DOI: 10.15439/2019F267
ISSN 2300-5963 ACSIS, Vol. 18

Multilingual Knowledge Base Completion by Cross-lingual Semantic Relation Inference

Nadia Bebeshina-Clairet
LIRMM, University of Montpellier,
860 rue de St Priest 34095 Montpellier
France
Email: clairet@lirmm.fr

Mathieu Lafourcade
LIRMM, University of Montpellier,
860 rue de St Priest 34095 Montpellier
France
Email: lafourcade@lirmm.fr

***Résumé*—In the present paper, we propose a simple endogenous method for enhancing a multilingual knowledge base through the cross-lingual semantic relation inference. It can be run on multilingual resources prior to semantic representation learning. Multilingual knowledge bases may integrate preexisting structured resources available for resource-rich languages. We aim at performing cross-lingual inference on them to improve the low resource language by creating semantic relationships.**

I. Introduction

HIGHLY structured knowledge bases (KBs) such as lexical semantic networks (LSNs) contain various connectivity patterns that can be learned as node features using dedicated frameworks i.e. node2vec [10]. However, semantic relations are often unequally distributed over such knowledge resources. Some of the language partitions may benefit from integrating structured resources which are more easily available for "rich" languages i.e. Princenton WordNet (PWN) [7], ConceptNet [21], YAGO [22] for English, RezoJDM [12] for French.

Unlike large factual KBs, the LSNs explicitly represent taxonomic relations ($hypernymy$, $meronymy$), predicate-argument relations, typical characteristic, and possibly other relation types ($entailment$, causal relations) as well as polysemy (through synsets, refinements). A meta-information related to the relation weight (power of association i.e. in RezoJDM), a confidence score linked to the origin of the relations integrated from some existing resources (i.e. in ConceptNet), annotation as well as negatively weighted relations that explicitly model "noise" (relation considered as false, i.e. RezoJDM) may be attached to the LSN relations in the framework of a particular model. Thus, automated semi-structured approaches to the multilingual LSN building represent a hard task : when available, models may vary from one language to another. For instance, the modeling of meronymy relations may reflect different vision of this relation type. In ConceptNet, the meronymy is represented as a *hasPart* relation. PWN introduces the distinction between part (*mammal*→*mouth*), substance (*wine*→*alcohol*), and member (*bee*→*bee colony*) meronymy. RezoJDM model includes all the relations covered by PWN and adds the holonymy relation (*cutlet*→*beef*).

II. State of the Art

Cross-lingual relationship inference benefits from active research efforts. State of the art inference in KBs include rule-based and machine learning approaches. In the framework of the large KBs such as NELL [3], several approaches centered on the equivalence between entities and relationships have been introduced. For instance, authors in [11] describe the experience of merging several monolingual editions of NELL. Authors in [14] detail the statistical relational learning on knowledge graphs (KGs) and point out the importance of type constraints and transitivity as well as other statistical patterns or regularities, "which are not universally true but nevertheless have useful predictive power". Similar to [24], they base their method mainly on large scale KBs such as Nell [3], KnowItAll [6], YAGO [22] or DeepDive [20].

The endogenous rule-based inference process has been studied by Zarrouk (2015) and Ramadier (2016) in the framework of RezoJDM, the LSN for French. Their methods rely on the relationships and relationship meta-information that are already present in this LSN in order to propose the new ones following one of the following schemes : deduction, induction (which benefit from taxonomic relations), abduction (exploiting semantic similarity), and inference by refinement. Gelbukh (2018) introduced a comparable inference mechanism to enrich a collocationnal knowledge base by suggesting new collocations through the inference by abduction (where semantic similarity is calculated on the basis of PWN [7]).

KBs completion can be made using embedding strategies where latent spaces allow modeling candidate facts as resulting from latent factors. RESCAL [15] and TransE [1] propose such approaches. RESCAL performs collective learning using the latent components of the tensor factorization. In other words, the entity neighborhood is used to predict an unknown relation between this entity e_1 and some other entity e_2 knowing that some other entities similar to e_1 (in terms of their neighborhood) are connected to e_2 through the relation type t. The TransE method models relationships by interpreting them as translations in the embedding space and relies on low-dimensional embeddings of the entities. This system associates some vector depending on the relationship type to the vector of this relationship *tail* (source). This allows learning only one low-dimensional vector for each entity and each relationship.

Markov Random Fields (MRF) based approach to KBs completion has been proposed with first-order logic representation through a *Markov logic network* in [17]. MRF-based approach with probabilistic logic representation has been introduced in [2]. Path ranking approaches based on random walk i.e. [5] are also being explored.

III. Context

To conduct the experiment, we built a multilingual lexical-semantic network (MLSN) inspired by the RezoJDM for the food domain. At the time of our writing, the MLSN contains 821 781 nodes (terms) and 2 231 197 arcs (relationships). It is a directed, typed, and valuated graph. The MLSN k sub-graphs correspond to each of the k languages (English, French, Russian, and Spanish). A specific sub-graph fulfills the role of the *interlingual pivot*. The MLSN nodes may correspond to one of the following types : lexical items (i.e. *garlic*) ; interlingual items (also called *covering terms*) that are not necessarily labeled in a human readable way ; relational items (i.e. relationship reifications such as *salad[r_has_part]garlic*) ; category items modeling categories, parts of speech or other morpho-syntactic features (i.e. *Verb :Present*, *Noun :AccusativeCase*).

During the MLSN set up and building, the hypothesis introduced in [16] concerning the "non separation" between general and domain specific knowledge has been considered. This hypothesis states that during the semantic analysis of domain specific texts that relies on background knowledge, the presence of general common sense knowledge in addition to the domain specific knowledge improves the performance of such analysis. Thus, general commonsense knowledge information has been integrated into our resource from the existing LSNs i.e. PWN, ConceptNet etc. This integration has been "guided" by the domain specific comparable corpora (96 083 cooking recipes, about 8 300 000 words) as its vocabulary has been used as a seed.

A relation $r \in R$ is a sextuplet $r =< s, t, type, v, l_s, l_t >$ where s and t correspond respectively to the source and the target term of the relation. The relation $type$ is a typical relation type. It may model different features such as taxonomic and part-whole relations (*r_isa, r_hypo, r_has_part, r_matter, r_holo*), possible predicate-argument relations (typical object *r_object*, location *r_location*, instrument *r_instr* of an action), "modifier" relations (typical characteristic *r_carac*, typical manner *r_manner*) and more [1]. The relationship valuation v corresponds to the characteristics of the relation which are its weight, confidence score, and annotation. The relation weight may be negative in order to model noise and keep the information about erroneous relations easy to access programmatically so they could not affect the inference processes. The confidence score is a score attributed to a particular data origin (external resource, inference process). In practice, this feature is an array as different origins may provide the same relation. The confidence information is provided as an argument to the function that maps from some external knowledge resource to the MLSN. In case of relation calculated by an inference process, it corresponds to the precision evaluated on a sample of candidate relations returned by this process. To *annotate* a relation we add a complementary information that allows qualifying this relation. The labels l_s and l_t correspond to the language (sub-graph) labels.

As it has been difficult to set up the pivot using a multilingual embedding (joining multiple spaces, one per language) as well as to avoid pairwise alignment based on combinatorial criteria, the pivot has been started as a natural one using the English edition of DBNary [19]. It incrementally evolves based on sense alignments (obtained through external resources or by inference) between the languages of the MLSN to become interlingual. It can be considered as a union of word senses lexicalized or identified in the languages covered by the MLSN. Such progressive pivot building allows reducing the artificial contrast phenomena defined by [18] as a discriminatory information loss linked to the divergent conceptualization and lexicalization observed in different languages. Even though we assume the pivot as being interlingual, it is still close to a natural one. As a result, the inference mechanisms we detail are suitable for the architecture by transfer.

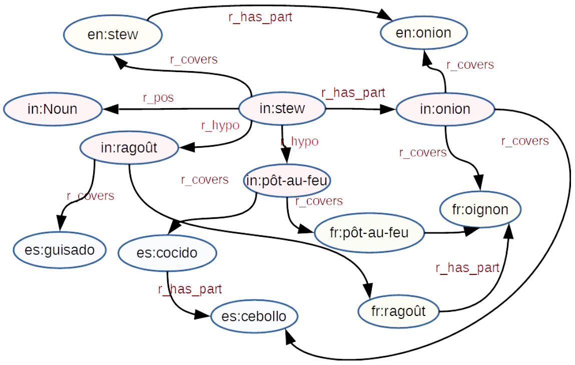

FIGURE 1. Pivoted MLSN architecture. *in* precedes interlingual terms.

The MLSN has been set up and populated through term and relationship extraction from corpora using state of the art pattern based techniques and lexical semantic patterns which add semantic constraints i.e. lookup for semantic relationships in the MLSN to the syntactic pattern) ; by integrating weakly structured domain specific information (glossaries, vocabulary lists, nutritional composition of food items)as well as structured lexical semantic resources (such as RezoJDM, PWN, ConceptNet) and inference mechanisms.

The table I shows that the global inference impact is higher in the context of the relationship types that are hard to yield by integration from structured resources or identified in corpora. The relationship typed *refinement*(*r_raff*) connecting terms to their senses will be described in section IV. The pivot coverage (how many terms from a given MLSN sub-graph are connected to the pivot ?), the presence of semantic relationships and sense

[1] We also introduced more specific relation types such as *r_entailment*, *r_cause*, *r_telic_role*, *r_incompatible*, *r_before*, *r_after* etc.

TABLE I
MLSN RELATIONSHIP ACQUISITION.

R_{type}	**corpus**	**integ.**	**before inf**	**inf**	**prod**
r_isa	67 894	544 632	612 526	27 546	4%
r_hypo	688	797 783	798 471	41 053	5%
r_has_part	662 737	172 287	835 024	48 015	6%
r_matter	606	35 597	36 203	1 893	5%
r_holo	224	67 081	67 305	51 360	76%
r_object	42 280	29 262	71 542	15 512	22%
r_carac	8 300	69 236	77 536	9 521	12%
r_manner	2 854	3 250	6 104	250	4%
r_location	2086	3 573	5 659	146	3%
r_instr.	58	2 738	2 796	402	14%
r_refinement	221	29 441	29 662	182 135	614%
Overall	788 344	1 754 484	2 542 828	377 816	15%

refinements in a MLSN sub-graph also determine the success of a relationship inference process. Semantic information is easier to obtain from monolingual external resources. Thus, the exogenous data and semantic relationship acquisition are mostly monolingual. As a result, some terms may not be covered by the pivot. As the semantic relationships are used by the inference mechanism for logical filtering, when a MLSN sub-graph has numerous semantic relationships the inference precision is higher.

IV. CROSS-LINGUAL SEMANTIC RELATION INFERENCE

Principle - In this section we detail the inference of new semantic relations in one lexicalized part of the MLSN from the ones existing in another MLSN part (sub-graph). In a pivoted resource, the relations are first inferred into the pivot (ascending inference). Second, they are inferred in other sub-graphs (descending inference). In transfer-based resources where lexicalized sub-graphs are directly connected to each other, the inference process would directly apply to the source and target languages and rely on translation links between those. Thus, the proposed inference process is independent from the architecture of the resource (transfer or pivot based). It also may be considered as independent from the expressiveness of the multilingual resource as we define for and test it on a very expressive MLSN with numerous relation types. **Monolingual context -** In the monolingual context, the mechanisms of inference by deduction, induction, abduction, and inference with sense refinements apply. These processes have been described in [25]. In case of transitive semantic relations (i.e. hyperonymy, hyponymy), **deduction** and **induction** can be implemented. These inference schemes propose to a term some relevant relations detained by its hypernyms or hyponyms based on the transitivity of these taxonomic relations. For (nearly) synonyms, the **abduction** procedure is chosen. The abduction yields a set of terms similar to the term T then proposes the neighborhood relations detained by these terms to T. In order to calculate the similar terms more finely, in addition to calculating Jaccard similarity score, weighted Jaccard such as in [12] or some other similarity measure, we consider semi-relations (Typed ingoing and outgoing relations from/to a neighbor) shared by a pair of similar (synonymous) terms. **Inference with sense refinements** exploits the sense refinement of polysemous terms. When the senses (we also call *refinements*) are modeled, it is possible to verify whether they are semantically related to the opposite term of the relation to be inferred.

To give an example, we may consider the french term *soupe* and its refinements $\{$*soupe>potage, soupe>neige, soupe>repas*$\}$("soup>broth", "soup>melted snow", "soup>meal") and a new candidate relation we want to infer (relation obtained either by deduction, induction or abduction) *soup*$\xrightarrow{r\_carac}$ *chaud*(hot). In order to automatically accept such relation, we may check if one of soup refinements is semantically connected to *chaud* or one of its refinements : *soupe>potage*$\xrightarrow{r\_isa}$*liquide* & *liquide*$\xrightarrow{r\_carac}$*chaud.*

Multilingual context - In the context of the cross-lingual semantic relation inference, we use the *r_covers* relations to identify the semantic relations that correspond to the premises of the inference rules.

The relations typed *r_covers* link an interlingual term to the lexicalized terms that it covers. We may suppose that during the ascending ($language \rightarrow pivot$) and the descending ($pivot \rightarrow language$) processes we deal with the equivalent terms. Due to the discrepancies between languages and to the fact that our recent interlingual pivot is still close to the natural one, one lexicalized term may have multiple covering terms and *vice versa*. Therefore, we consider the *r_covers* relation as a cross-lingual variant of a (possibly) incomplete synonymy. Given that, the case of inference that applies can be either abduction or inference by refinement. For the abduction case in the ascending multilingual context, the relation to be inferred is considered as an abduction rule instance. We transform its source and target terms into the sets which may contain interlingual and lexicalized terms. Then, we explore the neighborhood of the intersection between the obtained sets. If the intersection between the typed semi-relations is sufficient (we empirically set the threshold to 3), the relation from the lexicalized subgraph is proposed for the terms in the interlingual pivot (and vice versa while performing a descending inference).

The case of polysemy is processed as if it was an "inference with sense refinements" case. It checks by triangulation the presence of semantic relations between the "refinements" of a term (the different covering or covered terms) and the opposite term of the relation to be inferred. A simplified example of the Russian term `pryanik` for which we are looking to infer relations typed *r_has_part* thanks to the "fr" MLSN subgraph illustrates the inference mechanism. The distinction between the sense refinements of *pain d'épices* in French can be modeled at the interlingual level as two interligual refinements of the interlingual term `in:`*gingerbread* that are `in:`*gingerbread>cake* and `in:`*gingerbread>biscuit.* The inference is a twofold process. The relationships from the "fr" subgraph are inferred into the pivot using the interlingual terms that cover the *pain d'épices* neighbors : such as `in:`*sugar*$\xrightarrow{r\_covers}$ `fr:`*sucre*, `in:`*ginger*$\xrightarrow{r\_covers}$ `fr:`*gingembre*, etc. Then, the relations are inferred from the pivot to the "ru" subgraph. As `pryanik` in Russian culinary tradition has the soft cookie texture (this information is available from semantic relations of `pryanik` and from translation links where `pryanik` is linked to both refinements of *pain d'épices*, the distinction observed in French is not relevant for Russian. Thus, the descending inference process proposes candidate relations of the general interlingual term `in:`*gingerbread* to `pryanik`. As the general term detains the relationships of its refinements, `pryanik` yields all the relationships of *pain d'épices* that can be represented on the interlingual level and persist after logical/statistical filtering.

The abduction scheme generates a lot of candidate relationships. Therefore, a filtering strategy significantly improves the precision. First, we apply part-of-speech pre-filtering can be used depending on relation types. For instance, in the case of the relation typed r_carac (typical characteristic) the source term must be a noun whereas the target term must be an adjective (i.e. *cake*$\xrightarrow{r\_carac}$*sweet*). Second, we use the statistical filtering as the relations of the MLSN can be analyzed in terms of their number, weight, and origin. The weight w corresponds to the crowd-sourced weight or to the default weight. Similar to ConceptNet, we introduced the information regarding the confidence given to the structured resource from which a given relation has possibly been integrated or to the endogenous inference. Thus we attach the *origin* information to the relationships. It took the form of an array of strings (naming the different processes that provided the relation) to which we associate an array of confidence scores $\psi = \{i_1, i_2, ...i_n\}$ where $i_j \in [0;1]$. The size of the set of semi-relations shared by the terms ϕ is also taken into account. For the positively or negatively weighted relation the filtering function is calculated as follows for $w \in \mathbb{Z}$ and $|\psi| > 0$:

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

In a mature MLSN, the relation inference algorithm becomes a simple lifting and descending algorithm where no significant filtering to be applied.

Experiment - We tested our approach on all the semantic relations and languages present in the MLSN. The table II lists the results of the descending inference process. The results are presented in terms of number of relations in the source partition (**#orig**), number of candidate relations (**#cand**), number of accepted relations after filtering (**#acc**), productivity of the algorithm (**prod**), acceptance rate (**%acc**, the percentage candidate relationships that verify the inference rule premises and conclusion and subsist after filtering), and precision (**pr**) which has been manually evaluated on a sample of 500 accepted relations (per type). This type of manual evaluation has been chosen due to the difficulty to find a well balanced reference for evaluation. As we integrated the main LSNs for the languages covered by the MLSN, we presumably infer the relationships that are not explicitly represented in such structured resources. The range **r** has been introduced to express how close a given process is situated to the "gold" productivity (100%). Indeed such "gold" productivity would mean that the sense based alignment is sufficient for a given term. The table II lists the results for the main semantic relations of the Russian and Spanish sub-graphs and details the evolution of the number of semantic relations.

TABLE II
DESCENDING INFERENCE OF SEMANTIC RELATIONS.

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%
TOTALS	-	172 433	359 572	532 005	**+208%**

The logical filtering concerns only a subset of relation types to be checked m times (according to the branching factor of the term). Thus, the global complexity of the logical filtering would be $O(m \times n^2)$. La *average complexity* would correspond to the average degree observed in the MLSN at the time of our writing : $d_{av} = 4 \Rightarrow O(16 \times m)$.

Towards the Sense-based Alignment - The MLSN refinement relations allow modeling the "use" senses of a term. The refinement corresponds to maximal cliques (calculus) or to the contributions (GWAP). For the french term *baguette*, we distinguish the sense "bread" as opposed to "direction", "stick", and "magic wand". The glossed refinement corresponds to this sense is *baguette>pain*. Thus, we have the following structure in the MLSN : *baguette*$\xrightarrow{r\_raff}$*baguette>pain*$\xrightarrow{r\_glose}$*pain*. A glossed refinement may be itself refined and glossed. In the case of a resource that already possesses refinement relations, it is possible to infer some cross-lingual new refinements from the existing ones. The 30% refinement rate of the MLSN pivot has been obtained using this process.

When the term has multiple covering terms, the descending inference pattern can be applied. We consider that the covering terms are potentially linked to the gloss. First, we temporarily label the potential senses using the labels of the covering terms. Second, we group the redundant senses and choose the gloss. the recently started experiment with this pattern allowed producing the first batch of 2 535 sense refinements in Russian whereas 1 800 refinements have been yielded for this language using the glossed refinement technique.

V. CONCLUSION

We introduced a simple endogenous method for cross-lingual semantic relation inference to improve structured KBs such as MLSN. Given a certain coverage in terms of translation links, it allows enhancing the under-resourced parts of a lexical semantic resource from the rich ones. Even though they benefit from translation resources and tools, some "rare" languages are not covered by any rich lexical semantic resource. To some extent, the method is beneficial for domain specific MLSNs. It allows rich semantic modeling which provides a semantically structured representation for the fine grained semantic analysis (including word sense disambiguation) and statistical representation learning.

References

[1] Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States.*, pages 2787–2795, 2013.

[2] Matthias Bröcheler, Lilyana Mihalkova, and Lise Getoor. Probabilistic similarity logic. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, UAI'10, pages 73–82, Arlington, Virginia, United States, 2010. AUAI Press.

[3] Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell. 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*, 2010.

[4] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2d knowledge graph embeddings. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 1811–1818, 2018.

[5] Xin Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 601–610, New York, NY, USA, 2014. ACM.

[6] Oren Etzioni, Michael Cafarella, Doug Downey, Ana-Maria Popescu, Tal Shaked, Stephen Soderland, Daniel S. Weld, and Alexander Yates. Unsupervised named-entity extraction from the web: An experimental study. *Artificial Intelligence*, 165(1):91 – 134, 2005.

[7] Christiane Fellbaum. *WordNet An Electronic Lexical Database*. The MIT Press, Cambridge, MA ; London, 1998.

[8] J. Ferber. *Les systèmes multi-agents: vers une intelligence collective*. InterEditions, Paris, 1995.

[9] Alexander F. Gelbukh. Inferences for enrichment of collocation databases by means of semantic relations. *Computación y Sistemas*, 22(1), 2018.

[10] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. *KDD : proceedings. International Conference on Knowledge Discovery & Data Mining*, 2016:855–864, 2016.

[15] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. A three-way model for collective learning on multi-relational data. In *Proceedings of the 28th International Conference on Machine Learning*, ICML'11, pages 809–816, USA, 2011. Omnipress.

[11] Jerónimo Hernández-González, Estevam R. Hruschka Jr., and Tom M. Mitchell. Merging knowledge bases in different languages. In *Proceedings of TextGraphs-11: the Workshop on Graph-based Methods for Natural Language Processing*, pages 21–29, Vancouver, Canada, August 2017. Association for Computational Linguistics.

[12] Mathieu Lafourcade. *Lexique et analyse sémantique de textes - structures, acquisitions,calculs, et jeux de mots. (Lexicon and semantic analysis of texts - structures, acquisition, computation and games with words)*. Montpellier, 2011.

[13] Mathieu Lafourcade and Lionel Ramadier. Semantic RelationExtraction with Semantic Patterns: Experiment on Radiology Report. In *LREC 2016 Conference on Language Resources and Evaluation*, volume 10th of *LREC 2016 Proceedings*, Portorož, Slovenia, May 2016.

[14] Maximilian Nickel, Kevin Murphy, Volker Tresp, and Evgeniy Gabrilovich. A review of relational machine learning for knowledge graphs: From multi-relational link prediction to automated knowledge graph construction. *CoRR*, abs/1503.00759, 2015.

[16] Lionel Ramadier. *Indexation and learning of terms and relations from reports of radiology*. Theses, Université de Montpellier, November 2016.

[17] Matthew Richardson and Pedro Domingos. Markov logic networks. *Mach. Learn.*, 62(1-2):107–136, February 2006.

[18] Gilles Sérasset. Dbnary: Wiktionary as a lmf based multilingual rdf network. In *LREC*, 2012.

[19] Gilles Sérasset. DBnary: Wiktionary as a Lemon-Based Multilingual Lexical Resource in RDF. *Semantic Web – Interoperability, Usability, Applicability*, pages –, 2014. To appear.

[20] Jaeho Shin, Sen Wu, Feiran Wang, Christopher De Sa, Ce Zhang, and Christopher Ré. Incremental knowledge base construction using deepdive. *Proc. VLDB Endow.*, 8(11):1310–1321, July 2015.

[21] Robert Speer and Catherine Havasi. Representing general relational knowledge in conceptnet 5. In *LREC Proceedings*, 2012.

[22] Fabian Suchanek, Gjergji M Kasneci, and Gerhard M Weikum. Yago: A Core of Semantic KnowledgeUnifying WordNet and Wikipedia. In *16th international conference on World Wide Web*, Proceedings of the 16th international conference on World Wide Web, pages 697 – 697, Banff, Canada, May 2007.

[23] Kristina Toutanova and Danqi Chen. Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality*, pages 57–66, Beijing, China, July 2015. Association for Computational Linguistics.

[24] Quan Wang, Bin Wang, and Li Guo. Knowledge base completion using embeddings and rules. In *Proceedings of the 24th International Conference on Artificial Intelligence*, IJCAI'15, pages 1859–1865. AAAI Press, 2015.

[25] Manel Zarrouk, Mathieu Lafourcade, and Alain Joubert. Inference and Reconciliation in a Crowdsourced Lexical-Semantic Network. In *CICLING: International Conference on Intelligent Text Processing and Computational Linguistics*, number 14th, Samos, Greece, March 2013.