



# **Heuristics for Recovery from Residual Ambiguity and Incongruity in the Semantic Interpretations of Texts**

**Sergei Nirenburg, Stephen Beale and Marjorie McShane**

**Working Paper 02-05**

**March 8, 2005**

**Institute for Language and Information Technologies  
University of Maryland Baltimore County**

## Abstract

In this paper we present the heuristic algorithms at the core of a text meaning analyzer based on the ontological-semantic approach (e.g., Nirenburg and Raskin 2004). The full analyzer takes as input natural language texts and through many levels of analysis produces text meaning representations (TMRs) formulated in a specially developed ontology-based metalanguage that covers semantic as well as discourse and pragmatic information. We concentrate here on the process of deriving basic semantic “who-did-what-to-whom” dependency structure from a text. In presenting the material, we concentrate on the motivation for the introduction of specific algorithms (methods) and on their heuristic content.

## Introduction

The ability to interpret – extract, represent and manipulate – meaning in natural languages is the holy grail of computational linguistics and natural language processing (NLP). The problem is notoriously difficult. Any serious system building and experimentation in this area requires large-scale knowledge acquisition efforts that cannot be fully automated within the state of the art. It is for these reasons that many researchers avoid working on broad-coverage text meaning interpretation. We believe that it is not worthwhile to put a moratorium on such work until heuristic knowledge acquisition becomes fully automatic. (Indeed, this latter capability may never be truly attained.) So, while work on the enabling capabilities for text meaning analysis (e.g., tokenization, morphological and syntactic analysis, studies in lexical semantics, and on automating knowledge acquisition using empirical stochastic methods etc.) are important, working on comprehensive meaning analysis directly is no less crucial.

In this paper we present the heuristic algorithms at the core of a text meaning analyzer based on the ontological-semantic approach (e.g., Nirenburg and Raskin 2004). The full analyzer takes as input natural language texts<sup>1</sup> and through many levels of analysis produces text meaning representations (TMRs) formulated in a specially developed ontology-based metalanguage that covers semantic as well as discourse and pragmatic information.

We concentrate here on the process of deriving basic semantic “who-did-what-to-whom” dependency structures from a text. Neither the non-semantic steps in the analysis process nor the static knowledge resources used and their acquisition processes are reported in this paper. See Nirenburg and Raskin (2004), Chapters 6-8 for an overview. The currently available static knowledge resources relevant for

---

<sup>1</sup> At present, we concentrate on English but in the past we developed an extensive Spanish analysis system and smaller-coverage systems for Chinese, Arabic and Persian.

the discussion in this paper include an ontology of about 6,000 concepts with, on average, 16 properties defined for each and a lexicon of about 12,000 lexical units of English plus a special lexicon of named entities of about 350,000 units.

In presenting the material, we concentrate on the motivation for the introduction of specific algorithms (methods) and on their heuristic content. All the algorithms described in the paper are implemented and have been used in the analyzer. We are currently conducting a glass-box evaluation of the analyzer.<sup>2</sup>

## Basic Semantic Dependency Building

The stage is set for the operation of the core semantic dependency building process when a) the input text has been processed for named entities, dates, time, punctuation, part of speech tags and syntactic dependencies; b) the lexicon entries for all the senses of all the words and phrases identified in the input have been instantiated in the workspace for the nascent TMR; and c) the heads of the propositions in the input text have been determined using a procedure that relies largely on the syntactic dependency structure as its heuristic source. The goal of the basic semantic dependency construction algorithm is to derive a) the most appropriate combination of word and phrase senses in the workspace and b) the most appropriate semantic dependency structure that takes into account the entire input space. The semantics-oriented zones of lexicon entries in our approach describe lexical meaning in terms of the same metalanguage that is used for defining the ontology, the TMRs and remembered instances of ontological concepts recorded in a special database called fact repository (FR).<sup>3</sup>

For the purposes of semantic interpretation at the proposition level, the most important inter-concept relations in the ontology are **case roles**, properties that relate an event-type concept to its typical participants – e.g., agents, beneficiaries, themes, etc.<sup>4</sup> The semantic zones of the lexicon

---

<sup>2</sup> Should the paper be accepted for presentation, we would be able to include relevant evaluation results in the final copy.

<sup>3</sup> The presence of a language-independent metalanguage separates our approach from other approaches, e.g., Gildea and Jurafsky (2000), Mohit and Narayanan’s (2003), Harabagiu and Maiorano (2000) or Schubert (2002) who basically operate with uninterpreted English text strings. Using a metalanguage not only allows for significant resource savings in multilingual tasks but also better supports both general and referential reasoning.

<sup>4</sup> There are many terminological variants for the notion of case role and a variety of inventories of actual case roles, with subtly (and sometimes not so subtly) different semantics.

entries instantiated in nascent TMRs contain case role constraints, and the initial processing step in lexical disambiguation and semantic dependency construction involves matching these constraints (also known as selectional restrictions). The simplest algorithm for processing selectional restrictions involves matching the constraints listed with every case role of every word sense of the proposition head with every available candidate argument, that is every word sense (lexical meaning) of every noun phrase head in the input clause.

Consider a toy sample input *The baker makes cookies*. Our system’s current lexicon has 37 senses of the verb *make*, though all but one of them are “light” which means that they describe idiomatic or semi-idiomatic constructions such as *make amends*, *make a turn* or *make friends*. Each of the senses of *make* will have constraints on AGENT<sup>5</sup> and THEME listed in the lexicon. The “literal” sense of *make* (the one actually relevant for this input) will be interpreted by the ontological concept MANUFACTURING-ACTIVITY with the ontological concept HUMAN listed as the selectional constraint in its AGENT case role and ARTIFACT in its THEME case role (most of the other senses of *make* will list much narrower constraints, often even narrower than those in a single ontological concept<sup>6</sup>). *Baker* and *cookie* have one sense each. The meaning of *baker* is expressed by BAKER, a descendant of HUMAN in the ontological graph, while the meaning of *cookie* is expressed by BAKED-FOOD, a descendant of ARTIFACT. The simplified algorithm above matches the selectional restrictions on case roles in each of the senses of *make* against the senses of the nouns. It returns a single appropriate TMR fragment: an instance

```
MANUFACTURING-ACTIVITY
  AGENT BAKER
  THEME BAKED-FOOD
```

The basic algorithm is usually fortified with additional external knowledge. Thus, to enhance efficiency, the selection of candidate arguments for specific case roles can be guided by a projection (linking) from the previously established syntactic dependency. For example, a grammatical subject often links to the AGENT case role, so that the meaning of the subject of a clause will be the prime candidate for filling the AGENT slot of the corresponding propo-

<sup>5</sup> Ontological concepts (objects, events and properties) are written in SMALL CAPS. English lexemes and examples, in *italics*.

<sup>6</sup> In our approach, ontological concepts used in lexicon entries can introduce additional constraints. This enhances the expressive power of the semantic metalanguage without creating new concepts. From the point of view of content, this device is intended to preserve the interlingual nature of the ontology, as it were, relegating the quirks of lexical meaning to the lexicons of specific languages.

sition head. Linking information is stored in the lexicon entries for argument taking words. In this paper we concentrate on building semantic dependencies for proposition heads and their arguments; we will report on the algorithms used in our analyzer to derive meanings of phrasal constituents (e.g., noun phrases) elsewhere.

## Recovery

There are three basic outcomes of the selectional restriction matching process – one, more than one or no acceptable solutions may be found. If no proposition head sense has a matching argument for each of its case roles, then one faces a case of incongruity (IC). If there are more than one senses with a full complement of matching arguments, then we have a case of residual ambiguity (RA). If matching yields a single answer, the process is fully successful<sup>7</sup>. Cases of IC and RA trigger special “recovery” procedures, seeking to relax (IC) or tighten (RA) the original constraints. In both cases, an attempt should be made to modify the lexical meaning of the input words as little as necessary to obtain satisfactory results. Naturally, “recovered” matches are assigned a lower quality ranking than exact matches.

We have considered a broad variety of “recovery” means, and have implemented many of them in our current system. Specifically, our repair procedures use the following stored or dynamically derived heuristics:

1. multivalued selectional restrictions (both IC and RA);
2. lateral constraints (primarily RA);
3. constraint satisfaction on properties other than case roles (primarily RA);
4. dynamic context-sensitive adjustment of selectional restrictions (primarily RA);
5. co-occurrence preferences in the space of the TMR for the entire text (primarily RA);
6. preferences based on weighted distances among ontological concepts (both IC and RA).

Figure 1 very briefly summarizes the processing in our semantic analyzer. In what follows we describe the recovery procedures.

<sup>7</sup> Though we recognize the possibility of a “false positive”; that is, a single, yet incorrect answer.

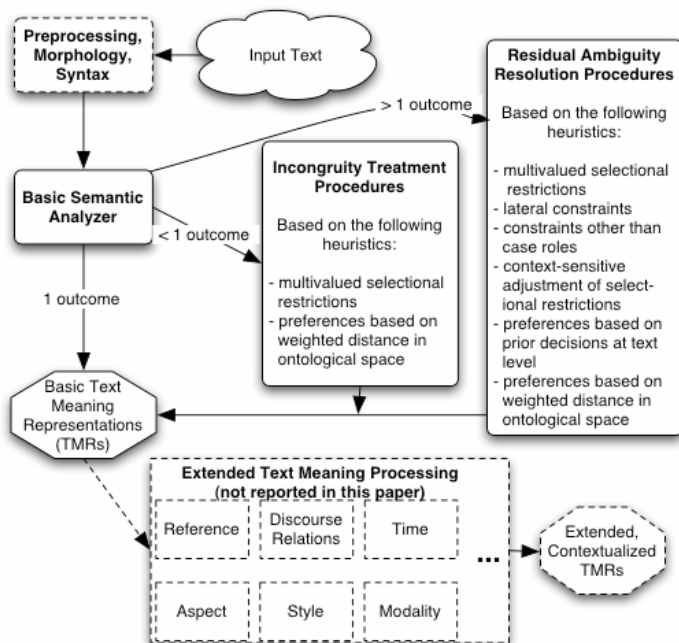


Figure 1. The overall architecture of the ontological-semantic analyzer. Means of recovery from incongruity and residual ambiguity are presented in detail. Elements in dotted-line boxes are not discussed in this paper.

Unlike any system we know, our approach uses **multi-valued selectional restrictions** (Heuristic 1). Three levels (or “facets”) are defined for each ontological property – DEFAULT (the strictest), SEM and RELAXABLE-TO (the weakest). For example, the DEFAULT constraint on the AGENT of the ontological concept BAKE is the disjunctive set (BAKER CHEF); the SEM constraint is HUMAN (an ontological ancestor of BAKER). RELAXABLE-TO can be used to encode typical metonymies (we also have a dynamic metonymy processing mechanism described below). For example, the RELAXABLE-TO constraint for the AGENT of BAKE allows a selectional restriction match on ORGANIZATION (to allow for the metonymical usage as in *This restaurant bakes excellent desserts*). RELAXABLE-TO constraints can also be used to record non-typical, yet still literal meanings (usually referred to as cases of sortal incongruity), such as the THEME in *The child baked the action figure* (since any PHYSICAL-OBJECT (of appropriate size - see below) can be put into an oven and baked). The initial basic algorithm only uses DEFAULT and SEM constraints for matching selectional restrictions, giving a “bonus” in the ranking to a match on the former: In RA cases, matches on DEFAULT constraints are preferred over those on SEM. RELAXABLE-TO is invoked in IC cases to relax the original constraints, and a match on this facet incurs a ranking penalty.

**Lateral constraints** (Heuristic 2) hold between the fillers of the different case roles of a predicate. They are used in RA cases. Consider the input *The doctor performed the operation*. *Operation* is ambiguous between PERFORM-SURGERY and MILITARY-ACTION (among others; these two are the relevant mappings for this sentence). The routine

“vertical” matching of selectional restrictions on the proposition head with the meanings of appropriate case roles does not rule out one of these senses. However, the fillers of the AGENT-OF property of DOCTOR include, among other EVENTS, PERFORM-SURGERY. This will allow us to apply a lateral constraint between the meanings of two case roles to choose the PERFORM-SURGERY meaning of *operation* with a relatively high confidence, since the resulting TMR will have DOCTOR as the AGENT of the PERFORM-SURGERY event – the very situation encoded in the knowledge resources.<sup>8</sup>

Consider now the sentence *The farmer put the pig in the pen*. *Pen* is ambiguous between a WRITING-IMPLEMENT sense and an ENCLOSURE sense. The initial matching of vertical selectional restrictions will fail and lead to an RA case. Indeed, the appropriate sense of *put* maps (after disambiguation) into CHANGE-POSITION; the ontology constrains its AGENT case role filler to ANIMAL, its THEME, to PHYSICAL-OBJECT and its DESTINATION, (to be filled by the meaning of *pen*), to PLACE.<sup>9</sup> Unfortunately, this latter constraint does not help to disambiguate *pen*, since both its senses map to descendants of PHYSICAL-OBJECT and will equally satisfy the DESTINATION: PLACE constraint. Lateral constraints do not work either (they would if the ontology recorded that ENCLOSURE is a typical DESTINATION or LOCATION of PIG). The system then attempts to use lateral constraints across **non-case-role property fillers** (heuristic 3).

In this case, we compare the typical size of PIG with the typical sizes of a WRITING-IMPLEMENT and a ENCLOSURE. How do we know that this is exactly the comparison we need? This constraint check on the relative size of the THEME of an EVENT and its DESTINATION is triggered by a general inference rule associated with the spatial sense of *in*. The typical size (actually, LENGTH; we also use typical measurements for HEIGHT, but fillers for that property are more sparsely acquired in the current ontology) of PIG is 1 meter. The typical size of WRITING-IMPLEMENT is 0.1 meters. The typical size for ENCLOSURE is weakly constrained to a broad range of between 10 and 1000 meters. In this example, the preferred outcome is clear. In general, this kind of word sense disambiguating capability is also claimed by approaches that use standalone (non-ontological) semantic markers – for example the “prag-

<sup>8</sup> Of course, there can always be unexpected contexts in which any listed constraints do not work, even when relaxed. Such contexts (the grist, for example, for the garden-path example mill) pop up more frequently as inputs become shorter. Indeed, most garden path examples are a single sentence. Text-level heuristics, in TMRs and ontological scripts (typical complex events) work better on longer inputs.

<sup>9</sup> The fact that CHANGE-POSSESSION requires DESTINATION disambiguates the PP attachment ambiguity.

matic codes” of LDOCE as well as by statistical corpus-based techniques. We use such techniques when lateral constraints cannot be used due to lacunae in the ontology.<sup>10</sup>

Our algorithm for treating RA cases by finding a more specific proposition head satisfying all selectional restrictions than the one used in the input has been reported in Mahesh et al. (1997). The algorithm effectively **dynamically tightens the selectional restrictions** (Heuristic 4) for inputs that were originally more vague than possible. For example, if among the senses of the word *range* in the lexicon two are closely related and refer to *STOVE* and *OVEN*, respectively, then the input *The cake was made using a range* will lead to an RA case. The algorithm then checks whether any of the children of *MANUFACTURING-ACTIVITY* (the meaning of *make*) has *CAKE* (or as close an ancestor as possible) among the legal fillers of its *THEME* and either *OVEN* or *STOVE* (but not both) among the fillers of its *INSTRUMENT*. If such a “lowest common anchor” is found, then the proposition head is effectively changed to it (thus, the system “corrects” English usage of the author), and the disambiguated semantic dependency is returned. In the above example, *BAKE* is such a lowest common anchor for *CAKE* and *OVEN* (but not for *CAKE* and *STOVE*). Using *BAKE* instead of *MANUFACTURING-ACTIVITY* disambiguates the input.

Let us at this point venture beyond single-sentence examples (Heuristic 5). Consider the following text fragment: *The doctor was called to the hospital. The operation was performed at IAM. Doctor and hospital set up a strong medical context for the passage. This can be used to disambiguate operation (either a SURGERY or a MILITARY-ACTIVITY) in the second sentence. This is similar to the lateral constraint example presented in the above; however, the context for lateral constraint processing is local to a proposition, and was only determined by overt links in the ontology. Here, the mechanism is more general.*

We have developed static knowledge resources that identify typical concepts used in specific domains. A list of concepts in the domain of medicine, for example, was built automatically by starting out with several obvious medically-related concepts, such as *DOCTOR* and *HOSPITAL*, and then recursively adding concepts that can be reached from the seed concepts by traversing links corresponding to a subset of their ontological properties. The links traversed centrally include case roles (*AGENT*, *THEME*, *BENEFICIARY*, *LOCATION*, etc.) as well as such important properties as *PRECONDITION*, *EFFECT* and *PURPOSE*. The resulting list of concepts describing the domain was edited by hand. It includes events such as *PERFORM-SURGERY*, typical actors in a medical domain such as *NURSE*, typical locations

such as *EMERGENCY-ROOM* and typical *THEMES*, e.g., different body parts.

While analyzing a text, a rough guess as to its general topic can be generated by matching concepts in the TMRs to the lists of concepts describing a domain. Each candidate domain receives a “context-score” between 0.0 and 1.0 which is continually updated as new TMRs are processed. A score of 1.0 indicates that the text is certainly in the topic area. When a case of RA is encountered, word senses that are in a domain with a high context score are preferred.

This knowledge acquisition and processing methodology seems similar to that used in less knowledge-intensive approaches, including statistical methods for determining a context (Yarowsky 1992) and methods that use a human-oriented resources such as WordNet (Moldovan & Novitschi 2002). The central difference is that we use disambiguated concepts from TMRs of previously analyzed sentences as the input to the context-identification procedures. Also, the topic-indicating clues are knowledge-based and are determined by their inclusion in detailed descriptions of complex events, ontological scripts, related to the topic.

We use a similar approach in a more dynamic fashion. The input: *ABC Corporation filed for bankruptcy. The court was located in Baltimore* will trigger an instance of the bankruptcy script stored in the ontology. This script contains information about sub-events common in a bankruptcy context, as well as typical participants and locations, such as courts. Once a script is triggered, all the information in it can be used to disambiguate the text. In this case, the *COURTROOM* sense of *court* can be preferred over the *GAME-COURT* sense. Using scripts in this way is a more dynamic application than the use of the static lists created in the first text-level method described above. Using scripts is also a more reliable method because it is more constrained: it only uses information within the script and does not attempt to extend that information by traversing all the related links in the ontology.

Consider the input: *George W. Bush arrived in London this afternoon. The president will attend a State Dinner at Buckingham palace tomorrow evening. George W. Bush* will be linked to a fact in our fact repository (a database of remembered instances of ontological events and objects) which contains information about him, in particular that he is the *PRESIDENT* of the United States. In our lexicon, *president* is ambiguous between the *PRESIDENT* (of a republic) and *PRESIDENT-CORPORATION* readings. Our existing reference resolution procedures (not addressed in this paper) will actually be able to resolve this case of definite reference automatically. However, we are also able to use information in the fact repository to develop groups of facts relating to a topical domain, similar to that described above. For example, events, people and places associated with George W. Bush can be used as additional heuristics

---

<sup>10</sup> Statistical routines in our approach are embedded in a semantic analyzer and serve as providers of heuristics for it.

for disambiguation. Since fact repository elements are stored as disambiguated TMR fragments, they can be used in a manner similar to the knowledge used in the previous two text-level methods described above.

The last heuristic described in this paper is **preferences based on weighted distances among ontological concepts** (Heuristic 6). In the case where there are no constraints recorded in the knowledge sources (for example, the constraints of Heuristics 1, 2 and 3), we use the Ontosearch program (Onyshkevych 1997) to determine the strength of the semantic association between any two constituents in a proposition. Ontosearch finds the ontological path (chain of relations) between any two concepts, and calculates the weight (score) of each path that reflects the strength of the association between two concepts. The cumulative score is a function of the length of the path and of the cost of traversing a particular relation link. Both the traversal costs and the combination of evidence function were induced empirically. The system was trained using simulated annealing on the 40 most frequently used ontological properties.

As applied in the semantic analyzer, Ontosearch serves as the source of heuristics for selectional restriction matching. The weighted distance between each pair of property fillers is calculated. Most often, each such pair will be composed of the filler of a case role property of the proposition head on the one hand and a candidate filler from among the word senses for propositional arguments on the other. The highest scoring set of connections within each set of candidate pairings will be returned provided all the comparisons yield at least one score above an empirically established minimum quality threshold. If no score in a set of candidates is above the threshold, the system returns the case of IC. Using the threshold cutoff is the dynamic counterpart of the content of the RELAXABLE-TO facet of properties that effectively defines the boundaries of constraint relaxation.

To help disambiguate the input *The hospital performed the operation*, we can examine the ontological connection between HOSPITAL and the two senses of *operation*: PERFORM-SURGERY and MILITARY-ACTIVITY. For HOSPITAL/PERFORM-SURGERY, Ontosearch returns the following chain:

```

ONTOSEARCH(HOSPITAL, PERFORM-SURGERY) →
  HOSPITAL
    IS-A MEDICAL-BUILDING
  MEDICAL-BUILDING
    LOCATION-OF PERFORM-SURGERY

```

In other words, the ontology contains information that a hospital is a place where surgeries are performed. The score for the HOSPITAL/MILITARY-ACTIVITY relationship is much lower. In this way we can determine how closely

two constituents of a proposition are related in the ontology. Of course this score reflecting the strength of association between constituents will receive relatively less weight than selectional restrictions of heuristics 1-3, but it is a useful, knowledge-based method of determining the most likely reading in cases of RA.

We also utilize the Ontosearch program to **dynamically process metonymy**. In fact, processing metonymies is simply subsumed in our approach by the use of Ontosearch for any case of IC. Indeed, Ontosearch has been statistically trained for use in helping to resolve IC cases of any provenance, not only metonymy. Consider the sentence *The hospital performed the surgery*. The literal constraint on the AGENT of a PERFORM-SURGERY is a DOCTOR (RELAXABLE-TO HUMAN). HOSPITAL does not literally meet this constraint. A metonymic resolution to this problem asks, "How can a HOSPITAL stand for a DOCTOR?" We can query Ontosearch to find the answer:

```

ONTOSEARCH(HOSPITAL, DOCTOR) →
  HOSPITAL
    IS-A MEDICAL-BUILDING
  MEDICAL-BUILDING
    LOCATION-OF PERFORM-SURGERY
  PERFORM-SURGERY
    AGENT DOCTOR

```

The score returned by Ontosearch can be used to judge between potentially metonymic readings of different word senses. In this case, we are able to add the inference to the resulting TMR that there was actually a doctor that performed the surgery, even though that information was not present in the input sentence. Another side effect of this process can be semi-automatic augmentation of the lexicon. For example, HOSPITAL could be actually added to the RELAXABLE-TO facet of the AGENT of PERFORM-SURGERY.

## Discussion

Even after all the available repair procedures have been used, the outcome can still remain unsatisfactory. There can be several reasons for this state of affairs. For example, the input may contain a case of genuine sortal incongruity. In the present system, the input *The baby ate a newspaper* will lead to an unresolvable case of IC because the meaning of *newspaper* and the selectional restriction on the THEME of INGEST, do not match even on the RELAXABLE-TO facet, while the dynamic relaxation using Ontosearch returns a result below the quality threshold. Of course, if the above input is deemed meaningful, the corresponding constraints can be added to the lexicon, so that the next time a similar input will yield a successful analysis. The input may also involve metaphor or other non-

literal language that could not be treated the way metonymies or, more generally, IC cases are treated in the system. In fact, even before attempting to treat metaphor, the system must detect it and distinguish it from an IC case.

A different kind of problem arises due to the presence of unknown words in the input: unknown words do not violate selectional restrictions for the simple reason that for them no static constraints are listed in the lexicon. At this time, our treatment of unknown words is rather simple: we treat them all as nouns<sup>11</sup> and check selectional restrictions on them “one-directionally” that is, unknown words match any restrictions in their proposition head.

Finally, the input can contain legitimate but unrecorded senses of words already in the lexicon. We try our best to avoid this eventuality through careful lexicon acquisition, as a dynamic solution for this problem is difficult: similarly to the case of metaphor, it is not easy automatically even to recognize that an extra sense is, in fact, needed.

While all of the procedures and heuristics described in this paper have been implemented (in the case of the static knowledge, there is, of course, a continual need for further acquisition), we have not yet experimented with different control structures for the entire recovery mechanism. This will be part of our future work.

As will be very important work on evaluating the system. Limited-extent proposition-level evaluation shows that with only multivalued selectional restrictions involved in the recovery process and given correct preprocessing results, the system disambiguates 87% of word sense ambiguity and establishes correct semantic dependency in 90% of the cases on inputs where average sentence length was over 25 words, with an average of 3.5 senses per word and an average of 1.33 unknown words per sentence. We plan to extend the evaluation to include the full complement of the recovery procedures. We will report the results of this evaluation in the final version of this paper.

## References

- Gildea, D. and D. Jurafsky. 2000. Automatic Labeling of Semantic Roles. Proceedings ACL-00, Hong Kong.
- Harabagiu, S. and S. Maiorano. 2000. Acquisition of Linguistic Patterns for Knowledge-Based Information Extraction. Proceedings of LREC-2000. Athens, Greece.
- Mahesh, K., S. Nirenburg, and S. Beale. 1997. If You Have It, Flaunt It: Using Full Ontological Knowledge for Word Sense Disambiguation. Proceedings of TMI-97, Santa Fe, NM.

---

<sup>11</sup> This approximation is, actually, reasonable: it has been estimated that over 85% of words in a language are nouns; also, our lexicon has a much higher proportion of verbs – their entries were acquired first because argument-taking entities are more important for the support of lexical disambiguation than other lexical elements.

Mohit, B. and S. Narayanan. 2003. Semantic Extraction with Wide-Coverage Lexical Resources. In Proceedings of HLT/NAACL 2003, Edmonton, Canada.

Moldovan, D. and A. Novischi. 2002. Lexical Chains for Question Answering. Proceedings of COLING 2002. Taipei, Taiwan.

Nirenburg S. and V. Raskin. 2004. *Ontological Semantics*. Cambridge, MA: MIT Press.

Onyshkevych, B. 1997. An Ontological-Semantic Framework for Text Analysis. Unpublished Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

Schubert, L.K. 2002. Can we derive general world knowledge from texts?. Proceedings of HLT-02, San Diego, CA.

Yarowsky, D. 1992. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. Proceedings of COLING-92, Nantes, France.