LooPings: a Look at Semantic Similarities

Adama Sow[1] and Jean-Rémi Bourguet[2,3]

[1] Département Génie Informatique et Télécommunications
Ecole Polytechnique de Thiès (EPT) – Senegal
adama.sow.4@ulaval.ca
[2] Dipartimento di Scienze Politiche ed Ingegneria dell'Informazione
Università degli Studi di Sassari (UNISS) – Italy
[3] Núcleo de Estudos em Modelagem Conceitual e Ontologias
Federal University of Esprito Santo (UFES) – Brazil
jrbourguet@inf.ufes.br

Abstract. Semantic similarities is a cross-field research in Natural Language Processing and Ontologies with some possible fallouts in Artificial Intelligence. Formerly, similarities were computed following a syntactical treatment to support case-based reasoning. Textual similarities are now guided by semantic machineries, offering various ways to compute relatedness measures. In this paper, we present both a logical and a visual framework aiming to reason with them. For that reason, we introduced $\mathcal{FLH}^{\pm}$, a fragment of description logic underpinning the well-known lexical database Wordnet. We illustrated this framework with the path length relatedness, one of the historical similarity measures occurring in a taxonomy. The core of our framework orchestrates the computation of similarity scores supported by REVERB, STANFORD CORENLP and WORDNET:SIMILARITY APIs and interfaces global similarities in graphical way by positioning them on segments. We also depicted some experimental results to confront our computational framework with some empirical data.

Keywords: Semantic Similarity, Ontology, Description logic, Natural language processing, Interface, Empirical data

1 Introduction

Reasoning with similarities is seen as one of the crucial steps in Artificial Intelligence. Turing, in his paper *Computing Machinery and Intelligence* [1], suggested that a machine having passed the so-called test, should appear as a human 70% of the time after five minutes of conversation. From Joseph Weizenbaum, and his seminal proposal *Eliza* [2] in 1966, until the last generation of programs *Jabberwacky* and *Cleverbot* developed by Rollo Carpenter [3] during the last decade, the treatment of similarities attracted a lot of interest to tackle this issue. Formerly, one classical way was to deal with Natural Language Processing (NLP), focusing on syntactic similarities through a case-based reasoning machinery.

These last years, the Semantic Web [4] has given complementary resources in terms of knowledge bases, offering new standards to deal with lexical databases. The semantic dimension of similarities is now guided by well-established and moderated repositories of knowledge including ontological layers (see [5]), but few attempts was performed to combine the classical NLP-based technologies together with semantic web infrastructures in order to analyze the limits of such interactions (see for example [6]). Moreover, if some methods were also proposed to compute global scores of similarities between two phrases based on local semantic similarities of their components (see [7,8]), we remark that the usual way to output a set of similarities between a phrase and a set of phrases is an ordered list of items (e.g. the popular search engines), and very few approaches were proposed in the literature in order to screen the similarity scores in different ways (see [9,10]).

In this paper, we present a logical and visual framework to represent and reason with textual relatedness measures guided by ontologies. In order to bridge the gap between natural language and ontologies, we introduced a fragment of Description Logic (DL), underpinning the well-known lexical database WORDNET [11]. With this fragment, we can properly redefine some relatedness measures historically used in taxonomies. To gain place, we only introduced here the path length relatedness measure denoted `plr`. Nevertheless, other historical similarity measures considering the maximum depth in taxonomy (see `lch` in [12]), the depth of the least common subsumer (see `wup` in [13]) or the supported information content (see `jcn` in [14] and `lin` in [15]) have also been integrated in this framework.

After some preliminaries to introduce our logical framework in section 2, we describe in section 3 both our computation and an interface to screen the similarities between a phrase and a set of phrases. Finally, in section 4 we present an experimental validation to confront the theoretical similarities with some experimental ones.

2 Preliminaries

DL is a well-known family of formal knowledge representation models. Semantic languages used on the web to share knowledge (e.g. RDFS [16], OWL [17] and OWL2 [18]) have all some direct underpinning logics that are fragments of DL. The core interpretation of a DL in first order logic was given by Baader and Nutt in [19] as follows:

Definition 1. *Let $\mathbf{C}$ the set all the atomic concepts and $\mathbf{R}$ the set all the atomic roles, an interpretation $\mathcal{I}$ is an ordered pair $(\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ such that:*
- $\Delta^{\mathcal{I}}$ is the domain, i.e. a non-empty set of individuals,
- $\cdot^{\mathcal{I}}$ is the interpretation function which maps:
- each atomic concept $A \in \mathbf{C}$ to a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$,
- each atomic role $r \in \mathbf{R}$ to a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

In this paper, we chose to introduce $\mathcal{FLH}^{\pm}$ defined in Definition 2 as an extension of $\mathcal{AL}$ [20] using transitivity and hierarchy for roles but removing negation, intersection, limited existential and value restrictions.

Definition 2. *Let an interpretation $\mathcal{I}$, $\{A, B\} \subseteq \mathbf{C}$ and $\{r, s\} \subseteq \mathbf{R}$*

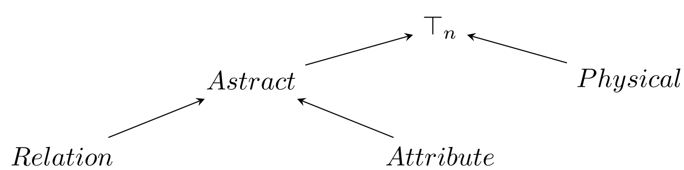
$$
\begin{array}{rcll}
\top^{\mathcal{I}} & = & \Delta^{\mathcal{I}} & \texttt{top concept} \\
\bot^{\mathcal{I}} & = & \emptyset & \texttt{bottom concept} \\
(A \sqsubseteq B)^{\mathcal{I}} & = & A^{\mathcal{I}} \subseteq B^{\mathcal{I}} & \texttt{inclusion axiom} \\
(r \sqsubseteq s)^{\mathcal{I}} & = & r^{\mathcal{I}} \subseteq s^{\mathcal{I}} & \texttt{role hierarchy} \\
(r^{+})^{\mathcal{I}} & = & (a,b) \in r^{\mathcal{I}} \wedge (b,c) \in r^{\mathcal{I}} \rightarrow (a,c) \in r^{\mathcal{I}} & \texttt{transitivity} \\
(\texttt{dom}(r) \equiv A)^{\mathcal{I}} & = & \forall a, b \in \Delta^{\mathcal{I}}.a \in A^{\mathcal{I}} \cup (a,b) \notin r^{\mathcal{I}} & \texttt{domain} \\
(\texttt{ran}(r) \equiv B)^{\mathcal{I}} & = & \forall a, b \in \Delta^{\mathcal{I}}.b \in B^{\mathcal{I}} \cup (a,b) \notin r^{\mathcal{I}} & \texttt{range}
\end{array}
$$

As redefined in [19], a Terminology Box or a TBox in DL is a finite set of axioms, with no symbolic name (equality whose left-hand side is an atomic concept), which is defined more than once. Princeton's WordNet is a lexical database for the English language [21]. A decade after the creation of this database, van Assem [22] proposed a conversion of WordNet in OWL. Example 1 presents a sample of TBox underpinning WordNet in OWL introducing particularly the concept Synset[4] and the role h[5].

Example 1 (Sample of WordNet *TBox).*
Synset $\sqsubseteq \top$, $\texttt{dom}(\mathsf{h}) \equiv$ Synset, $\texttt{ran}(\mathsf{h}) \equiv$ Synset, h^{+}.

In the terms of Baader and Nuts [19], an Assertional Box or an ABox is a specific state of affairs of an application domain in terms of concepts and roles. Example 2 depicts a graph representing a sample of ABox underpinning WordNet in OWL where $\top_n$ represents the individual root of all the individual nouns present in WordNet, the edge represents synsets (e.g. Synset$(Astract)$) and the directed arcs represent the hyponym relation assertions between two individuals (e.g. $\mathsf{h}(Attribute, Abstract)$). Note that the WordNet Abox is partitioned in different set, dealing with different parts of speech (nouns, verbs, adjectives and adverbs).

Example 2 (Sample of WordNet *ABox).*



The idea of computing relatedness measures between concepts occurring in a taxonomy is an old issue (see [23]). Few approaches (see [24,25]) have described the common relatedness measures through a DL formalism. According to Rada [26], a path length is founded on a node-counting scheme concerning the smallest specified role counting between two individuals. Given an interpretation and two individuals, we redefined the shortest path length function denoted $\texttt{L}$ as follows:

[4] Short for "Synonym set" to call a non-empty set of synonyms.
[5] Short for "hyponym" to call the well-known "hyponymy" relation.

Definition 3. *Let an interpretation $\mathcal{I}$, $\{u_k, v_l\} \in \Delta^{\mathcal{I}}$ and $\mathsf{h} \in \mathbf{R}$*

$$\texttt{L}(u_k, v_l) = min(2n\text{-}(k{+}l))\ \textit{s.t.}\ \forall(i,j) \in [\![k,n]\!] \times [\![l,n]\!].\ \mathsf{h}(u_i, u_{i+1}) \wedge \mathsf{h}(v_j, v_{j+1}) \wedge u_n{=}v_n$$

Note that if a practitioner deals with the transitive closure of the WORDNET ABox, a min max operator may be used due to the transitivity of the relation h. Example 3 lists the shortest path length between individual nouns introduced in Example 2. We denote $\top_v$ the individual root of all the verbs. The ABox being partitioned, note that no verb is involved in a hyponym relation with a noun (and vice-versa).

Example 3 (Shortest Path Length).

$\texttt{L}(Relation, Abstract) = 1$ $\qquad$ $\texttt{L}(Relation, Attribute) = 2$
$\texttt{L}(Physical, Physical) = 0$ $\qquad$ $\texttt{L}(Physical, \top_v) \mapsto \infty$

As reintroduced by Pedersen [27], the path length based relatedness score is equal to the inverse of the shortest path length between two concepts. Naturally, it is inversely proportional to the number of nodes along the shortest path between the synsets. The path length based relatedness denoted `plr` is defined as follows:

Definition 4. *Let an interpretation $\mathcal{I}$, $\{u, v\} \in \Delta^{\mathcal{I}}$*

$$\texttt{plr}(u, v) = \left(\frac{1}{1+\texttt{L}(u,v)}\right)$$

If no path exists (e.g. between a verb and a noun) we consider that L tends to ∞, and then the path length based relatedness score approaches 0. The highest possible score occurs when the two synsets are the same, in which case the score is 1.

Example 4 (Path Length based Relatedness).

$\texttt{plr}(Relation, Abstract) = 1/2$ $\qquad$ $\texttt{plr}(Relation, Attribute) = 1/3$
$\texttt{plr}(Physical, Physical) = 1$ $\qquad$ $\texttt{plr}(Physical, \top_v) \mapsto 0$

In the next section, we present our method to compute similarity scores between two phrases based on the local semantic similarities of their components (e.g. `plr`), moreover we depict a way to screen them through a graphical user interface.

3 Computational Framework

The finality of our framework is to screen similarity scores between one phrase and a set of phrases. We chose to investigate the way transforming a phrase in a set of triples before performing the computation of similarities. We selected the API REVERB, an information extractor for massive corpora (working without pre-specified vocabulary). In [28], REVERB was judged with a extraction precision of 0,8 or higher in at least 30% of the time, substantially outperforming others extractors like TEXTRUNNER (see [29]) and WOE (see [30]). Formally and for the following, we can see a phrase p as a set of triples $\{t_1, \ldots, t_n\}$ where $t_l = \langle t_l^1, t_l^2, t_l^3 \rangle$ with $t_l^k \in \Delta^{\mathcal{I}}$. Thereafter, we define the global score between two phrases as the maximum of the averages of similarities between the components from the triples (average of similarities between subjects, objects, and complements). Moreover, we arbitrary avoid to take into account two kinds of scores: scores of 0, it is the case when at least one entry is not present in the lexical base of WORDNET (see section 2), and the score of 1 (similar strings) only if the individuals in comparison are present in a stop words set denoted θ.

Definition 5. *Let θ and two phrases p, q, the Score of similarity is defined as follows:*

$$\mathtt{S}(p,q) = \max_{t_i,t_j}\Big(avg(\{\mathtt{s}(t_i^k,t_j^k) \mid \mathtt{s}(t_i^k,t_j^k) \neq 0 \wedge (\mathtt{s}(t_i^k,t_j^k) \neq 1 \vee t_i^k, t_j^k \notin \theta)\})\Big)$$

with $t_i \in p$, $t_j \in q$, $k \in [\![1,3]\!]$, $\mathtt{s} \in \{\mathtt{plr}, \mathtt{lch}, \mathtt{wup}, \mathtt{jcn}, \mathtt{lin}\}$ and $\mathtt{S}$ the upper case of $\mathtt{s}$.

To support this computation, we used the STANFORD CORENLP API [31] to perform the lemmatization of the components from the triples. The similarities calculus of the lemme were performed by WORDNET:SIMILARITY (developed by Pedersen et al. in [32] and redesigned in Java by Shima [33]).

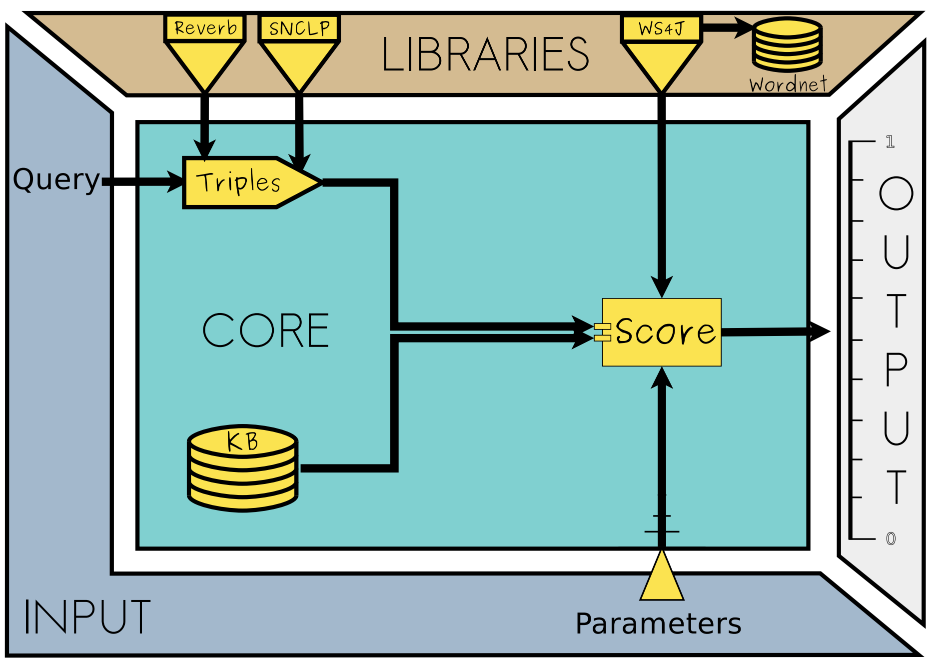


Fig. 1: The system description of `LooPings`

All this treatment is included in the framework **`LooPings`**[6] (**L**exical and **o**ntological **o**bservations to **P**lot **ing**athering **s**imilarities). Looking at the system description of Fig. 1, **`LooPings`** is decomposed in several modules:

[6] The framework **`LooPings`** is available at http://nemo.inf.ufes.br/?p=1185.

`INPUT` takes a phrase and some parameters (e.g. local measure, stop words list, etc.).
`LIBRARIES` comprise REVERB, STANFORD CORENLP and WS4J API (bundled with WORDNET).
`CORE` manages the computation.
`OUTPUT` is devoted to screen and to plot the similarity scores.

Fig. 2: An interface of `LooPings`

One of the most challenging steps is the connection between the NLP APIs REVERB and WS4J. The main aspect that made this compatibility a little immature is the fact that REVERB can output expressions (several raw words) while WS4J accepts only a single element as input (a lemme or an expression of lemmes linked themselves in one string by underscores). Then, we had to perform a treatment from our raw phrases, following some basic heuristics, as described through the example below:

▷ **What can we learn from the neural networks of C.elegans to understand human brains?**
The outputted triple is ⟨*we,learn from,the neural networks of C.elegans*⟩.

1. Lemmatization:
   ⟨*we,learn from,the neural network of C.elegans*⟩
2. Chunking:
   ⟨*we,learn from,the neural_network C.elegans*⟩
3. Stop words removing:
   ⟨*we, learn, neural_network C.elegans*⟩
4. Cutting:
   ⟨*we, learn, neural_network*⟩;⟨*we, learn, C.elegans*⟩

Note that during the step of the removal, stop words were never removed from the triples when it involved an empty set for one component. Fig. 2 presents the interface of LooPings, where practitioners can visualize and confront the semantic similarities. Note that this framework is oriented to integrate queries and SPARQL [34] interpretations allowing a possible integration of other semantic web technologies. The similarity scores are represented on a segment [0,1].

The last section is dedicated to the confrontation between the theoretical scores with some experimental ones.

4 Experimental Validation

In order to deal with real queries, we used STACK EXCHANGE API allowing us to extract from a website some structured data (in JSON format) about different question-and-answer themes. The total number of available queries for the Cognitive Science section of STACK EXCHANGE website was 1245, we decided to extract the 1200 first queries (w.r.t. the extraction order), from each of them we stored ids (from 1 to 4697) and queries. Formally, a query referenced by an id is denoted q_{id}, moreover we call series S_j a set of queries: $S_j \subseteq \{q_{id} | 1 \leq id \leq 4697\}$. A tricky aspect was the fact that ReVerb outputted nothing in 70.75 percent of queries. So, we decided to divide, the 1200 queries in 12 series of 100 queries, finally giving a median series of 28 queries. Due to space limitation, we will depict in this article only two series (A and B).

Once the extraction was performed, we created an on-line semantic recall based experience. For each series, we designed a witness query. We attached a questionnaire to each series and sent them by e-mail to mailing lists of PhD students. The requirement for participating was to be proficient in English. Questionnaires were designed as follows:

1. Instruction to read carefully a witness query.
2. Instruction to read the series of queries and to check 1, 2 or 3 queries among them, (at least 1, at most 3), the most similar, for the user, to the witness queries.
3. Instruction to partially preorder the selected query(ies).

We succeeded in obtaining 50 volunteers. We gave accumulated marks for queries in each series w.r.t the preorders given by the volunteers. For example if a volunteer selected only one query q_i in the series, the accumulated mark for q_i was increased by 6, the remaining possible situations are listed below:

$$
\begin{array}{cc}
q_i \succ q_j \rightsquigarrow q_i(+4); q_j(+2) & q_i \sim q_j \rightsquigarrow q_i, q_j(+3) \\
q_i \succ q_j \succ q_k \rightsquigarrow q_i(+3); q_j(+2); q_k(+1) & q_i \sim q_j \sim q_k \rightsquigarrow q_i, q_j, q_k(+2) \\
q_i \succ q_j \sim q_k \rightsquigarrow q_i(+3); q_j, q_k(+1.5) & q_i \sim q_j \succ q_k \rightsquigarrow q_i, q_j(+2.5); q_k(+1)
\end{array}
$$

The maximum mark is translated to an experimental score of 1, after what all the other marks have been transposed in experimental scores by cross-multiplications. As depicted in Fig. 3, we took the PLR score as a reference in order to observe how the other scores behaved following it. Thereafter, we describe what is globally remarkable in the behavior of the theoretical similarities.

– The higher the `PLR` is, the lower is the difference between scores founded on path lengths (`LCH`, `WUP`) and scores founded on information contents (`JCN`, `LIN`).
– `WUP` and `LIN` react in the same way in the case of a sudden increase or drop.
– `LIN` and `JCN` scores have a behavior of exponential shape in all the series.

Nevertheless, there are some limitations for this framework. We relate for example the case of a saturation. Series B is remarkable by the fact that 5 queries are outputted with the maximal score.

▷ **Where could I find psychological experiments tools?**
The outputted triples were ⟨*I,find,psychological*⟩*;* ⟨*I,find,experiment*⟩*;* ⟨*I,find,tool*⟩

Here, the systematic presence of two stop words “I” and “find” make that all the similarity scores are 1 iff one of the words “psychological”, “experiment” or “tool” is present in one of the extracted triples.

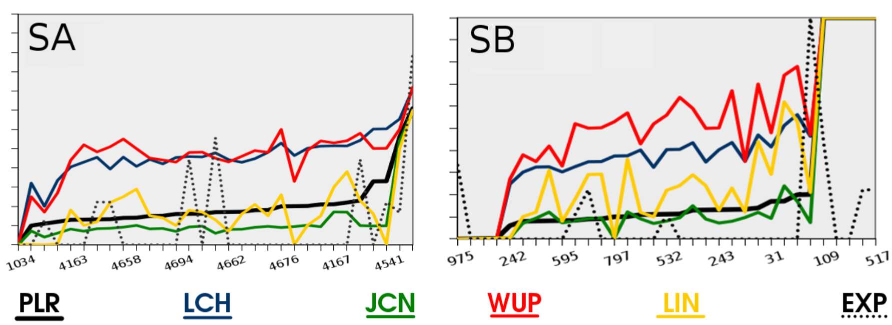


Fig. 3: Some plots of `LooPings`

This experience showed some natural limits concerning our approach. The main issue seems to be the automatic semantic annotation step (REVERB) and the matching step with WORDNET individuals (through WORDNET:SIMILARITY).

5 Conclusion

In this paper, we presented both a computational and a visual framework to support semantic similarities guided by ontologies. The first contribution was to redefine some taxonomy-based relatedness measures (e.g. path length relatedness) in a DL fragment ($\mathcal{FLH}^{\pm}$) we introduced.
The core of our framework orchestrates the computation of similarity scores supported by REVERB, STANFORD CORENLP and WORDNET:SIMILARITY APIs and interfaces results in graphical way through segments. Moreover, we plotted the behavior of our computations by designing a semantic recall-based experience to confront empirical similarities with the theoretical ones. Some research perspectives for this work would

be to extend **LooPings** in two ways. The first concerns the core of our framework by interpreting scores founded on other relations like for instance the mereology in WORDNET. The second is to integrate other repositories to support the similarities or other frameworks to compute similarities directly on semantic web languages (e.g. using the technology of QAKiS [35]).

References

1. Alan Mathison Turing. Computing machinery and intelligence. *Mind*, 59(236):433–460, 1950.
2. Joseph Weizenbaum. Eliza;a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1):36–45, 1966.
3. Rollo Carpenter and Jonathan Freeman. Computing machinery and the individual: The Personal Turing Test. Technical report, Jabberwacky, 2005.
4. Tim Berners-Lee, James Hendler, Ora Lassila, et al. The semantic web. *Scientific american*, 284(5):28–37, 2001.
5. Andreas Hotho, Steffen Staab, and Gerd Stumme. Ontologies improve text document clustering. In *Proceedings of the 2003 IEEE International Conference on Data Mining*, pages 541–544. IEEE Computer Society, November 19-22, 2003.
6. Viviana Mascardi, Angela Locoro, and Fabrizio Larosa. Exploiting prolog and nlp techniques for matching ontologies and for repairing correspondences. In *Proceedings of the 24th Convegno Italiano di Logica Computazionale*, pages 10–24, 2009.
7. Yuhua Li, David McLean, Zuhair Bandar, James O'Shea, and Keeley A. Crockett. Sentence similarity based on semantic nets and corpus statistics. *IEEE Trans. Knowl. Data Eng.*, 18(8):1138–1150, 2006.
8. Palakorn Achananuparp, Xiaohua Hu, and Xiajiong Shen. The evaluation of sentence similarity measures. In *Proceedings of the 10th International Conference on Data Warehousing and Knowledge Discovery*, DaWaK '08, pages 305–316. Springer-Verlag, 2008.
9. Oren Zamir and Oren Etzioni. Grouper: A dynamic clustering interface to web search results. *Comput. Netw.*, 31(11-16):1361–1374, May 1999.
10. Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. 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 3111–3119, 2013.
11. George A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38:39–41, 1995.
12. Claudia Leacock and Martin Chodorow. Combining local context and wordnet similarity for word sense identification. *WordNet: An Electronic Lexical Database*, pages 265–283, 1998.
13. Zhibiao Wu and Martha Stone Palmer. Verb semantics and lexical selection. In James Pustejovsky, editor, *32nd Annual Meeting of the Association for Computational Linguistics, 27-30 June 1994, New Mexico State University, Las Cruces, New Mexico, USA, Proceedings*, pages 133–138. Morgan Kaufmann Publishers / ACL, 1994.
14. Jay Jiang and David Conrath. Semantic similarity based on corpus statistics and lexical taxonomy. In *Proc. of the Int'l. Conf. on Research in Computational Linguistics*, pages 19–33, 1997.
15. Dekang Lin. An information-theoretic definition of similarity. In *Proceedings of the Fifteenth International Conference on Machine Learning*, ICML '98, pages 296–304, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc.

16. RDFS Working Group. Resource description framework schema (rdfs). `www.w3.org/TR/rdf-schema/`.
17. OWL Working Group. Web ontology language (owl). `www.w3.org/TR/owl-features/`.
18. OWL2 Working Group. Web ontology language 2(owl2). `www.w3.org/TR/owl2-overview/`.
19. Franz Baader and Werner Nutt. Basic description logics. In *Description Logic Handbook*, pages 43–95. Cambridge University Press, 2003.
20. Klaus Schild. A correspondence theory for terminological logics: Preliminary report. In *In Proc. of IJCAI-91*, pages 466–471, 1991.
21. Christiane Fellbaum. *WordNet: an electronic lexical database*. MIT Press, 1998.
22. Mark van Assem, Aldo Gangemi, and Guus Schreiber. Rdf/owl representation of wordnet. Technical report, W3C, 2008.
23. Amos Tversky. Features of similarity. *Psychological Review*, 84(4):327–352, 1977.
24. Peter C. Weinstein and William P. Birmingham. Comparing concepts in differentiated ontologies. In *Proceedings of the Twelfth Workshop on Knowledge Acquisition, Modeling and Management (KAW'99)*, 1999.
25. Alexander Borgida, Thomas J. Walsh, and Haym Hirsh. Towards measuring similarity in description logics. In *Proceedings of the 2005 International Workshop on Description Logics (DL2005), Edinburgh, Scotland, UK, July 26-28, 2005*, volume 147. CEUR-WS.org, 2005.
26. Roy Rada, Hafedh Mili, Ellen Bicknell, and Maria Blettner. Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(1):17–30, 1989.
27. Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. Wordnet::similarity: Measuring the relatedness of concepts. In *Demonstration Papers at HLT-NAACL 2004*, HLT-NAACL–Demonstrations '04, pages 38–41, Stroudsburg, PA, USA, 2004. Association for Computational Linguistics.
28. Anthony Fader, Stephen Soderland, and Oren Etzioni. Identifying relations for open information extraction. In *Proceedings of the Conference of Empirical Methods in Natural Language Processing (EMNLP '11)*, Edinburgh, Scotland, UK, July 27-31 2011.
29. Michele Banko, Michael J Cafarella, Stephen Soderland, Matt Broadhead, and Oren Etzioni. Open information extraction from the web. In *IN IJCAI*, pages 2670–2676, 2007.
30. Fei Wu and Daniel S. Weld. Open information extraction using wikipedia. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10, pages 118–127. Association for Computational Linguistics, 2010.
31. Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60, 2014.
32. Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. Wordnet: Similarity - measuring the relatedness of concepts. In *Proceedings of the 19th National Conference on Artifical Intelligence*, AAAI'04, pages 1024–1025. AAAI Press, 2004.
33. Hideki Shima. Wordnet similarity for java (ws4j). `https://code.google.com/archive/p/ws4j/`.
34. SPARQL Working Group. Sparql 1.1 overview. `www.w3.org/TR/sparql11-overview/`.
35. Elena Cabrio, Julien Cojan, Alessio Palmero Aprosio, Bernardo Magnini, Alberto Lavelli, and Fabien Gandon. Qakis: an open domain QA system based on relational patterns. In Birte Glimm and David Huynh, editors, *Proceedings of the ISWC 2012 Posters & Demonstrations Track, Boston, USA, November 11-15, 2012*, volume 914 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2012.