Computing Semantic Relatedness of GermaNet Concepts

Iryna Gurevych and Hendrik Niederlich

EML Research gGmbH, Schloss-Wolfsbrunnenweg 31c, 69118 Heidelberg

{gurevych,niederlich at eml-research.de}

Type of the paper GermaNet-Workshop

Computing Semantic Relatedness of GermaNet Concepts

Iryna Gurevych and Hendrik Niederlich

We present a system designed to compute the semantic relatedness between a pair of GermaNet concepts (word senses). Five different metrics have been implemented. Three of them are information content based and incorporate the knowledge of the GermaNet hierarchy (Resnik 1995, Jiang & Conrath 1997, Lin 1998). Two metrics constitute the application of a Lesk algorithm (Lesk 1986) to artificial conceptual glosses generated from GermaNet. We show that four metrics correlate very well with a set of human judgments of semantic relatedness. We conclude with implementation issues and a description of a graphical user interface to compute the semantic relatedness of German words.

Wir stellen ein System zur Auswertung der semantischen Beziehung für die Paare der GermaNet-Konzepte (Wortlesarten) vor. Fünf verschiedene Maße wurden implementiert. Drei von ihnen basieren auf dem Informationsgehalt und beziehen das Wissen über die GermaNet-Hierarchie mit ein (Resnik 1995, Jiang & Conrath 1997, Lin 1998). Zwei Maße stellen eine Anwendung des Lesk-Algorithmus (Lesk 1986) auf künstliche Definitionen der Konzepte dar, die automatisch aus dem GermaNet generiert wurden. Wir zeigen, dass vier Methoden eine hohe Korrelation mit den menschlichen Bewertungen der semantischen Beziehung aufzeigen. Es folgen einige Anmerkungen zur Implementierung sowie die Beschreibung einer graphischen Benutzerschnittstelle zur Auswertung der semantischen Beziehung zwischen Wortbedeutungen in GermaNet.

1. Introduction

Semantic relatedness metrics specify to what degree the meanings of two words (word senses) are related to each other. While the words *Glas* and *Becher* display a fairly close semantic relatedness, the relation between *Glas* and *Juwel* is less close according to human judgments. The system should, then, allow to determine the degree of semantic relatedness between e.g. the concepts *Glas*, *Becher* and *Juwel*.

As soon as the information about semantic relatedness becomes available, it can be employed in various natural language processing (NLP) tasks.

Patwardhan et al. (2003) employ semantic similarity metrics to measure the semantic similarity between all word senses of a given word pair, and thus disambiguate them in a given context. McCarthy et al. (2004) combined the use of a thesaurus automatically acquired from raw textual corpora and WordNet based similarity metrics to find predominant word senses in untagged text. Gurevych & Strube (2004) apply WordNet semantic similarity measures to assess the relevance of utterances given a specific dialogue and automatically construct spoken dialogue summaries. Hirst & Budanitsky (2004) investigate the usefulness of semantic similarity on the problem of spelling correction, where real-world spelling errors are detected and corrected automatically on open-class words. Further applications of semantic relatedness can be designed in the areas of document classification and information retrieval.

2. Motivation

Computing semantic relatedness of concepts requires a broad-coverage knowledge source, such as GermaNet (Kunze 2004). In this resource, nouns, verbs and adjectives are structured into hierarchies of *is-a* relations. Also, it encodes information about additional lexical and semantic relations, e.g. hypernymy, meronymy, antonymy, etc.

Ever since large-scale computational resources, such as e.g. Roget's Thesaurus and WordNet (Fellbaum 1998) have become available, there has been extensive research on semantic relatedness for the English language. This was facilitated by the availability of the software to embed semantic relatedness metrics into computational applications, e.g. WordNet::QueryData (Rennie 2002) and WordNet::Similarity Package (Pedersen 2003).

To our knowledge, explorations of semantic relatedness for the German language are almost non-existent, especially with GermaNet word senses as a basis. On the one hand, there are generally fewer large-scale computational knowledge sources like dictionaries for German. On the other hand, if they are available as is the case with GermaNet, off-the-shelf tools to access it are lacking. As a consequence, building NLP applications utilizing semantic relatedness is not as straightforward as it is the case for English and involves a considerable programming effort.

Our work aims to bridge the gap between the metrics of semantic relatedness and GermaNet, thus making the knowledge in GermaNet accessible to NLP applications. Direct re-implementation of semantic relatedness metrics developed for WordNet on the basis of GermaNet turns out to be a non-trivial task. While sharing many design principles with WordNet, GermaNet also displays several divergent features (Kunze & Lemnitzer 2002). Some of them, such as a small number of textual conceptual definitions (glosses) in GermaNet, have crucial consequences for some metrics. For example, a Lesk metric (Lesk 1986) is based on word overlaps in conceptual glosses.

In the paper, we present an architecture of the system for computing semantic relatedness of the German word senses and discuss theoretical as well as technical issues in adapting individual metrics to GermaNet. We will touch upon related work in this field and conclude with evaluation results and our plans for further experiments and applications of semantic relatedness metrics.

3. GermaNet API

We evaluated the C-library distributed together with GermaNet V4.0 and the XML-encoded version of GermaNet (Lemnitzer & Kunze 2002). We built upon the latter as it makes the system portable across platforms. The XML version of GermaNet is parsed with the help of standard XML tools (e.g. Xerces http://xml.apache.org/) to create a JAVA object representing GermaNet. This object exists in two versions, the original one, where the information can be accessed by words, and the stemmed one, where the information can be accessed by word stems. Furthermore, we implement a range of JAVA-based methods for querying the data. These methods are organized around the notions of word sense and synset.

On the word sense (WS) level, we have the following methods:[1] *getAntonyms()*: retrieves all antonyms of a given WS; *getArtificial()*: indicates whether a WS is an artificial concept; *getGrapheme()*: gets a graphemic representation of a WS; *getParticipleOf()*: retrieves the WS of the verb that the word sense is a participle of; *getPartOfSpeech()*: gets the part of speech associated with a WS; *getPertonym()*: gives the WS that the word sense is derived from; *getProperName()*: indicates whether the WS is a proper name; *getSense()*: yields the sense number of a WS in GermaNet; *getStyle()*: indicates if the WS is stylistically marked; *getSynset()*: returns the corresponding synset; *toString()*: yields a string representation of a WS.

[1] We list the most important methods for accessing and manipulation of the GermaNet data.

On the synset level, we can retrieve the following information: *getAssociations()*: returns all associations; *getCausations()*: gets the effects that a given synset is a cause of; *getEntailments()*: yields synsets that entail a given synset; *getHolonyms(), getHyponyms(), getHypernyms(), getMeronyms()*: return all holonyms, hyponyms, hypernyms, and meronyms respectively; *getPartOfSpeech()*: returns the part of speech associated with word senses of a synset; *getWordSenses()*: returns all word senses constituting the synset; *toString()*: yields a string representation of a synset.

The metrics of semantic relatedness can, then, be designed employing this API. They are implemented as classes which use the API-methods on an instance of the GermaNet object (see Section 6).

4. Metrics of semantic relatedness

Depending on the type of knowledge from GermaNet and external knowledge, which particular metrics employ, they can be clustered into separate groups. Common to all our experimental metrics is the fact that they utilize the knowledge about the *is-a* hierarchy in GermaNet. This observation leads us to define all of them to be path-based (where the *path* is not limited to an *is-a* relation). Then, we differentiate between two kinds of path-based metrics. The first one translates the knowledge from GermaNet to some sort of free textual representations, which makes it accessible to standard language processing algorithms (e.g. dictionary based metrics). The second one makes use of the hierarchy in a more traditional manner, i.e. in order to determine the lowest common superordinate for a pair of word senses. It makes the application of information content based algorithms (metrics) possible. They additionally require some external corpus evidence to compute the information content of synsets. The following description is based on this classification of metrics.

Two metrics among the implemented ones constitute the application of a Lesk algorithm to the case of GermaNet. To compensate for the lack of glosses, we automatically generate artificial glosses (we call them *pseudo-glosses*), which stand as proxies for textual definitions of concepts. Three remaining metrics are information content based and incorporate the knowledge of the GermaNet hierarchy (Resnik 1995, Jiang & Conrad 1997, Lin 1998). All of the metrics take two words as input and produce a numeric score reflecting the degree of semantic relatedness between them. Each individual metric is called for all possible sense combinations of the two words. The sense combination of the highest semantic relatedness is, then, chosen.

4.1 Dictionary based metrics

While WordNet can be seen simultaneously as a conceptual network and as a machine-readable dictionary, GermaNet is rather a conceptual network. Only a small number of fairly short textual definitions of concepts are available, which makes it difficult to apply standard dictionary based metrics to GermaNet. To overcome this problem, we try to compensate for a small number of glosses in GermaNet (or any other conceptual hierarchy - the approach is not restricted to GermaNet indeed). We generate a textual definition of a given concept automatically, given the structure of the knowledge base. This idea is remarkable because it allows to apply a range of traditional text based algorithms, e.g. the Lesk metric under circumstances when textual definitions *proper* are not available (this is the case for many world languages, for which not so many resources are available as for English). We demonstrate that the results of semantic relatedness metrics operating on automatically generated glosses correlate very well with human judgments.

In generating artificial glosses, our goal is to represent the words which would normally be present in the glosses of a traditional dictionary. This is done by selecting particular concepts from the conceptual network and combining them together into a unique concept representation. A pseudo-gloss is therefore not a coherent piece of text (gloss). Rather it stands as a proxy including the most important concepts which we expect to appear in a real gloss.

The experiments reported were done with two different system configurations: one for radial glosses (all lexical-semantic relations of a given concept are taken into account, except hyponymy), and the other one for hypernym glosses (only hypernymy relation is considered). For the purpose of this paper, "Lesk (radial)" is an algorithm using the 1st system configuration and "Lesk (hypernym)" is an algorithm using the 2nd one. We experimented with a set of different parameters for generating pseudo-glosses, whose description and results will be published elsewhere. Table 1 presents examples of pseudo-glosses generated according to the system configurations mentioned above.[2]

[2] Note that the synset *Laienprediger* is subject to multiple inheritance in GermaNet, *Prediger* and *Laie*. Therefore, both concepts appear in generated pseudo-glosses.

	Bursche	**Bruder**
Radial glosses	*1. junger Mensch, Erwachsener, Bursche, Bub, Junge, Knabe, Bube, Kind, Jüngling*	*1. Bruder, Geschwister, Mitmensch, Familie, Verwandter* *2. LaienpredigerIn, Fachkraft, unausgebildeter Mensch, Geistlicher, Prediger, ausgebildeter Mensch, Bruder, Berufstätiger, Laie, Laienprediger* *3. christlicher Sakralbau, Kloster, Geistlicher, Ordensangehöriger, Mönch, Bruder, Mönchskloster, Berufstätiger, Laie, Glaubensgemeinschaft, Orden*
Hypernym glosses	*1. Bursche, Junge, Kind*	*1. Bruder* *2. unausgebildeter Mensch, Geistlicher, Prediger, Laie, Laienprediger* *3. Geistlicher, Ordensangehöriger, Mönch*

Table 1: Examples of pseudo-glosses for the word pair ``Bruder - Bursche''.

According to the Lesk algorithm, the relatedness of a pair of word senses $sim_{c1,c2}$ is, then, defined as a number of overlaps in their respective pseudo-glosses. For the examples given above, this will yield 0 for both radial and hypernym glosses as no overlaps exist.

4.2 Information content based metrics

Resnik (1995), *res* for short, defines semantic similarity between two words w_1 and w_2 as the information content value of their lowest common superordinate (LCS) as given in Equation 1:[3]

[3] For all methods, c_1 and c_2 are concepts (word senses) corresponding to w_1 and w_2.

$$sim(c1,c2) = \max_{c \in S(c1,c2)} \left[-\log p(c) \right]$$

Equation 1

where $S(c_1,c_2)$ is the set of concepts which subsume both c_1 and c_2 and $-log\ p\ (c)$ is the information content. The probability p is computed as the relative frequency of words (representing a specific concept) in a corpus.

As words of the German language are highly inflected, we revised the calculation of the information content and counted the occurrences of individual stems rather than word forms in the original algorithm. Detailed information about our implementation of this procedure for the German language can be found in Gurevych & Niederlich (2004). All information content based metrics for GermaNet, thus, employ the values computed according to a modified information content calculation from a German stem frequency list. The stem frequency list was compiled on the basis of a German newspaper corpus (taz) with about 172 million running tokens.

Resnik's metric assumes that semantic similarity between concepts can be quantified on the basis of the information shared between them. In this case, the GermaNet hierarchy is used to determine the lowest common subsumer for a pair of concepts. If multiple inheritance occurs, there exist more than one lowest common subsumers (LCS). As noted by Pedersen et al. (2004), the “best” LCS can be determined according to three different criteria: the LCS for a pair of concepts with the highest information content, the LCS found at the highest depth, and the LCS that results in the shortest path between a pair of concepts. In our case, the one with the highest information content value is selected. This maximizes the semantic relatedness of the two concepts and, thus, best fits the original Resnik’s definition of semantic similarity.

The next method is that of Lin (1998), referred to as *lin*. He defined semantic similarity using a formula derived from the information theory. This measure is sometimes called a universal semantic similarity measure as it is supposed to be application-, domain-, and resource independent. According to this method, the similarity is given in Equation 2:

$$sim(c1,c2) = \frac{2 \times \log p(lcs(c1,c2))}{\log p(c1) + \log p(c2)}$$

Equation 2

Jiang & Conrath (1997) proposed to combine edge- and node-based techniques in counting the edges and enhancing it by the node-based calculation of the information content as introduced by Resnik (1995). The method is abbreviated as *jcn*. The distance between two concepts c_1 and c_2 is formalized as given in Equation 3:

$$dist(c1, c2) = IC(c1) + IC(c2) - 2 \times IC(lcs(c1, c2))$$

Equation 3

where *IC* is the information content value of a concept, and $lcs(c_1,c_2)$ is the lowest common subsumer of the two concepts. As *jcn* returns distance (ranging from 0 to 34.348 in our data) rather than semantic relatedness, the results have to be post-processed. This transformation should normalize the scores as their distribution is skewed by the 16 out of 57 word pairs, which do not have any LCS, and are assigned the maximum distance value. We employ the hyperbolic tangent function to perform the normalization according to Equation 4:

$$sim_{(c1,c2)} = 1 - (\tanh(dist_{(c1,c2)} \times c))$$

Equation 4

where *c* is a constant and $dist_{(c1,c2)}$ is the distance value of a given word pair. To determine the value of *c*, we use an average human semantic relatedness score from the test dataset as $dist_{(c1,c2)} = dist_{(avg)}$ and set $sim_{(c1,c2)} = .5$.[4] After solving Equation 4, we get the following formula:

$$c = \tanh^{-1}(0.5) / dist_{avg}$$

Equation 5

5. Evaluation

Following Resnik (1995), semantic relatedness was typically evaluated with the help of correlation analysis by other researchers. In this analysis, semantic relatedness scores generated according to particular metrics are correlated with human judgments of semantic relatedness on a set of test word pairs. This way, researchers seek to understand and explain the relations existing between sets of scores produced by particular measurements. E.g. one measurement (the 1st

[4] Not necessarily corresponding to an arithmetic average.

variable) is produced by a human subject (the so-called *gold standard*) and the second one (the 2nd variable) is produced by a particular automatic scoring method. Semantic relatedness of words (the phenomenon studied) is operationalized with the help of sample word pairs, whose semantic relatedness is estimated (either by humans or a computer). Correlations, then, indicate the strength of a linear association between sets of values and serve as an empirical indication of a possible relation between two variables.

We designed an experiment aimed to create an evaluation dataset for the German language. 24 human subjects (native speakers of German) were asked to rate 65 word pairs (containing nouns) on the scale from 0 to 4 for their semantic relatedness. The word pairs represented the dataset by Rubenstein & Goodenough (1965) translated to German. We used a broader definition of semantic relatedness than the one used in the previous work on semantic similarity, where it was defined primarily via the synonymy relation. In particular, the subjects were free to consider any type of semantic relation including association when judging semantic relatedness. The average correlation coefficient for human subjects yielded .8098. The statistical reliability of the judgments is significant. The average correlation value also represents the upper bound of performance in the evaluation of semantic relatedness metrics.

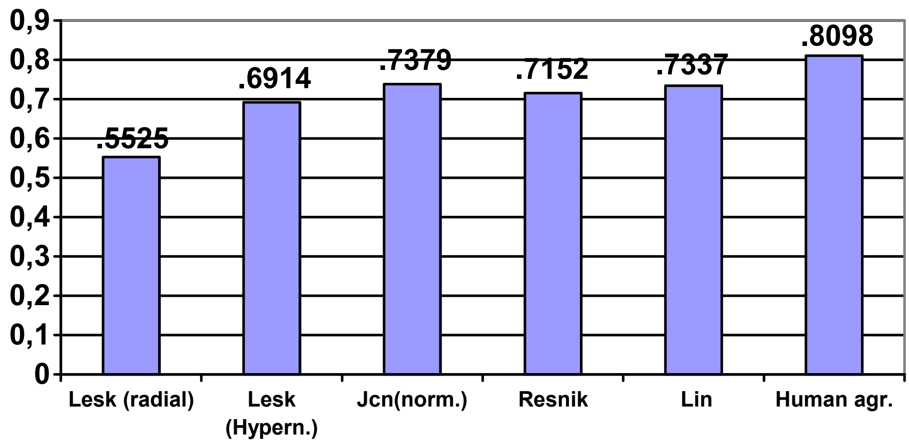

Figure 1: Correlation coefficients for five metrics contrasted with annotator agreement.

In Figure 1, we present final evaluation results for the implemented metrics of semantic relatedness as they have been described in Section 4. The evaluation is based on 57 out of 65 noun pairs as the rest was not represented in GermaNet.

The results are contrasted with the upper bound determined from experiments with human subjects. All of the information content based metrics perform rather well (given that the results of *jcn* are normalized). For the dictionary based metrics, the application of the Lesk algorithm to hypernym glosses performs superior to radial glosses. Notably, the information about the hypernymy relations, turns out to be more useful, given the current modeling of GermaNet, than the inclusion of all lexical-semantic relations. Also, it is quite remarkable that the gloss based metric performs on the scale comparable to the information content based ones. As opposed to them, the gloss metric does not require any corpus evidence, which is difficult and expensive to obtain in many cases. Instead, it only uses the structure of the conceptual hierarchy itself, and is, thus, easily portable to other knowledge bases.

6. Technical issues

6.1 Notes on implementation

RelatednessComparator is a class which takes two words as input and returns a numeric score indicating the degree of semantic relatedness for the given word pair. Five semantic relatedness metrics have been implemented as the descendants of this class. We designed an additional class *PseudoGlossGenerator* automatically generating pseudo glosses on the basis of the conceptual hierarchy as the textual definitions of concepts in GermaNet are small in number. Two of the algorithms, then, operate on the generated pseudo glosses. They are based on the Lesk algorithm (Lesk 1986). The rest of the metrics are classes derived from *InformationBasedComparator*, which is in its turn derived from the class *PathBasedComparator*. They make use of the GermaNet hierarchy as well as the information content values of GermaNet concepts.

We revised the proposal by Resnik (1995) to compute information content values from German texts. A set of utilities have been implemented for this purpose. These programs use the TreeTagger (Schmid 1997) to compile word frequency lists for particular parts of speech, such as nouns, verbs and adjectives. As German is a highly inflected language, the word frequency lists are, then, transformed into stem frequency lists. For this purpose, a free JAVA implementation of the Porter stemmer for German is employed (see http://snowball.tartarus.org/). Two further options for morphological analysis are provided. The first option is to use no stemmer, i.e. information content is computed directly from the word frequency lists. The second option allows to perform morphological analysis rather than stemming. In this case, an external lexicon-based package for morphological analysis is employed (Neumann & Piskorski 2002). A detailed analysis of the impact of using different approaches to morphological analysis as a pre-processing step in computing information content of GermaNet concepts is left to future work. Also, we would like to investigate the impact of the corpus size (very large versus middle size) upon the calculation of information content.

6.2 Semantic relatedness GUI

A graphical user interface was developed to interactively experiment with a set of semantic relatedness measures (Gurevych & Niederlich 2005). The user can enter two words together with their part of speech and specify one of the five methods. Then, the system displays the corresponding word stems, possible word senses according to GermaNet, glosses generated for these word senses, or their information content values. Furthermore, possible combinations of word senses for the two words are created and returned together with various diagnostic information specific to each of the metrics. This may be e.g. word overlaps in glosses for the Lesk based metrics, or lowest common superclasses and the respective information content values, depending on what is appropriate for a particular metric. After all, the best word sense combination for the two words is determined and this is compactly displayed together with a semantic relatedness score. The user interface allows to save the detailed analysis in a text file for off-line inspection.

7. Related work

Unfortunately, a straightforward comparison of our results with other related work on semantic relatedness is difficult. On the one hand, previous works concentrated on semantic similarity rather than semantic relatedness (Hirst & Budanitsky 2004). On the other hand, our research has been done for the German language with GermaNet concepts as the sense inventory. The implementation of our system comprising different methods for estimating semantic relatedness was inspired by the work by Pedersen et al. (2004) and their semantic similarity software. The approach to comparative evaluation of alternative methods can be tracked down to the work by Resnik (1995) and Hirst & Budanitsky (2004). However, they use a dataset with human judgments of semantic similarity in the evaluation. Contrary to that, we conducted a special experiment with human subjects to create a dataset for evaluating semantic relatedness and test the methods on its basis.

8. Conclusions

We presented an implemented system for computing semantic relatedness of pairs of GermaNet concepts. The system employs five different methods for measuring semantic relatedness. For the most part, the results of the metrics correlate well to a set of human judgments about semantic relatedness. The Lesk algorithm applied to pseudo glosses performs better when it uses *hyperhym* glosses as opposed to *radial* glosses. The results of the *jcn* metric have to be additionally normalized in order to obtain reliable results.

Our work demonstrates that a straightforward re-implementation of semantic similarity metrics proposed for WordNet is not always possible on the basis of GermaNet. This is due to some discrepancies in the modeling decisions of GermaNet, but also due to some peculiarities of the German language, e.g. as far as computing the information content is concerned. Special pre-processing components become necessary, such as morphological analysis, to achieve an accurate mapping of word frequencies to information content of German word senses.

In the future, we would like to conduct further research including the following parameters:

- a fine-grained analysis for computing information content: contrasting different approaches to morphological processing, i.e. no stemming, rule-based stemming, a lexicon-based morphological analysis; effects of the corpus size upon the information content calculation and GermaNet coverage;

- computing semantic relatedness of verbs and adjectives (current experiments were restricted to nouns only);

- the deployment of semantic relatedness metrics in natural language processing applications, such as e.g. information retrieval and document classification.

Acknowledgments

The work presented here has been funded by the Klaus Tschira Foundation. We would like to thank Michael Strube for his valuable comments concerning this work and Lothar Lemnitzer for the German word frequency list.

References

Fellbaum, Christiane (Ed.) (1998). WordNet: An Electronic Lexical Database*. Cambridge, Mass.: MIT Press.*

Gurevych, Iryna & Hendrik Niederlich (2004). Semantic Relatedness of GermaNet Word Senses*. Technical Report at EML Research gGmbH, Heidelberg, Germany.*

Gurevych, Iryna & Hendrik Niederlich (2005). A User Interface for Estimating Semantic Relatedness of Words. A system demonstration at the GLDV'2005 conference, Bonn, Germany, 30 March - 01 April.

Gurevych, Iryna & Michael Strube (2004). Semantic similarity applied to spoken dialogue summarization. In Proceedings of the 20th International Conference on Computational Linguistics, Geneva, Switzerland, 23 - 27 August 2004*, pp. 764-770.*

Hirst, Graeme & Alexander Budanitsky (2004). Correcting real-word spelling errors by restoring lexical cohesion. Natural Language Engineering*. To appear.*

Jiang, Jay J. & David W. Conrath (1997). Semantic similarity based on corpus statistics and lexical taxonomy. In Proceedings of the 10th International Conference on Research in Computational Linguistics (ROCLING)*. Tapei, Taiwan.*

Kunze, Claudia (2004). Lexikalisch-semantische Wortnetze. In K.-U. Carstensen, C. Ebert, C. Endriss, S. Jekat, R. Klabunde & H. Langer (Eds.), Computerlinguistik und Sprachtechnologie. Eine Einführung*, pp. 423-431. Heidelberg, Germany: Spektrum Akademischer Verlag.*

Kunze, Claudia & Lothar Lemnitzer (2002). GermaNet - representation, visualization, application. In Proceedings of the International Conference on Language Resources and Evaluation (LREC), Las Palmas, Canary Islands, Spain, 29 - 31 May*, pp. 1485-1491.*

Lemnitzer, Lothar & Claudia Kunze (2002). Adapting GermaNet for the Web. In Proceedings of the first Global WordNetConference, Central Institute of Indian Languages. Mysore, India*, pp. 174-181.*

Lesk, Michael (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In Proceedings of the 5th Annual International Conference on Systems Documentation, Toronto, Ontario, Canada, June*, pp. 24-26.*

Lin, Dekang (1998). An information-theoretic definition of similarity. In Proceedings of the 15th International Conference on Machine Learning, San Francisco, Cal.*, pp. 296-304.*

McCarthy, Diana, Rob Koeling, Julie Weeds & John Carroll (2004). Finding predominant senses in untagged text. In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, Barcelona, Spain, 21-26 July 2004*, pp. 280 - 287.*

Neumann, Günter & Jakub Piskorski (2002). A shallow text processing core engine. Journal of Computational Intelligence*, 18(3):451-476.*

Patwardhan, Siddharth, Satanjeev Banerjee & Ted Pedersen (2003). Using measures of semantic relatedness for word sense disambiguation. In Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics, Mexico City, Mexico*, pp. 241-257.*

Pedersen, Ted (2003). Semantic Similarity Package*. http://www.d.umn.edu/~tpederse/similarity.html.*

Pedersen, Ted, Siddharth Patwardhan & Jason Michelizzi (2004). Wordnet::Similarity - Measuring the relatedness of concepts. In Demonstrations of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Boston, USA, 2-7 May 2004*, pp. 267-270.*

Rennie, Jason (2002). WordNet-QueryData-1.35*. http://search.cpan.org/dist/WordNet-QueryData/.*

Resnik, Phil (1995). Using information content to evaluate semantic similarity in a taxonomy. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, Montréal, Canada, 20-25 August 1995*, Vol. 1, pp. 448-453.*

Rubenstein, Herbert & John Goodenough (1965). Contextual correlates of synonymy. Communications of the ACM*, 8(10):627-633.*

Schmid, Helmut (1997). Probabilistic part-of-speech tagging using decision trees. In Daniel Jones & Harold Somers (Eds.), New Methods in Language Processing*, Studies in Computational Linguistics, pp. 154-164. London, UK: UCL Press.*