Resolving syntactic ambiguities with lexico-semantic patterns: an analogy-based approach(*)

Simonetta Montemagni
Par.O.La sas
via del Borghetto 35
Pisa, 56100, ITALY
simo@ilc.pi.cnr.it

Stefano Federici
Par.O.La sas
via del Borghetto 35
Pisa, 56100, ITALY
stefano@ilc.pi.cnr.it

Vito Pirrelli
ILC-CNR
via della Faggiola 32
Pisa, 56126, ITALY
vito@ilc.pi.cnr.it

Abstract

A system for the resolution of syntactic ambiguities is illustrated which operates on morpho-syntactically ambiguous subject-object assignments in Italian and tries to find the most likely analysis on the basis of the evidence contained in a knowledge base of linguistic data automatically extracted from on-line resources. The system works on the basis of a set of straightforward analogy-based principles. Its performance on a substantial corpus of test data extracted from real texts is described.

1 Introduction

In this paper, a system for the resolution of syntactic ambiguities is illustrated: SenSOR - the SEmaNtic Subject-Object disambiguatoR - operates on morpho- syntactically ambiguous subject-object assignments in Italian and tries to find the most likely analysis by using the evidence contained in a knowledge base of linguistic data automatically extracted from machine readable dictionaries (MRDs), both taxonomic information and example sentences. The system works on the basis of a set of straightforward analogy-based principles which have been used for a wide range of NLP applications (Pirrelli et al., 1992; Montemagni et al., 1994; Federici et al., 1996). Both inherent semantic properties of words (as embodied by taxonomical relationships) and word distributional properties (as attested in example sentences) are exploited as a clue to the most likely Subject-Object Assignment (SOA).

We start with an illustration of the parsing problem, to move on to a consideration of the nature of the lexico-semantic knowledge usable for its solution. SenSOR's Knowledge Base (KB) is then described, together with the function which projects ambiguous SOAs onto KB in the search for the best candidate analogue. Two different tests of SenSOR's performance are illustrated and discussed in some detail. Finally, further improvements are sketched and other possible applications of the system envisaged.

2 The problem

A crucial problem in parsing Italian is the assignment of subject and object relations to sentence constituents. It is often the case that grammatical relations cannot be assigned unambiguously on the basis of morpho-syntactic information only: in the sentence *il bambino legge il libro* 'the child reads the book' agreement information is not decisive for SOA since both nominal constituents agree with the verb. On the other hand, word order information cannot be relied on conclusively due to the freedom allowed in the ordering of sentence constituents in Italian, where virtually all permutations of verb, subject and object are possible.

The ambiguities stemming from this freedom are ubiquitous and represent a problem for any NLP system dealing with Italian, a problem to whose resolution a wide variety of factors, both linguistic (i.e. phonological, morphological, syntactic, lexico-semantic and pragmatic) and extra-linguistic (i.e. based on world knowledge), contributes. Here, we concentrate on how morpho-syntactically ambiguous SOAs can be solved on the basis of lexico-semantic knowledge; in particular, the focus is on the lexico-semantic restrictions that a verb or a noun imposes on its context.

3 On the nature of lexico-semantic knowledge

The lexico-semantic knowledge used for our purposes consists of typical Verb-Subject/Object (VSO) co-occurrence patterns, automatically acquired from MRDs, whose single elements are expressed as individual words. These patterns are of a heterogeneous nature and represent both literal and figurative (metonymical or metaphorical) usages: for instance, the set of typical objects of the verb *bere* 'drink' ranges from *caffè* 'coffee' to *bicchiere* 'glass' and *benzina* 'gasoline'. They correspond to more or less constrained word combinations which include idiosyncratic collocations (as in *commettere-assassinio* 'commit-murder', where *assassinio* cannot be replaced by any semantically equivalent word such as *uccisione* 'killing') as well as regular co-occurrence restrictions (as in *aprire-porta* 'open-door', where *porta* is a typical instance of the class of "openable" objects).

Such a typology of data raises the question of whether and how well patterns acquired from dictionaries and expressed in terms of individual words can be used for disambiguation purposes. In the recent literature, Tsujii et al. (1992) and Hindle and Rooth (1993) among others use pure word co-occurrence information acquired from textual corpora for dealing with parsing problems such as prepositional phrase attachment. Lexical acquisition is carried out on large textual corpora by means of statistical techniques, whereby each pair of co-occurring words is supplied with a measure of their association strength.

Clearly, when it comes to data acquired from dictionaries it makes no sense to rely on measures of association strength, as each pattern is usually attested only once. Moreover, the overall number of different extracted patterns is certainly smaller than if the acquisition were performed on large textual corpora. This makes the issue of generalisation over individual patterns more crucial than with data acquired from textual corpora.

Traditional example-based techniques address this issue by using a thesaurus to calculate the match between the input to be disambiguated and the known examples. Yet, the thesaurus captures only certain kinds of relationships between words, namely synonymy and hyperonymy, relations which do not exhaust the many salient distinctions affecting word co-occurrence. This point has been raised in the example-based literature by Nirenburg et al. (1993) and Uramoto (1994): in order to complement thesaural relationships, Nirenburg et al. resort to morphological information and antonyms, whereas Uramoto uses "conjunctive relationships" acquired from corpora (i.e. words, such as sequences of actions, which appear together in coordinated structures).

Sets of semantically similar words can also be inferred from their distribution (Harris, 1968): each noun can be characterised according to the verbs it appears with as a subject or object; similarly, each verb can be classified on the basis of the subject or object nouns it co-occurs with. Words occurring in the same context can be classified as semantically similar, although the similarity in this case is grounded on some covert properties, as shown by Hindle (1990) who derived semantically related classes of nouns from their syntactic distribution. In the example-based literature, a context-based approach is adopted by Uramoto (1994) but is restricted to the treatment of unknown words only.

Summing up, patterns automatically extracted from MRDs are heterogeneous in their nature and constitute a relatively small collection of data which need be generalised over somehow. Clearly, any self-learning algorithm trained on them has to proceed with care, through stepwise cautious inferences rather than sweeping generalisations. An appropriate inferential strategy should exploit both inherent semantic properties of words (as reflected by thesaural relationships) and distributional similarities.

4 Resolution of ambiguous SOAs

SenSOR is a specialised version of a general purpose language-learning system (Federici, 1991; Pirrelli, 1993; Federici et al., 1994). Analogy is the fundamental principle which lies behind its functioning and architecture. Broadly speaking, generalisation by analogy is defined as the inferential machinery through which an unknown linguistic object (the "target object" or TO) is seen as an analogue to already known objects (the "base objects" or BOs) so that whatever piece of (linguistic) information is acquired about the latter can be used to make predictions about the former too. For the present purposes, this means that an ambiguous SOA is solved on the basis of its analogy to already familiar sentences whose grammatical relations are already known.

SenSOR's inferential routine requires: i) a structured data set of BOs constituting KB; ii) a TO to be interpreted; iii) a best-analogue(s) function projecting TO onto KB for the best analogue(s) to be selected.

4.1 The internal architecture of KB

The internal organisation of KB plays a crucial role in the inferential routine of the system.

4.1.1 Representation of base objects

Let us first consider how the VSO patterns acquired from a dictionary source are formalised and stored. In the current KB all patterns consist of two elements: a Verb and a Subject (VS pattern), or a Verb and an Object (VO pattern). Each pattern is assigned a two-level representation consisting of a set of "inherent features" describing its elements (context-independent representation), and a set of "relational features" specifying the role of each element within the lexico-semantic context described by the pattern (context-dependent representation), as exemplified in (1) below:

(1) *LEGGERE-LIBRO/O* 'READ-BOOK/O'

inherent features	relational features
Verb(*LEGGERE*)	**Predicate**(*LEGGERE*)
Noun(*LIBRO*)	**Object**(*LIBRO*)

In the top line, an identifying label is provided (*LEGGERE-LIBRO/O*), where "/O" (short for "Object") specifies the syntactic relation of the noun relative to the verb. In the left column, the inherent features of each element of the pattern are specified. Since (1) illustrates the simplest possible pattern, i.e. a word co-occurrence pattern, Noun and Verb slots are filled in by actually occurring lexemes only, encoded in capital letters. In the right column, containing relational features, *LEGGERE* is characterised as the "predicate" with respect to *LIBRO*, and *LIBRO* as the "object" of *LEGGERE*.

More complex patterns can be envisaged: e.g. taxonomical information can be associated with nouns and/or verbs, as shown in (2):

(2) *LEGGERE-LIBRO/O* 'READ-BOOK/O'

inherent features	relational features
Verb(*LEGGERE interpretare/H*)	**Predicate**(*LEGGERE interpretare/H*)
Noun(*LIBRO pubblicazione/H*)	**Object**(*LIBRO pubblicazione/H*)

Taxonomical information is encoded in lowercase: *interpretare/H* means that the verb *interpretare* 'interpret' is the H(yperonym) of *LEGGERE*. More generally, an "/X" suffixing a lowercase string specifies the sort of semantic relation (X) linking the suffixed string with the actual lexeme (in uppercase).

4.1.2 Core patterns

The patterns considered so far do not exhaust the typology of linguistic information possibly stored in KB: more abstract patterns, generalising over actually attested ones, are also stored on a par with actual co-occurrence patterns, to be used by the inferential routine of the system. These abstract patterns, called "core patterns", contain the amount of redundant information conveyed by the attested evidence and are automatically extracted by the system through the analogy-based mechanism known as "core extraction".

A core pattern is extracted from two analogous patterns which share a certain amount of information. In practice, for an analogy between two linguistic objects to be recognised as relevant (thus triggering core extraction) a match is to be found between the elements of the pattern at both levels of inherent and relational features. For instance, a relevant analogy is found between the VO patterns in (3) and (4) below, since the two have at least one inherent feature and one relational feature in common for the same element type (noun or verb). The corresponding extracted core is shown in (5):

(3) *PRENDERE-SEDE/O* 'TAKE_UP-RESIDENCE/O'

inherent features	relational features
Verb(*PRENDERE*)	**Predicate**(*PRENDERE*)
Noun(*SEDE luogo/H*)	**Object**(*SEDE luogo/H*)

(4) *PRENDERE-STANZA/O* 'TAKE_UP-ROOM/O'

inherent features	relational features
Verb(*PRENDERE*)	**Predicate**(*PRENDERE*)
Noun(*STANZA luogo/H*)	**Object**(*STANZA luogo/H*)

(5) *PRENDERE-luogo/H/O* 'TAKE_UP-place/H/O'

inherent features	relational features
Verb(*PRENDERE*)	**Predicate**(*PRENDERE*)
Noun(*luogo/H*)	**Object**(*luogo/H*)

(5) expresses the following generalisation: when the verb *PRENDERE* co-occurs with a noun having *luogo* as a H(yperonym), then this noun is interpreted as the object. Core extraction is thus used as a step towards abstracting away from actual words in the direction of a more semantically-grounded form of analogy. Note incidentally that, for reasons of computational efficiency, core extraction does not apply to core patterns, but to attested patterns only.

4.1.3 Paradigmatic structures

Another important feature of the internal structure of KB relates to the existence of "paradigmatic structures" of nouns and verbs based on their distribution in KB patterns: nouns which are subject of the same verb exhibit, likewise objects of the same verb, a sort of semantic similarity; the same can be said of verbs which take the same subjects (or the same objects). The nature of this similarity varies from case to case and remains implicit in the different groupings. Consider the two sets of patterns below:

(6) a. *SALIRE-FUMO/S*
b. *SALIRE-MAREA/S*
c. *SALIRE-PREZZO/S*
d. *SALIRE-QN/S*
e. *SALIRE-STRADA/S*

(7) a. *ABBASSARSI-TEMPERATURA/S*
b. *ALZARSI-TEMPERATURA/S*
c. *CALARE-TEMPERATURA/S*
d. *OSCILLARE-TEMPERATURA/S*
e. *SCENDERE-TEMPERATURA/S*

The core of the VS patterns in (6) is represented by *SALIRE* (whose translation in 6a-e varies from 'rise' to 'come in', 'go up', 'climb'), or more precisely by the set of both inherent and relational features associated with it. Similarly, the core of the VS patterns in (7) above is *TEMPERATURA/S* 'temperature/S'. Let us focus now on the elements which are NOT shared in (6) and (7). *FUMO/S* 'smoke/S', *MAREA/S* 'tide/S', *PREZZO/S* 'price/S', *QN/S* 'someone/S' and *STRADA/S* 'road/S' on the one hand, *ABBASSARSI* 'fall', *ALZARSI* 'rise', *CALARE* 'drop', *OSCILLARE* 'fluctuate' and *SCENDERE* 'fall' on the other hand are in complementary distribution relative to their core (i.e. they represent parallel choices in the same context). Tables in (8) and (9) below illustrate these word associations as modelled in KB:

(8)

SALIRE	*FUMO/S*
	MAREA/S
	PREZZO/S
	QN/S
	STRADA/S

(9)

ABBASSARSI	*TEMPERATURA/S*
ALZARSI	
CALARE	
OSCILLARE	
SCENDERE	

(8) and (9) are distributional "paradigms" where the core represents the invariant element common to all patterns and elements in complementary distribution are assigned distinct "paradigmatic slots".

4.2 Best-analogue(s) function

The best-analogue(s) function projects the target object TO onto KB in the search for the best candidate analogue. It operates as follows:

a) BO(s) whose inherent features are fully contained within the set of inherent features of TO is/are selected;

b) if more than one candidate analogue is found, the analogue which is specified for the greatest number of inherent features (hereafter referred to as the "best analogue") wins out over the others;

c) if, after steps a) and b) have been taken, no best analogue is found (i.e. either there is more than one best analogue or none), then control is passed to "paradigm extension" (see infra);

d) if also paradigm extension fails to provide a unique interpretation, then TO is left ambiguous.

More concretely, given the input sentence *il bambino legge il libro* 'the child reads the book', the system will be able to identify *il libro* as the object of *legge* on the basis of the pattern *LEGGERE-LIBRO/O* in (1) and/or (2) above. This is the case of a full match, since *leggere-libro* is an already known co-occurence pattern.

Suppose now that SenSOR has to disambiguate the verb-noun combination *prendere-dimora* 'take_up-residence', where *dimora* is supplied in input with its taxonym(s) but is not attested as such in KB. The interpretation of the target expression is still possible thanks to the presence of *luogo* 'place' among the taxonyms of *dimora*; i.e. TO is interpreted as a VO pattern on the basis of the core pattern *PRENDERE-luogo/II/O* in (5).

When neither a) nor b) are viable, the most likely interpretation is yielded on the basis of distributional criteria. Paradigms (8) and (9) above can be exploited to make inferences about SOAs through paradigm extension, a process which involves the tentative correlation of two words which are not seen as co-occurring within KB. Paradigm extension is defined as follows: if an element A (whether verb or noun) shares a core with another element B (of the same type), then B is allowed to inherit the paradigmatic slots of A (if any).

Suppose that the word combination *salire-temperatura* 'rise-temperature' is to be interpreted by SenSOR. The system will entertain two mutually exclusive hypotheses, with *temperatura* as either subject or object of *salire*. Consider the first hypothesis. The paradigms of *SALIRE* and *TEMPERATURA/S* are in (8) and (9) above. Paradigm extension checks whether there exists a nonempty intersection between the paradigmatic slots of *SALIRE* and those of the verbs in the paradigm of *TEMPERATURA/S*. In order to support the *SALIRE-TEMPERATURA/S* hypothesis the system needs to find a co-occurrence pattern where at least one of the fillers of the paradigmatic slots of *SALIRE* (i.e. *FUMO/S*, *MAREA/S*, etc.) co-occurs with one of the fillers of the paradigmatic slots of *TEMPERATURA/S* (i.e. *ABBASSARSI*, *SCENDERE*, etc.) as shown in (10) below:

(10)

SALIRE	FUMO
	QN
SCENDERE	MAREA
	PREZZO
	STRADA
	TEMPERATURA

where the relevant paradigms appear to share the slots *MAREA/S*, *PREZZO/S* and *STRADA/S*. From this it follows that the hypothesis *temperatura* being the subject is justifiable through paradigm extension. As to the hypothesis with *TEMPERATURA* as an object, no supporting paradigmatic evidence is found in KB.

The same process also accounts for more irregular cases, such as the co-selection of nouns and semantically light verbs (also known in the literature as "support verbs"). Moreover, it can be used to justify cases of metaphor and metonymy (Montemagni, 1995).

Summing up, in SenSOR analogy operates both at the level of the internal organisation of KB as a whole and in the search for the best analogue at the interpretation stage. The analogy-based inferential routine is rather weak and conservative at the acquisition as well as the interpretation stage. In acquisition, core patterns are extracted by matching directly attested patterns only (as opposed to already extracted cores). In interpretation, the use of directly attested evidence is always preferred over inferred evidence, i.e. over core patterns and extended paradigms. When the system is confronted with unknown combinations of words or even with unknown words (i.e. words which are not directly attested within KB but for which taxonomical information is available), it interprets them insofar as they are analogically related to already known word combinations.

5 Performance

Two different experiments were designed to test SenSOR's performance: in the first experiment, carried out on large amounts of data, KB was constituted by pure word co-occurrence patterns; the second experiment was meant to assess how and to what extent taxonomical information improves the system's performance.

5.1 Experiment 1

Pure word co-occurrence patterns represent the minimal input configuration the system can deal with. With these patterns, the system carries out the SOA task mostly via paradigm extension.

In this experiment, KB consists of 18,585 VS/VO patterns extracted from the Collins Italian-English (1985) and Garzanti (1984) dictionaries. The test was carried out on 4,279 morpho-syntactically ambiguous SOAs elicited from real texts. The results of this experiment are reported in the table below.

	n.	abs. %	UW		KE	
			n.	rel. %	n.	rel. %
OK	2,856	66.7	22	0.8	862	30.2
WRONG	251	5.9	-	-	24	9.5
AMBIG.	1,172	27.4	611	52.1	n.a.	n.a.
TOTAL	4,279	100.0	633	14.8	886	20.0

Correctness is 66.7% and the accuracy rate (whereby TO's which SenSOR leaves ambiguous are not counted in) is 92%. As to the unresolved cases, more than a half (i.e. 611 out of 1,172) contains an unknown word (i.e. UW). Note that a few cases of unknown words (22) are successfully interpreted on the basis of cores which are specified for verb or for noun information only (e.g. intransitive verbs are attested with subjects only).

It is interesting to note that about one third of TOs correctly interpreted by SenSOR are known examples (KE), although in some cases this previous knowledge is misleading. 24 TO's, already known by the system under a certain interpretation, are in fact used in the target context under a different intepretation. To give an example, this typically occurs with verbs undergoing the causative-inchoative alternation (e.g. *affondare* 'sink') where the object of the transitive reading of the verb (e.g. *nave* 'ship') can also be the subject of its intransitive reading. In cases like these, SenSOR is bound to fail, as they cannot be resolved on the basis of lexico-semantic knowledge, but require knowledge of the general context in which the sentence is uttered.

5.2 Experiment 2

In a second experiment, KB consists of 10,307 VS/VO patterns extracted from the Collins bilingual dictionary. For testing, we randomly extracted 500 ambiguous SOAs from the test corpus used in Experiment 1. For each element of the patterns, taxonomical information is specified, thus making possible generalisations over patterns sharing the same taxonomical information.

Since words in the acquired patterns are not disambiguated as to their word sense in the specific context, all their possible immediate taxonyms are specified, only one of which is appropriate to the context described by the pattern. In spite of the introduced "noise", the performance of SenSOR shows a statistically significant improvement on the previous experiment, as illustrated in the table below.

	n.	abs. %	UW		KE	
			n.	rel. %	n.	rel. %
OK	368	73.6	62	16.8	110	34.8
WRONG	54	10.8	7	12.9	3	13.6
AMBIG.	78	15.6	64	82.0	n.a.	n.a.
TOTAL	500	100.0	133	26.6	113	22.6

Correctness rises to 73.6% and the accuracy rate slightly falls down to 87.2%. Due to the presence of taxonomical information in KB, the system has more core patterns to rely on. As a result, the inferential role of paradigm extension is reduced with respect to Experiment 1 where extended paradigms play a more prominent role in supporting possible inferences.

6 Conclusions

At this stage of development, SenSOR's performance, remarkably satisfactory when only word co-occurrence patterns are used, still improves by adding one level of non-disambiguated taxonyms to words in the patterns. Looked at from this perspective, SenSOR seems to combine the advantages of two worlds: on the one hand, it is capable of drawing inferences from pure word co-occurrence patterns (as in strictly distributional approaches); on the other hand, it overcomes the bottleneck problem of data sparseness by exploiting taxonomical relationships to make the most of comparatively small collections of typical patterns of use (as in example-based techniques). Most notably, such a twofold strategy proves to be flexible enough to come to grips with a notoriously slippery source of information such as MRDs, which appear to have eluded so far most attempts to use their lexico-semantic information automatically (Velardi, 1991).

Trained on different kinds of data, SenSOR has the potential of dealing well with other sorts of syntactic ambiguity such as PP attachment, and with compound nouns interpretation or word sense disambiguation (Montemagni et. al. 96).

Several strategies can be pursued to refine its inferential power:

1. the immediate taxonyms of elements in the patterns can in their turn be associated with their own taxonyms, and so on and so forth up to the top of the taxonomy (if necessary); this would result in more powerful generalisations, since cores based on more general semantic information get extracted; clearly, this strategy requires disambiguated taxonyms;
2. the best-analogue(s) function can be made sensitive to the semantic granularity of both core patterns and extended paradigms, so that specific semantic evidence is always given priority over more general evidence (which is eventually used only for lack of the latter); this idea meets the need for keeping the inferential routine of the SenSOR as weak and conservative as possible;
3. VSO patterns can be supplied during training with frequency information of their occurrence in large textual corpora; frequency can then be used as a preferential cue for SenSOR to settle on the most plausible SOA.

(*) All ideas of this paper are the outcome of a joint effort. For the specific concerns of the Italian Academy only, Montemagni is responsible for sections 2, 3, 4.1.1, 4.1.3, Federici for sections 4.1.2, 5, 6 and Pirrelli for sections 1, 4, 4.2.

References

Collins Giunti Marzocco, 1985. *English-Italian Italian-English Dictionary*. Collins Giunti Marzocco, London Firenze.

Federici Stefano, 1991. SECS-NLC: A Self-Expandable Connectionist for Natural Language Comprehension, in *Proceedings of AAAI Spring Symposium*.

Federici Stefano, Pirrelli Vito, 1994. Linguistic Analogy as a Computable Process, in *Proceedings of NeMLaP*, Manchester, UK, pp. 8-14.

Federici Stefano, Pirrelli Vito, Yvon François, 1996. A Dynamic Approach to Paradigm-Driven Analogy, in Stefan Wermter, Ellen Riloff, Gabriele Scheler, (eds.), *Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing*, Springer Verlag, Berlin, pp. 385-396.

Garzanti, 1984. *Il Nuovo Dizionario Italiano Garzanti*. Garzanti, Milano.

Harris Zellig S., 1968. *Mathematical Structures of Language*. Wiley, New York.

Hindle Donald, 1990. Noun classification from predicate-argument structures, in *Proceedings of the 28th ACL*, Pittsburgh, PA, pp. 268-275.

Hindle Donald, Rooth Mats, 1993. Structural Ambiguity and Lexical Relations, in *Computational Linguistics*, vol. 19, n.1, March 1993. pp. 103-120.

Montemagni Simonetta, 1995. *Subject and Object in Italian Sentence Processing*. PhD thesis, University of Manchester Institute of Science and Technology (UMIST), UK.

Montemagni Simonetta, Federici Stefano, Pirrelli Vito, 1994. Who's what? Generalization by Analogy and the Problem of Subject/Object Identification in Italian, in *Proceedings of NeMLaP*, Manchester, UK, pp. 15-22.

Montemagni Simonetta, Federici Stefano, Pirrelli Vito, 1996. Example-based word sense disambiguation: a paradigm-driven approach, in *Proceedings of EURALEX-96*, Goteborg, Sweden, August 1996.

Nirenburg Sergei, Domashnev Constantine, Grannes Dean J., 1993. Two Approaches to Matching in Example-Based Machine Translation, in *Proceedings of TMI-93*, pp. 47-57.

Pirrelli Vito, 1993. *Morphology, Analogy and Machine Translation*. PhD thesis, Salford University, UK.

Pirrelli Vito, Federici Stefano, 1992. An analogical way to language modelling: MORPHEUS, in *Acta Linguistica Hungarica*, n. 41, 1992-93, pp. 235-263.

Tsujii Jun-ichi, Ananiadou Sofia, Arad Iris, Sekine Satoshi, 1992. Linguistic Knowledge Acquisition from Corpora, in Proceedings of International Workshop on Fundamental Research for the Future Generation of Natural Language Processing, 30-31 July 1992, Manchester, UK, pp. 61-81.

Uramoto Naohiko, 1994. A Best-Match Algorithm for Broad-Coverage Example-Based Disambiguation, in *Proceedings of COLING-94*, pp. 717-721.

Velardi Paola, 1991. Acquiring a Semantic Lexicon for Natural Language Processing, in Uri Zernik, (ed.), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey, pp. 341-367.