'état d'évolution du sous-graphe « ru » et « es » permet de justifier la démarche proposée.

Les résultats obtenus pour l'inférence ascendante montrent que la productivité d'inférence dépend de celle des autres processus de peuplement du RLSM_{PI} tels que l'intégration des relations sémantiques

type	l	#bef	#inf	#aft	ev
r_isa	ru	46 827	7 036	53 863	+14 %
	es	36 807	268 040	304 847	+828 %
r_has_p.	ru	65 772	3 682	69 454	+5 %
	es	10 166	56 883	67 049	+559 %
r_mat.	ru	5190	4230	9 420	+81 %
	es	4013	7 351	7 764	183 %
r_man.	ru	1 265	1 655	2 920	+131 %
	es	1 753	9 507	11 260	+542 %
r_loc.	ru	640	621	1 261	+97 %
	es	90	567	657	+630 %
by lang.	ru	119 694	17 224	136 918	+14 %
	es	52 739	342 348	395 087	+649 %
Totaux (moyenne)	-	172 433	359 572	532 005	**+208 %**

TABLE 2 – Descending inference of semantic relations. According to the size of the available corpora and external resources, the positive inference impact may vary.

à partir des ressources structurées. La précision obtenue par filtrage logique montre l'efficacité de ces filtrages dans le contexte multilingue. Cependant, la question de complexité peut être évoquée. Les prémisses d'une règle qui définissent le parcours à mettre en œuvre dans le cadre du filtrage logique sont destinées à réduire le nombre de relations à tester (et la complexité) en ne parcourant que certains types de relations. Cette opération de recherche des relations doit être exécutée m fois. Ainsi, la complexité globale de filtrage logique (sans spécification de type de filtrage) serait de $O(m \times n^2)$. La **complexité moyenne** correspondrait dans ce cas au degré moyen du RLSM_{PI} : $d_{av} = 4 \Rightarrow O(16 \times m)$.

3 Raffinements, cas particulier d'inférence endogène

Dans un réseau lexico-sémantique, la relation de raffinement permet de modéliser les sens d'usage d'un terme polysémique. Les raffinements correspondent à des cliques maximales (calcul) ou à des contributions des joueurs (jeu d'acquisition lexicale). Ils peuvent être nommés (glosés) ou non. Ainsi, pour le terme *baguette*, nous avons le sens « pain » par opposition aux autres sens (« encadrement », « bâton », « baguette magique »). Le raffinement glosé correspondant à ce sens est *baguette>pain*. Ainsi, nous avons la structure suivante dans notre ressource : $baguette \xrightarrow{r\_refinement} baguette{>}pain \xrightarrow{r\_glose} pain$. Un raffinement glosé peut être raffiné ou raffiné et glosé à son tour. Dans le cas d'une ressource qui possède déjà des relations de raffinement, il est possible de se baser sur les raffinements existants dans une de ses partitions pour inférer les sens dans les partitions qui ne contiennent pas ce type de relation. L'inférence des raffinements glosés se déroule en deux temps. Dans le cadre du ***schéma ascendant*** tous les sens présents dans les partitions lexicalisées sous forme de raffinements glosés se retrouvent dans le pivot interlingue. Ce pivot peut alors être considéré comme union raisonnée des parties monolingues de la ressource. Ce processus mis en œuvre à partir des partitions « en » et « fr », a permis d'atteindre le taux de raffinement du pivot interlingue de 30%.

Lorsque pour un terme il n'existe pas de raffinement glosé interlingue, mais ce terme possède plusieurs termes couvrant, le ***schéma de raffinement descendant en absence de raffinement nommé*** peut être appliqué. On considère les différents termes couvrants comme termes potentiellement liés à la glose. Initialement, les sens sont provisoirement étiquetés (ex. : étiquettes des termes couvrants). Ensuite,

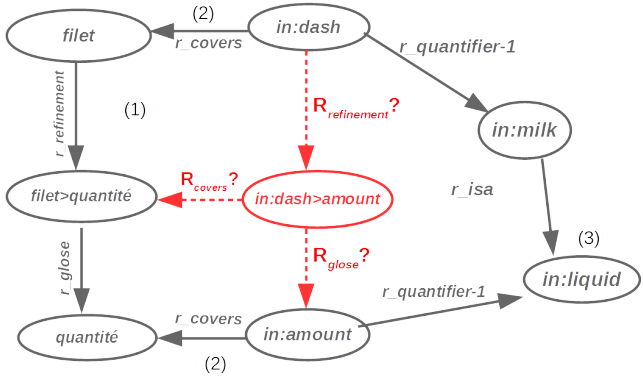

FIGURE 2 – Pour un terme raffiné (*filet*), son raffinement glosé (*filet>quantité*), la glose (*quantité*), le terme couvrant de terme raffiné (*in :dash*) et le terme couvrant de la glose (*in :amount*) : s'il existe une relation sémantique ou un chemin typé entre les termes couvrant le terme raffiné et la glose, un raffinement glosé interlingue peut être proposé. L'inférence implique la proposition de trois relations et la création d'un nouveau terme (raffinement interlingue).

cible	fr>int	en>int	intersection	total
int	8 558	33 930	254	31 752

TABLE 3 – Raffinements glosés interlingues acquis en appliquant le schéma ascendant.

on valide la distinction proposée en regroupant les sens redondants. L'étape finale du processus est le choix guidé par la sémantique de la glose appropriée dans la langue traitée. À ce jour ce schéma a permis de proposer 2 535 raffinements en russe tandis que 1 800 raffinements ont pu être obtenus pour cette langue grâce à l'inférence des raffinements glosés. Les mécanismes d'inférence proposés sont conçus pour améliorer le RLSM_{PI} de façon continue.

Conclusion

Dans le présent article, nous avons proposé une méthode d'inférence endogène des relations sémantiques. Nous avons ainsi exploité dans un cadre multilingue les méthodes conçues et précédemment déployées dans le cadre monolingue. Nous avons également proposé une méthode d'inférence des raffinements de sens à partir des raffinements existants et des contrastes observés dans la ressource. Cette méthode donne un rôle central au pivot évolutif amorcé en tant que pivot naturel et conçu pour devenir incrémentalement un pivot interlingue. Cette vision de pivot ainsi que l'inférence endogène des relations sémantiques confèrent à l'expérience que nous avons décrite un intérêt pour la construction des ressources qui impliquent des langues dites « peu dotées » pour lesquelles il n'existe que peu de ressources structurées et sémantiquement riches.

Références

BOLLACKER K., EVANS C., PARITOSH P., STURGE T. & TAYLOR J. (2008). Freebase : a collaboratively created graph database for structuring human knowledge. In *In SIGMOD Conference*, p. 1247–1250.

CARLSON A., BETTERIDGE J., KISIEL B., SETTLES B., JR. E. R. H. & MITCHELL T. M. (2010). Toward an architecture for never-ending language learning. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010*.

CORDIER A., DUFOUR-LUSSIER V., LIEBER J., NAUER E., BADRA F., COJAN J., GAILLARD E., INFANTE-BLANCO L., MOLLI P., NAPOLI A. & SKAF-MOLLI H. (2014). *Taaable : A Case-Based System for Personalized Cooking*, In S. MONTANI & L. C. JAIN, Eds., *Successful Case-based Reasoning Applications-2*, p. 121–162. Springer Berlin Heidelberg : Berlin, Heidelberg.

DONG Z., DONG Q. & HAO C. (2010). Hownet and its computation of meaning. In *Proceedings of the 23rd International Conference on Computational Linguistics : Demonstrations*, COLING '10, p. 53–56, Stroudsburg, PA, USA : Association for Computational Linguistics.

FELLBAUM C. (1998). *WordNet An Electronic Lexical Database*. Cambridge, MA ; London : The MIT Press.

FERBER J. (1995). *Les systèmes multi-agents : vers une intelligence collective*. Paris : InterEditions.

GELBUKH A. F. (2018). Inferences for enrichment of collocation databases by means of semantic relations. *Computación y Sistemas*, **22**(1).

HERNÁNDEZ-GONZÁLEZ J., HRUSCHKA JR. E. R. & MITCHELL T. M. (2017). Merging knowledge bases in different languages. In *Proceedings of TextGraphs-11 : the Workshop on Graph-based Methods for Natural Language Processing*, p. 21–29, Vancouver, Canada : Association for Computational Linguistics.

HUA W., WANG Z., WANG H., ZHENG K. & ZHOU X. (2015). Short text understanding through lexical-semantic analysis. In *International Conference on Data Engineering (ICDE)*.

HUANG C.-R., TSENG I.-J. E. & TSAI D. B. (2002). Translating lexical semantic relations : The first step towards multilingual wordnets. In *COLING-02 : SEMANET : Building and Using Semantic Networks*.

JERMSURAWONG J. & HABASH N. (2015). Predicting the structure of cooking recipes. In L. MÀRQUEZ, C. CALLISON-BURCH, J. SU, D. PIGHIN & Y. MARTON, Eds., *EMNLP*, p. 781–786 : The Association for Computational Linguistics.

LAFOURCADE M. (2007). Making people play for Lexical Acquisition with the JeuxDeMots prototype. In *SNLP'07 : 7th International Symposium on Natural Language Processing*, p.7̃, Pattaya, Chonburi, Thailand.

LAFOURCADE M. (2011). *Lexicon and semantic analysis of texts - structures, acquisition, computation and games with words*. Habilitation à diriger des recherches, Université Montpellier II - Sciences et Techniques du Languedoc.

LOUKACHEVITCH N., LASHEVICH G., GERASIMOVA A., IVANOV V. & DOBROV B. (2016). Creating russian wordnet by conversion. In *Dialog-2016*, p. 405–415, Moscow.

MÜLLER G. & BERGMANN R. (2015). Cookingcake : A framework for the adaptation of cooking recipes represented as workflows. In *Workshop Proceedings from The Twenty-Third International Conference on Case-Based Reasoning (ICCBR 2015), Frankfurt, Germany, September 28-30, 2015.*, p. 221–232.

RAMADIER L. (2016). *Indexation et apprentissage de termes et de relations à partir de comptes rendus de radiologie*. PhD thesis. Thèse de doctorat dirigée par Lafourcade, Mathieu Informatique Montpellier 2016.

SÉRASSET G. (2012). Dbnary : Wiktionary as a lmf based multilingual rdf network. In *LREC*.

SPEER R. & HAVASI C. (2012). Representing general relational knowledge in conceptnet 5.

TASSE D. & SMITH N. A. (2008). Sour cream :toward semantic processing of recipes. *T.R. CMU-LTI-08-005*, p.9̃.