Classifying functional relations in Factotum via WordNet hypernym associations*

Tom O'Hara[1] and Janyce Wiebe[2]

[1] Department of Computer Science
New Mexico State University
Las Cruces, NM 88003
`tomohara@cs.nmsu.edu`

[2] Department of Computer Science
University of Pittsburgh
Pittsburgh, PA 15260
`wiebe@cs.pitt.edu`

Abstract. This paper describes how to automatically classify the functional relations from the FACTOTUM knowledge base via a statistical machine learning algorithm. This incorporates a method for inferring prepositional relation indicators from corpus data. It also uses lexical collocations (i.e., word associations) and class-based collocations based on the WordNet hypernym relations (i.e., *is-subset-of*). The result shows substantial improvement over a baseline approach.

1 Introduction

Applications using natural language processing often rely predominantly upon *hierarchical* semantic relations (e.g., *is-a*, *is-subset-of*, and *is-part-of*), along with synonymy and word associations. These are readily available in lexical resources such as Princeton's WordNet [1] or can be extracted directly from corpora [2]. Other types of relations are important, although more difficult to acquire. These correspond to *dictionary differentia* [3], that is, the distinguishing relations given in definitions. Differentia provide information such as attributes, typical functions, and typical purpose. This paper shows how to infer such relations from examples in a knowledge base (KB). For the purpose of this work, the term *functional relations* refers to these non-hierarchical relations, excluding attributes.

The FACTOTUM semantic network [4] developed by Micra, Inc. makes explicit many of the functional relations in Roget's Thesaurus.[3] Outside of proprietary resources such as Cyc [5], Factotum is the most comprehensive KB

* Patrick Cassidy of Micra, Inc. kindly made Factotum available and provided valuable input on the paper. Michael O'Hara helped much with the proofreading. The first author is supported by a generous GAANN fellowship from the Department of Education. Some of the work used computing resources at NMSU made possible through MII Grants EIA-9810732 and EIA-0220590.

[3] Factotum is based on the public domain version of Roget's Thesaurus. The latter is freely available via Project Gutenberg (http://promo.net/pg), thanks to Micra, Inc.

with respect to functional relations. OpenCyc[4] does include definitions of many non-hierarchical relations. However, there are not many instantiations (i.e., relationship assertions), because it concentrates on the higher level of the ontology.

This paper is organized as follows: Section 2 presents more background on the usefulness of differentiating relations, and discusses the main differentiating relations in Factotum. Section 3 shows how corpora can be used to infer clue words for these relations. Section 4 presents results from experiments on classifying the functional relations in Factotum. Section 5 discusses related work. The last section summaries the paper's contributions and mentions areas for future work.

2 Background

2.1 Importance of non-hierarchical semantic relations

Distinguishing features play a prominent role in categorization. For instance, in Tversky's [6] influential *contrast model*, the similarity comparison incorporates factors specific to either term, as well as factors common to both, Tversky also conducted experiments [7] showing that, in certain cases, the distinctive features are given more weight than common ones. Similar results are reported by Medin et al. [8].

Conceptual knowledge for natural language processing is commonly organized into hierarchies called ontologies (e.g., the Mikrokosmos ontology for machine translation [9]). The concepts in these hierarchies are usually partially ordered via the instance and subset relations (i.e., *is-a* and *is-subset-of*). Each is a *relation of dominance*, which Cruse [10] considers as the defining aspect of hierarchies. He points out that an important part of hierarchies is the differentiation of siblings. This is the role of *conceptual differentia*, that is, the semantic relations that distinguish sibling concepts. Without these relations, the information in hierarchical lexicons would only indicate how the lexicalized concepts represented are ordered without indicating the differences among the concepts.

Manually-derived lexicons, such as the Mikrokosmos English lexicon [11], often contain differentia in the rich case-frame structures associated with the underlying concepts. This contrasts with semi-automatically derived lexicons such as WordNet [1], which emphasize the lexical hierarchy but not the underlying semantics. For instance, Mikrokosmos[5] averages about 2.4 properties per concept (including some inverse relations), whereas WordNet[6] only averages 1.3 (including inverses).[7] This suggests that the reason large-scale lexicons tend to incorporate less differentia is due more to the difficulty in acquiring the

[4] Version 0.7 of OpenCyc, a publicly available subset of Cyc (www.opencyc.org).

[5] 1998 version of Mikrokosmos (crl.nmsu.edu/Research/Projects/mikro/index.html).

[6] Version 1.7 of WordNet (www.cogsci.princeton.edu/˜wn).

[7] *Properties* refers to functional relations, attributes and part-whole relations (e.g., *is-member-meronym-of*), excluding just the instance and subset relations. WordNet 1.6 only averages 0.64 properties, so version 1.7 represents a substantial improvement.

information than to the relative worth of the information. Factotum compares favorably in this respect, averaging 1.8 properties per concept. OpenCyc provides the highest average at 3.7 properties per concept (with an emphasis on argument constraints and other usage restrictions).[8]

Hirst [12] advocates adding case structures to standard dictionaries, in the same manner that learner's dictionaries indicate verbal subcategorization frames. This would provide a common resource for more-detailed language knowledge, useful for humans as well as for computerized processing.

Work in formal semantics tends not to cover functional relations much, although there are some notable exceptions. Pustejovsky's Generative Lexicon theory accounts for them in his *qualia* structure [13]. Mel'čuk's Meaning Text Theory [14] accounts for them via lexical functions in his *Explanatory Combinatorial Dictionary*. Both of these theories are quite influential, adding more support that functional relations are desirable although perhaps difficult to acquire. Heylen [15] discusses the connection between the two theories.

2.2 Factotum

The Factotum semantic network [4] is a knowledge base derived initially from the 1911 version of Roget's Thesaurus. Part of purpose is to make explicit the relations that hold between the Roget categories and the words listed in each entry. It incorporates information from other resources as well, in particular the Unified Medical Language System (UMLS), which formed the basis for the initial set of semantic relations.

Figure 1 shows a sample from Factotum. This illustrates that the basic Roget organization is still used, although additional hierarchical levels have been added. The relations are contained within double braces (e.g., "{{has_subtype}}") and generally apply from the category to each word in the synonym list on the same line. Therefore, the line with "{{result_of}}" indicates that conversion is the result of transforming, as shown in the semantic relation listing that would be extracted.[9] There are over 400 relations instantiated in the semantic network. Some of these are quite specialized (e.g., *has-brandname*). In addition, there are quite a few inverse functions, since most of the relations are not symmetrical. Certain features of the semantic network representation are currently ignored during the relation extraction. For example, relation specifications can have qualifier prefixes, such as an ampersand to indicate that the relationship only sometimes holds.

Table 1 shows the most common relations in terms of usage in the semantic network, and includes others that are used in the experiments discussed later.[10]

[8] These figures are derived by counting the number of relations excluding the instance and subset ones. OpenCyc's comments and lexicalizations are also excluded (implicit in Factotum and WordNet). The count is then divided by the number of concepts.

[9] For clarity, some of the relations are renamed to make the directionality more explicit, following a suggestion for their interpretation in the Factotum documentation.

[10] The database files and documentation for the semantic network are available from Micra, Inc. via ftp://micra.com/factotum.

A6.1.4 CONVERSION (R144)
#144. Conversion.
N. {{has_subtype(change, R140)}} conversion, transformation.
{{has_case: @R7, initial state, final state}}.
{{has_patient: @R3a, object, entity}}.
{{result_of}} {{has_subtype(process, A7.7)}} converting, transforming.
{{has_subtype}} processing.
transition.
⇒

⟨change, *has-subtype*, conversion⟩	⟨change, *has-subtype*, transformation⟩
⟨conversion, *has-case*, initial state⟩	⟨conversion, *has-case*, final state⟩
⟨conversion, *has-patient*, object⟩	⟨conversion, *has-patient*, entity⟩
⟨conversion, *is-result-of*, converting⟩	⟨conversion, *is-result-of*, transforming⟩
⟨process, *has-subtype*, converting⟩	⟨process, *has-subtype*, transforming⟩
⟨conversion, *has-subtype*, processing⟩	

Fig. 1. Sample entry from Factotum with extracted relations

The functional relations are shown in boldface. The exclusion of the meronymic or part-whole relations (e.g., *is-conceptual-part-of*) accords with their classification by Cruse [10] as hierarchical relations. Note that the usage counts just reflect relationships[11] explicitly labeled in the KB data file. For instance, this does not account for implicit *has-subtype* relationships based on the hierarchical organization of the thesaural groups.

Table 2 shows the relation usage in WordNet version 1.7. This shows that the majority of the relations are hierarchical (*is-similar-to* can be considered as a hierarchical relation for adjectives). Therefore, the information in Factotum complements WordNet through the inclusion of more functional relations.

3 Inferring relation markers

Note that Factotum does not indicate the way the relationships are expressed in English. WordNet similarly does not indicate this, but does include definition glosses that can be used in some cases to infer the *relation markers* (i.e., generalized case markers). For example,

Factotum: ⟨drying, *is-function-of*, drier⟩

WordNet: {dry#1, dry_out#3} remove the moisture from and make dry
{dryer#1, drier#2} an appliance that removes moisture

Therefore, the Factotum relations cannot be used as is to provide training data for learning how the relations are expressed in English. This contrasts with corpus-based annotations, such as Treebank II [16] and FrameNet [17], where the relationships are marked in context.

[11] For clarity, *relationships* refers to relation instantiations, and *relations* to the types.

Relation	Usage	Description
has-subtype	37355	inverse of *is-a* relation
is-property-of	7210	object with given salient character
is-caused-by	3203	indicates force that is the origin of something
has-property	2625	salient property of an object
has-part	2055	a part of a physical object
has-high-intensity	1671	intensifier for the property or characteristic
has-high-level	1564	implication for the activity (e.g., intelligence)
is-antonym-of	1525	generally used for lexical opposition
is-conceptual-part-of	1408	parts of other entities (in case relations)
has-metaphor	1313	non-literal reference to the word
causes$_{mental}$	1208	motivation (causation in the mental realm)
uses	1157	a tool needing active manipulation
is-performed-by	1081	human actor for the event
performs$_{human}$	987	human role in performing some activity
is-function-of	983	artifact that passively performs the function
has-result	977	more specific type of *causes*
has-conceptual-part	937	generalization of *has-part*
is-used-in	930	activity or some desired effect for the entity
is-part-of	898	distinguishes part from group membership
causes	866	inverse of *is-caused-by*
has-method	830	method used to achieve some goal
is-caused-by$_{mental}$	810	inverse of *causes*$_{mental}$
has-consequence	785	causation due to a natural association
has-commencement	663	state that commences with the action
is-location-of	655	absolute location of an object
requires	341	object or sub-action necessary for an action
is-studied-in	331	inquires into any field of study
is-topic-of	177	document or other communication for the subject
produces	166	what an action yields, secretes, generates, etc.
is-measured-by	158	instrument or method for measuring something
is-job-of	117	occupation title for a job function
is-patient-of	101	action that the object participates in
is-facilitated-by	98	object or sub-action aiding an action
is-biofunction-of	27	biological function of parts of living things
was-performed-by	22	*is-performed-by* occurring in the past
has-consequence$_{object}$	21	consequence for the patient of an action
is-facilitated-by$_{mental}$	9	trait that facilitates some human action

Table 1. Sample relations from Micra's FACTOTUM. Boldface relations are used in the experiments in Section 4.

Relation	Usage	Description
has-hypernym	88381	superset relation
is-similar-to	22492	similar adjective synset
is-member-meronym-of	12043	constituent member
is-part-meronym-of	8026	constituent part
is-antonym-of	7873	opposing concept
is-pertainym-of	4433	noun that adjective pertains to
also-see	3325	related entry (for adjectives and verbs)
is-derived-from	3174	adjective that adverb is derived from
has-verb-group	1400	verb senses grouped by similarity
has-attribute	1300	related attribute category or value
is-substance-meronym-of	768	constituent substance
entails	426	action entailed by the verb
causes	216	action caused by the verb
has-participle	120	verb participle

Table 2. Relation usage in WordNet (version 1.7)

However, given the increased coverage of the web, the relation markers can be inferred. For example, each of the relationships can be used in proximity searches involving the source and target terms. For example, using AltaVista's Boolean search[12], this can be done via 'source NEAR target'. Unfortunately, this technique would require detailed post-processing of the web search results, possibly including parsing in order to extract the patterns. As an expedient, common prepositions[13] are included in a series of proximity searches to find the preposition occurring the most with the terms. For instance, given the relationship ⟨drying, *is-function-of*, drier⟩, the following searches would be performed.

drying NEAR tendril NEAR of
drying NEAR tendril NEAR to
...
drying NEAR tendril NEAR "because of"

To account for prepositions that occur frequently (e.g., 'of'), mutual information (MI) statistics [2] are used in place of the raw frequency when rating the potential markers. These are calculated as follows:

$$MI_{prep} = log_2 \frac{P(X,Y)}{P(X) \times P(Y)} \approx log_2 \frac{f(\text{source NEAR target NEAR prep})}{f(\text{source NEAR target}) \times f(\text{prep})}$$

Such checks are done for the 25 most common prepositions to find the preposition yielding the highest mutual information score. Using this metric, the top three

[12] AltaVista's Boolean search is available at www.altavista.com/sites/search/adv.
[13] The common prepositions are determined from the prepositional phrases assigned functional annotations in Penn Treebank II [16].

markers for the ⟨drying, *is-function-of*, drier⟩ relationship are 'during', 'after', and 'with'.

This technique can readily be extended to finding relation markers in foreign languages, such as Spanish, given a bilingual dictionary. Ambiguous translations pose a complication, but in most of these cases, similar relation markers should be likely unless the relations between the alternative meaning pairs diverge significantly.[14] For example, when the process is applied to the translated relationship for the example, namely ⟨secar, *is-function-of*, secarador⟩, the top three markers are 'con', 'de', and 'para'.

4 Classifying the functional relations

4.1 Methodology

Given the functional relationships in Factotum along with the inferred relation markers, machine learning algorithms can be used to infer what relation most likely applies to terms occurring together with a particular marker. Note that the main purpose of including the relation markers is to provide clues for the particular type of relation. Because the source term and target terms might occur in other relationships, associations based on them alone might not be as accurate. In addition, the inclusion of these clue words (e.g., the prepositions) makes the task closer to what would be done in inferring the relations from free text. In effect, this task is preposition disambiguation, using the Factotum relations as senses.

A straightforward approach for preposition disambiguation would use standard feature sets for word-sense disambiguation (WSD), such as those used in the Senseval competitions [19, 20]. These include syntactic features for the immediate context (e.g., the parts-of-speech of surrounding words). More importantly, WSD feature sets include semantic features based on collocations (e.g., word associations). The latter can be highly accurate, but might over-fit the data and generalize poorly. To overcome these problems, class-based collocations are also incorporated, using WordNet hypernym synsets.

Figure 2 gives the feature settings used in the experiments. These are similar to the settings used by the grling-sdm system in the first Senseval competition [21], except for the inclusion of the hypernym-based collocations.

Word collocation features are derived by making two passes over the training data. The first pass tabulates the co-occurrence counts for the words in a window around the target word and each of the classification values or categories (e.g., the preposition senses). These counts are used to derive a conditional probability estimate of each class value given the various potential collocates. Those exceeding a certain threshold are collected into a list associated with the class value, making this a "bag of words" approach. As shown in Figure 2, a potential collocate is selected whenever its co-occurrence with the class category increases

[14] Sidorov et al. [18] illustrate the differences that might arise for terms referring to non-adults in English, Spanish, and Russian.

Features:

POS_{source}:	part-of-speech of the source term
POS_{target}:	part-of-speech of the target term
Prep:	preposition serving as relation marker (or 'n/a' if not inferable)
WordColl_i:	true if context contains any word collocation for relation i
HypernymColl_i:	true if context contains any hypernym collocation for relation i

Collocation selection:

Frequency constraint:	$f(word) > 1$
Conditional independence threshold:	$\frac{p(c\|coll)-p(c)}{p(c)} >= 0.2$
Organization:	per-class-binary grouping [22]

Model selection:

Decision tree (via Weka's J48 classifier [23])

10-fold cross-validation

Fig. 2. Features used in semantic role classification experiments

the probability for the latter by 20%. The second pass determines the value for the collocational feature of each classification category by checking whether the current context has any of the associated collocation words. For the test data, only the second pass is made, using the collocation lists derived from the training data.

In generalizing this to a class-based approach, the potential collocational words are replaced with each of their hypernym ancestors from WordNet. Since the co-occurring words are not sense-tagged, this is done for each synset serving as a different sense of the word. (Likewise, in the case of multiple inheritance, each parent synset is used.) For example, given the co-occurring word "money", the counts would be updated as if each of the following tokens were seen, shown grouped by sense.

1. {medium_of_exchange#1, monetary_system#1, standard#1, criterion#1, measure#2, touchstone#1, reference_point#1, point_of_reference#1, reference#3, indicator#2, signal#1, signaling#1, sign#3, communication#2, social_relation#1, relation#1, abstraction#6}
2. {wealth#4, property#2, belongings#1, holding#2, material_possession#1, possession#2}
3. {currency#1, medium_of_exchange#1, monetary_system#1, standard#1, criterion#1, measure#2, touchstone#1, reference_point#1, point_of_reference#1, reference#3, indicator#2, signal#1, signaling#1, sign#3, communication#2, social_relation#1, relation#1, abstraction#6}

Thus, the word token 'money' is replaced by 41 synset tokens. Then, the same two-pass process described above is performed over the replacement tokens. Although this introduces noise due to ambiguity, the conditional-independence selection scheme [22] compensates somewhat (e.g., by selecting hypernym synsets that only occur with specific categories).

Figure 3 contains sample feature specifications from the experiments discussed in the next section. This shows that 'n/a' is used whenever a preposition marker for a particular relationship cannot be inferred. For brevity, the feature specification only includes collocation features for the most frequent relations. Sample collocations are also shown for the relations. In the word collocation case, the occurrence of 'similarity' is used to determine that the *is-caused-by* feature (HC_1) should be set on for the first two instances; however, there is no corresponding hypernym collocation due to conditional-independence filtering. Although 'new' is not included as a word collocation, one of its hypernyms, namely 'Adj:early#2', is used to determine that the *has-consequence* feature (HC_3) should be on in the last instance.

4.2 Results

For this task, the set of functional relations in Factotum are determined by removing the hierarchical relations (e.g., *has-subtype* and *has-part*) along with the attribute relations (e.g., *is-property-of*). In addition, in cases where there are inverse functions (e.g., *causes* and *is-caused-by*), the most frequently occurring relation of each inverse pair is used. This is done because the approach currently does not account for argument order. The boldface relations in the listing shown earlier in Table 1 are those used in the experiment. Only single-word source and target terms are considered to simplify the WordNet hypernym lookup. The resulting dataset has 5959 training instances. The dataset also includes the inferred relation markers, thus introducing some noise.

Table 3 shows the results of the classification. The combined use of both collocation types achieves the best overall accuracy at 71.2%, which is good considering that the baseline of always choosing the most common relation (*is-caused-by*) is 24.2%. This combination generalizes well by using hypernym collocations, while retaining specificity via word collocations. Note that the classification task is quite challenging, given the large number of choices and high entropy [24].

Experiment	*Accuracy*	*Stdev*
Word	68.4	1.28
Hypernym	53.9	1.66
Combined	*71.2*	1.78

# Instances:	5959
# Classes:	21
Entropy:	3.504
Baseline:	24.2

Table 3. Functional relation classification, using inferred prepositions along with source and target. The accuracy figures are averages based on 10-fold cross validation. The gain in accuracy for the combined experiment versus the word experiment is statistically significant at $p < 0.01$ (via a paired t-test).

Relationships from Factotum with inferred markers:

⟨similarity, *is-caused-by*, connaturalize⟩	n/a
⟨similarity, *is-caused-by*, rhyme⟩	by
⟨approximate, *has-consequence*, imprecise⟩	because
⟨new, *has-consequence*, patented⟩	with

Word collocations only:

Relation	POS_s	POS_t	Prep	WC_1	WC_2	WC_3	WC_4	WC_5	WC_6	WC_7
is-caused-by	NN	VB	n/a	1	0	0	0	0	0	0
is-caused-by	NN	NN	by	1	0	0	0	0	0	0
has-consequence	NN	JJ	because	0	0	0	0	0	0	0
has-consequence	JJ	VBN	with	0	0	0	0	0	0	0

Sample collocations:

is-caused-by	{bitterness, evildoing, monochrome, *similarity*, vulgarity, wit}
has-consequence	{abrogate, frequently, insufficiency, nonplus, ornament, useless}

Hypernym collocations only:

Relation	POS_s	POS_t	Prep	HC_1	HC_2	HC_3	HC_4	HC_5	HC_6	HC_7
is-caused-by	NN	VB	n/a	0	0	0	0	0	0	0
is-caused-by	NN	NN	by	0	0	0	0	0	0	0
has-consequence	NN	JJ	because	0	0	0	0	0	0	0
has-consequence	JJ	VBN	with	0	0	1	0	0	0	0

Sample collocations:

is-caused-by	{N:hostility#3, N:inelegance#1, N:humorist#1, V:stimulate#4}
has-consequence	{V:abolish#1, *Adj:early#2*, N:inability#1, V:write#2, V:write#7}

Combined collocations:
The combination of the above specifications:
that is, ⟨Relation, POS_s, POS_t, Prep, WC_1, ... WC_7, HC_1, ... HC_7⟩.

where POS_s and POS_t are the parts of speech for the source and target terms, and the relations for the word and hypernym collocations (WC_i and HC_i) follow:

1. *is-caused-by* 2. *is-function-of* 3. *has-consequence* 4. *has-result*
5. *is-caused-by*$_{mental}$ 6. *is-performed-by* 7. *uses*

Fig. 3. Sample feature specifications for the different experiment configurations. The collocation features are not shown for the low frequency relations.

5 Related Work

Recently there has a bit of work related to preposition disambiguation and semantic role classification. Litkowski [25] presents manually-derived rules for disambiguating 'of'; Srihari et al. [26] present manually-derived rules for disambiguating prepositions used in named entities. Gildea and Jurafsky [27], as well as Blaheta and Charniak [28], address the more general problem of assigning semantic roles to arbitrary constituents of a sentence. We provide a detailed comparison elsewhere [29], including other work in preposition disambiguation. Syntactic functional relations are important as well. Dini et al. [30] show how relations extracted from parse annotations facilitate word sense disambiguation.

Scott and Matwin [31] also use WordNet hypernyms for classification, in particular topic detection. Their approach is different in that they include a numeric density feature for each synset that subsumes words appearing in the document, potentially yielding hundreds of features. We just have a binary feature for each of the relations being classified. They only consider nouns and verbs, whereas we also include adjectives.[15] As with our approach, they consider all senses of a word, distributing the alternative readings throughout the set of features. Gildea and Jurafsky [32] instead just select the first sense for their hypernym features.

Factotum has been used in other language processing research. Cassidy [4] shows how control of inference might be done for Factotum and discusses its use in word sense disambiguation. Bolshakov et al. [33] discuss the translation of Factotum into Russian and the complications due to the mismatch in the lexicalization of various concepts. Gelbukh [34] shows how Factotum can be useful for word-sense disambiguation and related tasks (e.g., machine translation) via path-based distance measures derived from the network. Follow-up work [35] discusses additional tasks that can be solved via the path minimization approach, such as resolving prepositional phrase attachment. This also describes more customizations to the standard shortest-paths algorithms for use in language processing applications (e.g., dealing with the different types of links in the semantic network).

6 Conclusion

Factotum provides complementary information to that contained in WordNet and other lexical resources. This paper shows how automatic classification of the functional relations from this data can be done, using a combination of word and hypernym collocations. The approach achieves good accuracy (71.2%), which is nearly three times the baseline. We also illustrate how relation markers can be inferred using corpus-based techniques (via AltaVista's proximity search).

Recent work by Gildea and Jurafsky [32] illustrates the use of mappings from FrameNet's fine-grained relations to coarse-grained ones more commonly used

[15] The adjective hierarchy is augmented by treating *is-similar-to* as *has-hypernym*. Adverbs would be included, but there is no hierarchy for them. Adverbs are related to adjectives via *is-derived-from*, so future work might treat these as *has-hypernym*.

in computational linguistics. This suggests a method for converting annotations from one lexical resource to another. Future work will pursue this with Factotum and other knowledge bases such as OpenCyc. We will also investigate more fully the inference of relation markers for foreign languages (e.g., via proximity searches of the source and target terms from the translated semantic network produced by Gelbukh's technique [35]).

References

1. Miller, G.: Special issue on WordNet. International Journal of Lexicography **3(4)** (1990)
2. Manning, C.D., Schütze, H.: Foundations of Statistical Natural Language Processing. MIT Press, Cambridge, Massachusetts (1999)
3. Landau, S.: Dictionaries: The Art and Craft of Lexicography. Second edn. Cambridge University Press, Cambridge (2001)
4. Cassidy, P.J.: An investigation of the semantic relations in the Roget's Thesaurus: Preliminary results. In: Proc. CICLing '00. (2000)
5. Lenat, D.B.: Cyc: A large-scale investment in knowledge infrastructure. Communications of the ACM **38(11)** (1995)
6. Tversky, A.: Features of similarity. Psychological Review **84(4)** (1977) 327–352
7. Gati, I., Tversky, A.: Weighting common and distinctive features in perceptual and conceptual judgements. Cognitive Psychology **16** (1984) 341–370
8. Medin, D.L., Goldstone, R.L., Gentner, D.: Respects for similarity. Psychological Review **100** (1993) 252–278
9. Mahesh, K., Nirenburg, S.: A situated ontology for practical NLP. In: Proc. Workshop on Basic Ontological Issues in Knowledge Sharing. (1995) International Joint Conference on Artificial Intelligence (IJCAI-95), Aug. 19-20, 1995. Montreal, Canada.
10. Cruse, D.A.: Lexical Semantics. Cambridge University Press, Cambridge (1986)
11. Onyshkevych, B., Nirenburg, S.: A lexicon for knowledge-based MT. Machine Translation **10(2)** (1995) 5–57 Special Issue on Building Lexicons for MT.
12. Hirst, G.: Why dictionaries should list case structures. In: Proc. Conference on Advances in Lexicography. (1986) University of Waterloo, November.
13. Pustejovsky, J.: The Generative Lexicon. MIT Press, Cambridge, MA (1995)
14. Mel'čuk, I.A., Polguere, A.: A formal lexicon in the meaning-text theory (or how to do lexica with words). Computational Linguistics **13 (3-4)** (1987) 261–275
15. Heylen, D.: Lexical functions, generative lexicons and the world. In Saint-Dizier, P., Viegas, E., eds.: Computational Lexical Semantics. Cambridge University Press, Cambridge (1995) 125–140
16. Marcus, M., Kim, G., Marcinkiewicz, M.A., MacIntyre, R., Bies, A., Ferguson, M., Katz, K., Schasberger, B.: The Penn Treebank: Annotating predicate argument structure. In: ARPA Human Language Technology Workshop. (1994)
17. Fillmore, C.J., Wooters, C., Baker, C.F.: Building a large lexical databank which provides deep semantics. In: Proceedings of the Pacific Asian Conference on Language, Information and Computation. (2001) Hong Kong.
18. Sidorov, G., Bolshakov, I., Cassidy, P., Galicia-Haro, S., Gelbukh, A.: 'Non-adult' semantic field: comparative analysis for English, Spanish, and Russian. In: Proc. 3rd Tbilisi Symposium on Language, Logic, and Computation. (1999)

19. Kilgarriff, A., Palmer, M.: Introduction to the special issue on SENSEVAL. In *Computers and the Humanities* [36] 15–48
20. Edmonds, P., Cotton, S., eds.: Proceedings of the SENSEVAL 2 Workshop. Association for Computational Linguistics (2001)
21. O'Hara, T., Wiebe, J., Bruce, R.F.: Selecting decomposable models for word-sense disambiguation: The grling-sdm system. In *Computers and the Humanities* [36] 159–164
22. Wiebe, J., McKeever, K., Bruce, R.: Mapping collocational properties into machine learning features. In: Proc. 6th Workshop on Very Large Corpora (WVLC-98), Montreal, Quebec, Canada, Association for Computational Linguistics SIGDAT (August 1998) 225–233
23. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann (1999)
24. Kilgarriff, A., Rosenzweig, J.: Framework and results for English SENSEVAL. In *Computers and the Humanities* [36] 15–48
25. Litkowski, K.C.: Digraph analysis of dictionary preposition definitions. In: Proceedings of the Association for Computational Linguistics Special Interest Group on the Lexicon. (2002) July 11, Philadelphia, PA.
26. Srihari, R., Niu, C., Li, W.: A hybrid approach for named entity and sub-type tagging. In: Proc. 6th Applied Natural Language Processing Conference. (2001)
27. Gildea, D., Jurafsky, D.: Automatic labeling of semantic roles. In: Proc. ACL-00. (2000)
28. Blaheta, D., Charniak, E.: Assigning function tags to parsed text. In: Proc. NAACL-00. (2000)
29. O'Hara, T., Wiebe, J.: Classifying preposition semantic roles using class-based lexical associations. Technical Report NMSU-CS-2002-013, Computer Science Department, New Mexico State University (2002)
30. Dini, L., Tomaso, V.D., Segond, F.: Word sense disambiguation with functional relations. In: Proc. First International Conference on Language Resources and Evaluation (LREC). (1998) 28-30 May 1998, Granada, Spain.
31. Scott, S., Matwin, S.: Text classification using WordNet hypernyms. In Harabagiu, S., ed.: Use of WordNet in Natural Language Processing Systems: Proceedings of the Conference, Somerset, New Jersey, Association for Computational Linguistics (1998) 38–44
32. Gildea, D., Jurafsky, D.: Automatic labeling of semantic roles. Computational Linguistics **28(3)** (2002) 245–288
33. Bolshakov, I., Cassidy, P., Gelbukh, A.: Russian Roget: Parallel Russian and English hierarchical thesauri with semantic links, based on an enriched Roget's Thesaurus. In: Proc. Annual International Conf. on Applied Linguistics Dialogue-95, Moscow, Russia (1995) 57–60
34. Gelbukh, A.F.: Using a semantic network for lexical and syntactical disambiguation. In: Proc. CIC-97, Nuevas Aplicaciones e Innovaciones Tecnológicas en Computación. (1997) 352–366 Simposium Internacional de Computación, CIC, IPN, Mexico City, Mexico.
35. Gelbukh, A.F.: Using a semantic network dictionary in some tasks of disambiguation and translation. Technical report, CIC, IPN, Mexico (1998) Serie Roja, N 36.
36. Kilgarriff, A., Palmer, M., eds.: Computers and the Humanities: Special Issue on SENSEVAL. Volume 34 (1-2). Kluwer Academic Publishers, Dordrecht, the Netherlands (2000)