On the categorization of Cause and Effect in WordNet

Cristina Butnariu, Tony Veale

School of Computer Science and Informatics,
University College Dublin, Dublin 4, Ireland
{Ioana.Butnariu, Tony.Veale}@ucd.ie

Abstract. The task of detecting causal connections in text would benefit greatly from a comprehensive representation of Cause and Effect in WordNet, since previous studies show that semantic abstractions play an important role in the linguistic detection of semantic relations, in particular the cause-effect relation. Based on these studies on causality, and on our own general intuitions about causality, we propose a cover-set of different WordNet categories to represent the ontological classes of Cause and Effect. We also propose a corpus-based approach to the population of these categories, whereby candidate words and senses are identified in a large corpus (such as the Google N-gram corpus) using specific syntagmatic patterns. We describe experiments using the Cause-Effect dataset from the 2007 SemEval workshop to evaluate the most effective combinations of WordNet categories and corpus data. Ultimately, we propose extending the WordNet category of Causal-Agent with the word-senses identified by this experimental exploration.

Keywords: semantic relations, WN categorization, cause, effect, causality, syntagmatic patterns.

1 Introduction

Causality plays a fundamental role in textual inference, not just because it is intrinsic to notions of cause and effect, but also because it is central to the meaning of artifacts, agents, products (whether physical or abstract) and even natural phenomena. Artifacts possess a purpose, or telicity, that is causally defined, while agents are often defined by the products that they cause to exist, and natural phenomena like storms and other acts of god are typically conceptualized as intentional processes. Since each of these notions – agents, artifacts, products and natural phenomena – are all explicitly represented and richly specialized in a lexical ontology like WordNet [4], one can ask whether the concepts of Cause and Effect can and should be as richly represented in WordNet. Of course, since these concepts correspond to the nouns "cause" and "effect", they clearly are represented in WordNet. Indeed, WordNet represents different nuances of these concepts, distinguishing between cause-as-agent (or {causal-agent}) and cause-as-reason (or {cause, reason, grounds}) and effect-as-outcome and effect-as-symptom.

Nonetheless, these attempts at ontologizing causality are simultaneously too coarse-grained – insofar as they admit of too many specializations that are not meaningfully represented as causes or effects – or too under-developed – insofar as they are little more than ontological place-holders that have few meaningful specializations. For instance, because WordNet defines the concept Causal-agent as a hypernym of Person, concepts like Victim, Martyr and Casualty will be seen indirectly as agents of their own state, even when this view is counter to their true meaning (these concepts are clearly better defined as causal-patients, though WordNet lacks such a concept). Likewise, WordNet categorizes antacids and other medicinally helpful substances (such as antacids) as causal agents but denies this classification to unhelpful substances such as poisons and allergens, as well as to harmful weather phenomena (such as storms and earthquakes) that are readily conceptualized as major causes by humans. Similarly, WordNet 2.1 only provides four possible specializations of the symptom meaning of Effect when any number of other WordNet concepts can, in the right circumstance, by seen as symptoms. Indeed, only 30% of the concepts whose WordNet 2.1 gloss contains the phrase "that causes" are categorized as causal agents in WordNet, even though all the concepts are valid examples of causal agency.

WordNet would clearly benefit then from considerable house-cleaning under its categories of Cause (and Causal-Agent) and Effect. In this paper, we consider the effectiveness of WordNet in recognizing and capturing cause and effect relationships, by focusing on the cause-effect relation in the recent SemEval semantic-relations task (see [7]). While virtually all entrants in this task adopted a supervised machine-learning approach to the problem of detecting relations such as cause-effect between noun-pairs, we consider here how well WordNet, without training, can perform on this task when its basic causal repertoire is augmented with causally-indicative syntagmatic cues from a large corpus. In section 2 we briefly describe past-work on this topic, before presenting a purely WordNet-based approach to cause and effect in section 3. Causality is a highly contextual notion: a dinner plate is an effect (product) in the context of its construction, and a cause of pain when used as a projectile in the context of a domestic argument (see [12]). WordNet cannot hope to anticipate or reflect all of these contexts, but the language used in context-specific corpora may well reflect these causal nuances. In section 4 then, we present a corpus-based approach to identifying possible causes and effects in terms of lexico-syntactic patterns. Section 5 then presents an empirical evaluation of this corpus/WordNet combination. The paper concludes with some closing remarks in section 6.

2 Past Work

There have been many attempts in the computational linguistic communities to define and understand the Causality relation. Nastase in [11] defines causality as a general class of relations that describe how two occurrences influence each other. Further she proposes the following sub-relations of causality: *cause*, *effect*, *purpose*, *entailment*, *enablement*, *detraction* and *prevention*. She states that semantic relations can be expressed in different syntactic forms, at different syntactic levels. Hearst [8] states that "certain lexico-syntactic patterns unambiguously indicate certain semantic relations". The key issue then is to discover the most efficient patterns that indicate a certain semantic relation. These patterns can be either manually specified by linguists or discovered automatically from corpora. For instance, the subject-verb-object lexico-syntactic pattern (where subject and verb are noun-phrases) was used in [3] to detect causal relations in text, and from these patterns, automatically construct Bayesian for causal inference.

Girju proposes in [5] a classification of lexical patterns for mining instances of the causality relation from corpora, and describes a semi-automatic method to discover new patterns. She uses a general pattern *<NP1 verb NP2>* in combination with WordNet to impose semantic restrictions on NP1 (the Cause category) and NP2 (the Effect category). She defines the classes of Cause and Effect in WordNet terms as a patchwork of different synsets/categories. For Effect, she proposes a cover-set comprising the following synsets: {human_action, human_activity, act}, {phenomenon}, {state}, {psychological_feature} and {event}. However, she observes that the Cause class is harder to define in such terms of WordNet categories, since the notion of causality is frequently entwined with, and difficult to separate from, that of metonymy (e.g., does the poison cause death or the poisoner, or both? The gun or the gun-man?). She thus relies entirely on the intuitions already encoded in WordNet under the category of Causal-Agent. Girju then ranks the output patterns into five categories, according to their degree of ambiguity. She reports a precision of 68% when applying these patterns to a terrorism corpus.

The SemEval-2007 task 4 (see [7]) concerned itself with the classification of semantic relations between pairs of words in a given context. Seven semantic relations were proposed and a training dataset for each semantic relation (comprising positive and negative examples, the latter in the form of near misses) was collected from the web and classified by two human judges. The relation that interests us here is the Cause-Effect relation, which the task authors define as follows: "Cause-Effect(X,Y) is true for a sentence S if X and Y appear close in the syntactic structure of S and the situation described in S entails that X is the cause of Y." There are some restrictions imposed on X and Y: "X and Y can be a nominal denoting an event, state, activity or an entity, as a metonymic expression of an occurrence.” The data-set for this relation comprises 220 noun pairs (with WordNet sense-tags and associated context fragments), of which 114 pairs are positive exemplars and 106 are negative "near-miss" exemplars.

3 Defining Cause and Effect in WordNet terms

Following Girju, we should intuitively expect a variety of high-level WordNet abstractions to encompass a range of concepts that play an enabling role in achieving certain ends, and thus to contribute to the cover-set that defines the class of Causes. Recall that Girju limits the definition of Cause to the WordNet category {causal_agent}, a snapshot of which is presented in Figure 1.

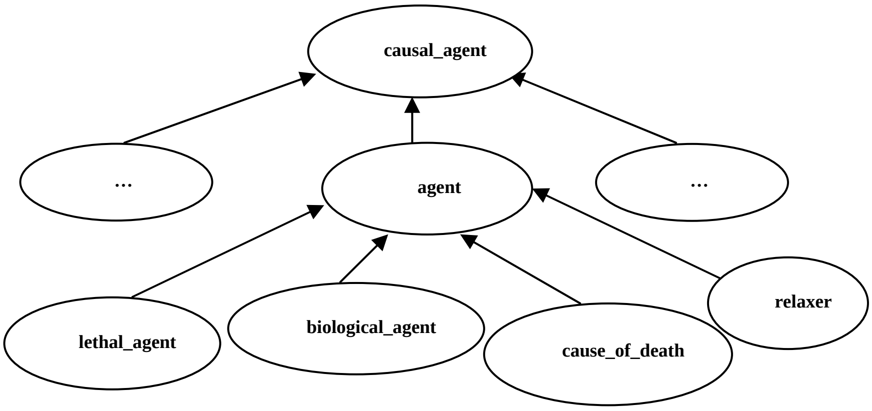

Fig. 1. The figure shows a fragment of the taxonomy for the lexical concept *{causal_agent}* in WordNet.

In contrast, we broaden the cover-set of Causes to include the following WordNet categories and their descendants: {causal_agent}, {psychological_feature}, {attribute}, {substance} (insofar as many are biological causal-agents), {phenomenon}, {communication} (insofar as they can drive agents to action), {natural_action} and {organic_process}. In contrast, the class of Effects should include: {psychological_feature}, {attribute}, {physical_process}, {phenomenon}, {natural_action}, {possession} and {organic_process}. The two cover-sets are similar because causes and effects typically interact as part of complex causal chains, so the causes of one effect are often themselves the effects of prior causes.

It is worth considering how well these WordNet-based cover-sets correspond to the exemplars of the SemEval dataset. Figure 2 reveals the coverage obtained for both the positive and negative exemplars by each WordNet category in the class of Causes. Note how the category Causal-Agent offers very little coverage for the positive exemplars (i.e., most of the actual causes in that data-set are not categorized as causal-agents in WordNet), and actually offers higher coverage for the negative exemplars (making it more likely to contribute to a classification error in the case of a near-miss).

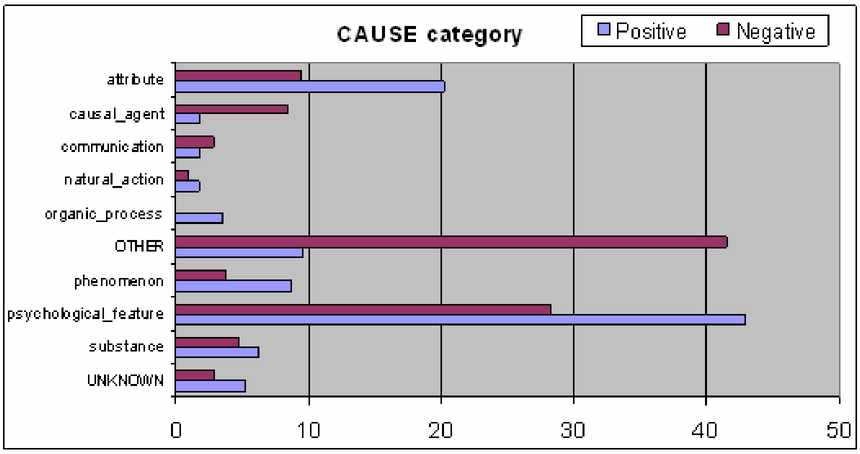

Fig. 2. The coverage (%) offered by different WordNet categories for SemEval positive and negative exemplars of the *Cause* class.

Figure 3 presents a comparable analysis for the WordNet categories that comprise the cover-set for the class of Effects. Note how the category {psychological_feature} looms large as both a Cause and an Effect in the SemEval data-set.

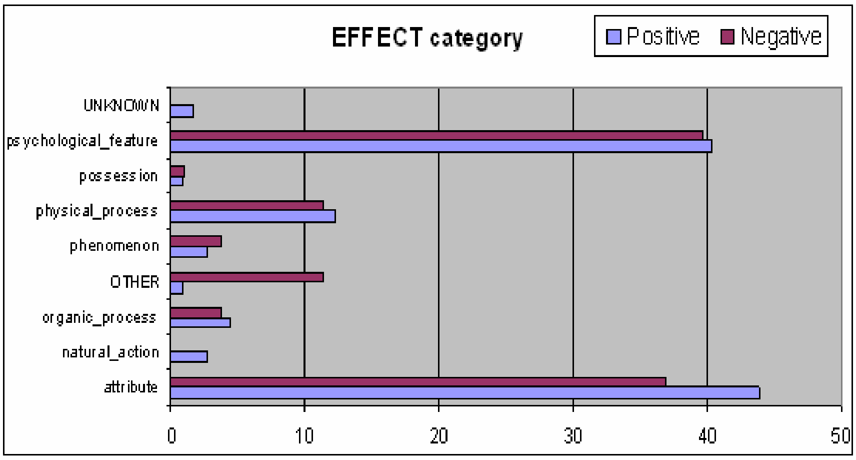

Fig. 3. The coverage (%) offered by different WordNet categories for SemEval positive and negative exemplars of the *Effect* class.

4 Defining Cause and Effect in Syntagmatic terms

Girju in [5] notes that certain lexico-syntactic patterns are indicative of causal relations in text, but that some patterns are more ambiguous than others. For instance, the patterns "NP2-causing NP1" and "NP1-caused NP2" are explicit and largely unambiguous cues to the interpretation of NP1 as a cause and NP2 as an effect. In contrast, Girju notes that "NP2-inducing NP1" and "NP2-generated NP1" are equally explicit but potentially more ambiguous patterns for identifying cause and effect in text. Nonetheless, the pattern "NP-induced NP" does occur quite frequently in large corpora, and does designate causes with high accuracy and low ambiguity. However, this triple of "NP-induced/inducing NP" produces a spare space of associations between different causes and effects, so it is more productive to consider each noun-phrase in isolation.

Thus, we look for the patterns "Noun-inducing" and "Noun-causing" in a large corpus to identify those nouns that can denote effects, as in the phrase "headache-inducing". Our corpus is the set of Google N-grams (see [1]), from which the above pairings can easily be mined. Similarly, we mine the patterns "Noun-induced" and "Noun-caused" from these n-grams to identify a large set of nouns that can denote causes, as in "caffeine-induced". In addition, we look to the patterns "-induced Noun" and "-caused Noun" to identify a further collection of possible effect nouns, and the patterns "-inducing Noun" and -causing Noun" to identify further cause nouns. In this way, we obtain 3,500+ nouns as denoting potential causes, and 4,200+ nouns as denoting potential effects. Table 1 presents the top-ranked (by frequency) causes and effects in this data, as well as the top-ranked causality pairs (i.e., cause associated with specific effect).

Table 1. Top-ranked (by frequency) *cause-effect* pairs, as well as isolated *causes* and isolated *effects*.

CAUSE-EFFECT pairs	CAUSE nouns	EFFECT nouns
(organism, disease)	Drug	apoptosis
(laser, fluorescence)	stress	disease
(noise, hearing)	radiation	cancer
(chemical, cancer)	exercise	changes
(agent, cancer)	self	cell
(exercise, asthma)	laser	increase
(collagen, arthritis)	human	activation
(bacteria, disease)	acid	asthma
(pregnancy, hypertension)	light	inhibition
(human, climate)	virus	odor

Because the Google N-grams corpus is not sense-tagged, we can only guess at the senses of the nouns in Table 1. However, if we assume that each noun is used in one of its two most frequent senses, then we can assign these nouns to various WordNet categories, as we did for the SemEval nouns in Figures 2 and 3. Following this heuristic assignment of senses, Figure 4 presents the distribution of cause nouns to different WordNet cause categories.

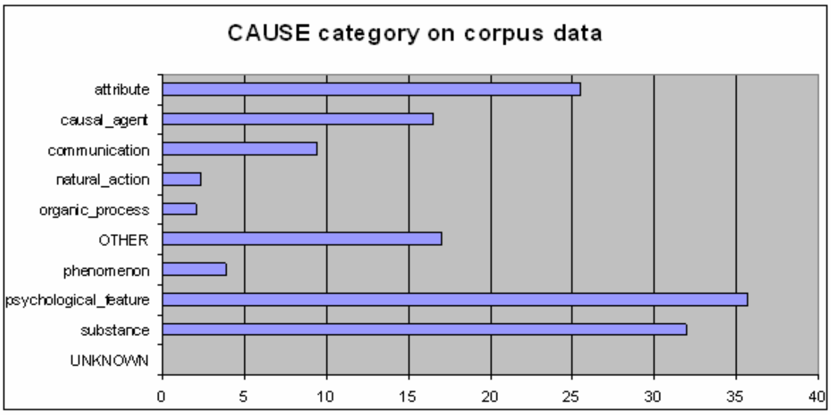

Fig. 4. The distribution of corpus-mined *cause* nouns to WordNet categories.

A comparable distribution for effect nouns is displayed in Figure 5.

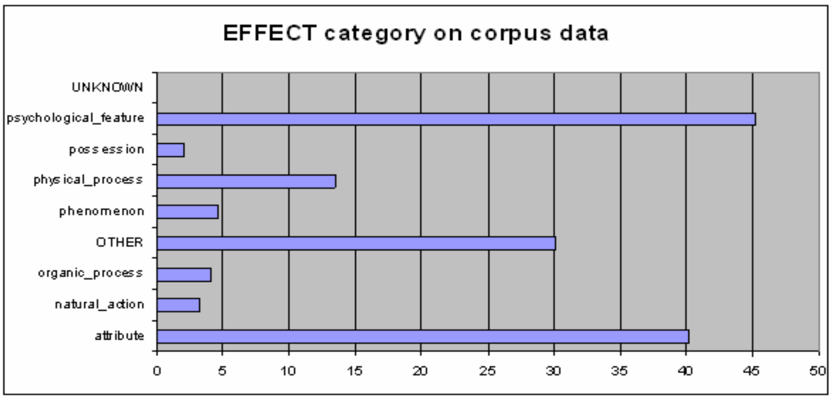

Fig. 5. The distribution of corpus-mined *effect* nouns to WordNet categories.

Because some noun senses belong to multiple categories, and because we use the two most frequent senses of each noun, the sum total of distributions in Figures 2 to 5 may exceed 100%. Note also that certain patterns are noisier than others. While "Noun-inducing" is a tight and rather unambiguous micro-context in which to recognize Noun as an effect, "-induced Noun" is more prone to leakage. For instance, "drug-induced liver failure" yields "drug" as an unambiguous cause, but mistakenly suggests "liver" as an effect. Given that "Noun-induced" is a more frequent pattern than "Noun-inducing", the set of nouns designed as effects is noisier than the set of nouns designed as causes. For this reason, the Other category in Figure 5 is more populous than the Other category in Figure 4. The most frequently misclassified nouns in the Effect class are: *protein*, *liver*, *gene*, *lung*, *acute*, *platelet*, *insulin*, *diabetic*, *skin*, *calcium*, *rat*, *cytotoxicity*, *genes*, *immune*, and *bone*.

5 Empirical results

We can test the approaches of section 3 and 4 in a variety of guises and combinations:

The WordNet-only approach (as described in section 3): a word pair <X,Y> can be classified as a Cause-Effect pairing if and only if any of the two most frequent senses of X fall under a synset in the Cause cover-set and any of the two most frequent senses of Y fall under a synset in the Effect cover-set.

The Corpus-only approach (as described in section 4): a word pair <X,Y> can be classified as a Cause-Effect pairing if and only if X is found in the set of nouns that have been identified as cause nouns (e.g., because the pattern "X-induced" was found in the corpus) and Y is found in the set of effect nouns (e.g., because the pattern "Y-inducing" or "-induced Y" was found in the corpus). In our experiments we test two different sets of corpus-mining patterns: a minimal set based on just two causation verbs, *induce* and *cause*, and an extended set comprising variations of the verbs *induce, cause, power, fuel, activate, enable, control* and *operate*.

The Hybrid approach (WordNet used in combination with corpus-derived data): a word pair <X,Y> can be classified as a Cause-Effect pairing if any of the two most frequent senses of X fall under a synset in the Cause cover-set *and* a synonym of one these two senses (i.e., any word from the same two synsets) is found in the set of corpus-derived cause nouns, and if any of the two most frequent senses of Y fall under a synset in the Effect cover-set *and* a synonym of one these two senses of Y (or Y itself) is found in the set of effect nouns. The hybrid approach is thus a logical conjunction of the WordNet and corpus approaches, but one that includes synonyms of the words X and Y, so the corpus-data of the latter is effectively smoothed and made less sparse.

Table 2 presents empirical results for each of these approaches on the SemEval cause-effect data-set and the All-true baseline which always guesses “true” (and thereby maximizes recall). Interestingly, the WordNet-only approach has the best overall performance (F-score), which accords with the observations of the SemEval organizers: the statistics show that WordNet plays an important role in the task of relation classification.

Table 2. Empirical results for cause-effect in SemEval data-set, where F = 2*P*R / (P+R).

	P	R	F	Total no
A. WordNet only approach	61.3	85	71.3	220
B. Corpus-only approach Using *{induce, cause}* patterns	54	60	62.3	220
C. Hybrid A+B approach	63.5	70	66.8	220
D. Corpus-only approach using *{induce, cause, power, fuel, activate, enable, control, operate}* patterns	51.6	83	63.6	220
E. Hybrid A+D approach	60	85	70.3	220
***All-true* baseline**	51.8	100	68.2	220

5.1 Analysis of Results

As the corpus yields a somewhat sparse and noisy data set of candidate cause and effect nouns, the corpus approach (B) that uses just *cause* and *induce* as causal markers achieves only 60% recall, with a low precision of 54%. The WordNet contribution in the Hybrid A+B approach boosts recall by 10% while also increasing precision. Recall is improved since the sparse corpus data is extrapolated by the use of WordNet synonyms; precision is also improved somewhat, over that of the WordNet-only approach (A) and the simple corpus approach (B) because WordNet's category restrictions help to filter out some noisy and misclassified effect nouns. Nonetheless, there is need for more corpus data to increase the recall of the hybrid approach even further. In the second corpus approach (D), recall is boosted by using patterns based on a broader list of causative verbs (see [9]) to identify cause and effect nouns: *{induce, cause, power, fuel, activate, enable, control, operate}*. Note that when WordNet Cause and Effect categories of (A) are used to filter noisy classifications in the hybrid approaches, this imposes a WordNet-based ceiling of 85% (i.e., the recall of A) on the recall of the hybrid approaches: the tradeoff results in a lower precision but a better F-measure overall.

Each approach in Table 2 (WordNet-alone, corpus-alone, and the combination of both) is unsupervised and does not avail of the WN sense information provided for nouns in the SemEval data-set. Our best F-measure is 71.3% and is comparable with the 72% F-measure obtained by the best performing system in the corresponding SemEval category (i.e., category A, in which competing systems do not avail of WordNet sense tags). The relatively low precision is largely explained by the fact that SemEval's negative examples are near misses rather than random examples of non-causal relationships. Our recorded precision is a lower-bound then for what one might expect on random word-pairings drawn from a real text.

6 Concluding Remarks

In this paper we presented three unsupervised approaches to the classification of causal-relations among noun-pairs: a corpus-based approach, an ontological WordNet-based approach, and a combination of both. The results achieved by these approaches on the SemEval dataset are encouraging, especially given the fact that these approaches do not apply machine-learning techniques to a training data-set. The WordNet categories which form the substance of the ontological approach, and which also contribute substantially to the combined approach, are hand-picked based on human intuitions about causality. However, a machine-learning approach to identifying these categories automatically is a topic of current research. As reflected in the superior performance of the WordNet-only approach, WordNet does have the capability to accurately represent high-level abstractions like Cause and Effect, and to do so in a non-trivial way that spans large numbers of more specific specializations.

Nonetheless, our results also bear out our initial observation that the WordNet category of Causal-agent is very weakly represented and in serious need of re-organization, at least if it is to properly serve its intended purpose. In the SemEval data analyzed here, the {causal_agent} category covers only 2% of the Cause instances in the positive exemplar set, and just 8% of the negative "near-miss" exemplars. Extension to this WordNet category can clearly be performed using intuition-guided ontological-engineering as well as corpus-based discovery. Based on our results then, we might ask which WordNet concepts should be included under the newly organized umbrella term of Causal-Agent, and under a new category, Causal-Patient? We suggest the word senses that satisfy approach E will make excellent candidates to populate these categories.

We next plan to extend the general approach described here to other classes of semantic relation, such as Content-Container, Part-Whole and Tool-Purpose, since these too combine a strong ontological dimension to their meaning with a strong usage-based (i.e., corpus-based) dimension. Overall, our results confirm that WordNet has a significantly useful role to play in the detection of semantic relations in text, but detection would be more efficient if WordNet could provide more insightful ontological classifications of the concepts underlying these relations. These ontological insights will come from using the existing structures of WordNet to hypothesize about, and filter, large quantities of relevant usage data in a corpus.

References

1. Brants, T., Franz, A.: Web 1t 5-gram version 1. Linguistic Data Consortium (2006)
2. Butnariu, C., Veale, T.: A hybrid model for detecting semantic relations between noun pairs in text. In: Proceedings of SemEval 2007, the 4th International Workshop on Semantic Evaluations. ACL 2007 (2007)
3. Cole, S., Royal, M., Valorta, M., Huhns, M., Bowles, J.: A Lightweight Tool for Automatically Extracting Causal Relationships from Text. In: Proceedings of IEEE (2006)
4. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA (1998)
5. Girju, R.: Text Mining for Semantic Relations. Ph.D. Dissertation, University of Texas at Dallas (2002)
6. Girju, R., Moldovan, M.: Text mining for causal relations. In: Proceedings of the FLAIRS Conference, pp. 360--364 (2002)
7. Girju, R., Nakov, P., Nastase, V., Szpakowicz, S., Turney, P.: SemEval 2007 Task 04: Classification of Semantic Relations between Nominals. In: Proceedings of SemEval 2007, the 4th International Workshop on Semantic Evaluations. ACL 2007 (2007)
8. Hearst, M.: Automated Discovery of WordNet Relations. In: WordNet: An Electronic Lexical Database and Some of its Applications. MIT Press (1998)
9. Khoo, C., Kornfilt, J., Oddy, R., Myaeng, S.H.: Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing. Literary & Linguistic Computing, vol. 13(4), pp. 177--186 (1998)
10. Lewis, D.: Evaluating text categorization. In: Proceedings of the Speech and Natural Language Workshop, pp. 312--318. Asilomar (1991)
11. Nastase, V.: Semantic Relations Across Syntactic Levels. Ph.D. Dissertation, University of Ottawa (2003)
12. Veale, T., Hao, Y.: A context-sensitive framework for lexical ontologies. The Knowledge Engineering Review Journal. Cambridge University Press (in press) (2006)