

Semantic tagging for resolution of indirect anaphora

R. Vieira¹, E. Bick², J. Coelho¹, V. Muller¹, S. Collovini¹, J. Souza¹, L. Rino³

UNISINOS¹, University of Denmark², UFSCAR³

renatav@unisininos.br, eckhard.bick@mail.dk, lucia@dc.ufscar.br

Abstract

This paper presents an evaluation of indirect anaphor resolution which considers as lexical resource the semantic tagging provided by the PALAVRAS parser. We describe the semantic tagging process and a corpus experiment.

1 Introduction

Bridging anaphora represents a special part of the general problem of anaphor resolution. As a special case of anaphora, it has been studied and discussed by different authors and for various languages. There are many problems in developing such studies. First, bridging is not a regular class, it seldom contains cases of associative and indirect anaphora (defined in the sequence); lexical resources such as Wordnet are not available for every language, and even when available such resources have proven to be insufficient for the problem. In fact, different sources of lexical knowledge have been evaluated for anaphora resolution (Poesio et al., 2002; Markert and Nissim, 2005; Bunescu, 2003). At last, corpus studies of bridging anaphora usually report results on a reduced number of examples, because this kind of data is scarce. Usually bridging anaphora considers two types: **Associative anaphors** are NPs that have an antecedent that is necessary to their interpretation (the relation between the anaphor and its antecedent is different from identity); and **Indirect anaphor** are those that have an identity relation with their antecedents but the anaphor and its antecedent have different head-nouns. In both associative and indirect anaphora, the semantic relation holding between the anaphor and its antecedent play an essential role for res-

olution. However, here we present an evaluation of the semantic tagging provided by the Portuguese parser PALAVRAS (Bick, 2000) (<http://visl.sdu.dk/visl/pt/parsing/automatic>) as a lexical resource for indirect anaphora resolution. We focus on indirect anaphors for two reasons, they are greater in number and they present better agreement features concerning human annotation.

2 Semantic Annotation with Prototype Tags

As a Constraint Grammar system, PALAVRAS encodes all annotational information as word based tags. A distinction is made between morphological, syntactic, valency and semantic tags, and for a given rule module (or level of analysis), one tag type will be regarded as primary (= flagged for disambiguation), while tags from lower levels provide unambiguous context, and tags from higher levels ambiguous lexical potentialities. Thus, semantic tags are regarded as secondary help tags at the syntactic level, but will have undergone some disambiguation at the anaphora resolution level. The semantic noun classes were conceived as *distinctors* rather than semantic definitions, the goal being on the one hand to capture semantically motivated regularities and relations in syntax, on the other hand to allow to distinguish between different senses, or to chose different translation equivalents in MT applications. A limited set of *semantic prototype* classes was deemed ideal for both purposes, since it allows at the same time similarity-based lumping of words (useful in structural analysis, IR, anaphora resolution) and context based polysemy resolution for an individual word (useful in MT, lexicography, alignment). Though we define *class hypernyms* as prototypes in the Roschian sense (Rosch, 1978)

as an (idealized) best instance of a given class of entities, we avoided low level prototypes, using <Azo> for four-legged land-animals rather than <dog> and <cat> for dog and cat races etc.). Where possible, systematic sub-classes were established. Semiotic artifacts <sem>, for instance are sub-divided into “readables” <sem-r> (book-prototype: *book, paper, magazine*), “watchables” <sem-w> (*film, show, spectacle*), “listenables” etc. The final category inventory, though developed independently, resembles the ontology used in the multilingual European SIMPLE project (<http://www.ub.es/~gilcub/SIMPLE/simple.html>). For the sake of rule based inheritance reasoning, semantic prototype classes were bundled using a matrix of 16 *atomic semantic features*. Thus, the atomic feature +MOVE is shared by the different human and animal prototypes as well as the vehicle prototype, but the vehicle prototype lacks the +ANIM feature, and only the bundle on human prototypes (<Hprof>, <Hfam>, <Hideo>,...) shares the +HUM feature (human professional, human family, human follower of a theory/belief/conviction/ideology). In the parser, a rule selecting the +MOVE feature (e.g. for subjects of movement verbs) will help discard competing senses from lemmas with the above prototypes, since they will all inherit choices based on the shared atomic feature. Furthermore, atomic features can themselves be subjected to inheritance rules, e.g. +HUM → +ANIM → +CONCRETE, or +MOVE → +MOVABLE. In Table 1, which contains examples of polysemic institution nouns, positive features are marked with capital letters, negative features with small letters¹. The words in the Table 1 are ambiguous with regard to the feature H, and since it is only the <inst> prototype that contributes the +HUM feature potential, it can be singled out by a rule selecting 'H' or by discarding 'h'. The parser's about 140 prototypes have been manually implemented for a lexicon of about 35.000 nouns. In addition, the ±HUM category was also introduced as a selection restriction for 2.000 verb senses (subject restriction) and 1.300 adjective senses (head restriction).

While the semantic annotation of common nouns is carried out by disambiguating a given lemma's lexicon-listed prototype potential, this strategy is not sufficient for proper nouns, due

¹furn=furniture, con=container, inst=institution

Ee = entities (±CONCRETE)					
Jj = ±MOVABLE					
Hh = ±HUMAN ENTITY					
Mm = ±MAS					
Ll = ±LOCATION					
polysemy spectrum					
Ee	j	Hh	m	Ll	<i>faculdade</i>
E		H		L	<inst> univ. faculty
e		h		l	<f-c> property
Ee	j	Hh	m	Ll	<i>fundo</i>
e		h		L	<Labs> bottom
E		H		L	<inst> foundation
e		h		l	<ac> <smP> funds
Ee	j	Hh	Mm	Ll	<i>indústria</i>
E		H	m	L	<inst> industry
e		h	M	l	<am> diligence
E	Jj	Hh	m	L	<i>rede</i>
	J	h			<con> net
	j	H			<inst> <+n> network
	J	h			<furn> hammock

Table 1: Feature bundles in prototype based polysemy

to the productive nature of this word class. In two recent NER projects, the parser was augmented with a pattern recognition module and a rule-based module for identifying and classifying names. In the first project (Bick, 2003), 6 main classes with about 15 subclasses were used in a lexeme-based approach, while the second adopted the 41 largely functional categories of Linguateca's joint HAREM evaluation in 2005 (<http://www.linguateca.com>). A lexicon-registered name like *Berlin* would have a stable tag (<civ> = civitas) in the first version, while it would be tagged as either <hum>, <top> or <org> in the second, dependent on context. At the time of writing, we have not yet tagged our anaphora corpus with name type tags, and it is unclear which approach, lexematic or functional, will work best for the resolution of indirect and associative anaphora.

3 Indirect Anaphora Resolution

Our work was based on a corpus formed by 31 newspaper articles, from Folha de São Paulo, written in Brazilian Portuguese. The corpus was automatically parsed using the parser PALAVRAS, and manually annotated for anaphoricity using the MMAX tool (<http://mmax.eml-research.de/>). Four subjects annotated the corpus. All annotators agreed on the antecedent in 73% of the cases, in other 22% of the cases there was agreement between three annotators and in 5% of the cases only two annotators agreed. There were 133 cases of

definite *Indirect anaphors* (NPs starting with definite articles) from the total of 1454 definite descriptions (near to 10%) and 2267 NPs.

The parser gives to each noun of the text (or to most of them) a semantic tag. For instance, the noun *japonês* [*japanese*] has the following semantic tags *ling* and *Hnat*, representing the features: human nationality and language respectively.

```
<word id="word_28">
  <n can="japonês" gender="M" number="S">
    <secondary_n tag="Hnat"/>
    <secondary_n tag="ling"/>
  </n>
</word>
```

The approach consists in finding relationships with previous nouns through the semantic tags. The chosen antecedent will be the nearest expression with the largest number of equal semantic tags. For instance, in the example below, the anaphor is resolved by applying this resolution principle, to *japonês - a língua*.

O Eurocenter oferece *cursos de japonês* em Kanazawa. Após um mês, o aluno falará modestamente *a língua*.

The Eurocenter offers *Japanese courses* in Kanazawa. After one month, a student can modestly speak *the language*.

As both expressions (japanese and language) hold the semantic tag “ling” the anaphor is resolved. For the experiments, we considered as correct the cases where the antecedent found automatically was the same as in the manual annotation (same), and also the cases in which the antecedent of the manual annotation was found further up in the chain identified automatically (in-chain). We also counted those cases in which the antecedent of the manual annotation was among the group of candidates sharing the same tags (in-candidates), but was not the chosen one (the chosen being the nearest with greater number of equal tags).

Indirect anaphora		
Results	#	% of Total
Same	25	19%
In-chain	15	11%
Total Correct	40	30%
In-candidates	9	7%
Unsolved	40	30%
Error	44	33%
Total	133	100%

Table 2: Indirect anaphor resolution

Table 2 shows the results of the indirect anaphor resolution. In 19% of the cases, the system found

the same antecedent as marked in the manual annotation. Considering the chain identified by the system the correct cases go up to 30%. The great number of unsolved cases were related to the fact that proper names were not tagged. Considering mainly the tagged nouns (about 93 cases), the correct cases amount to 43%). This gives us an idea of the quality of the tags for the task. We further tested if increasing the weight of more specific features in opposition to the more general ones would help in the antecedent decision process. A semantic tag that is more specific receives a higher weight. The semantic tag set has three levels, level 1, which is more general receives weight 1, level 2 receives 5, and level 3 receives 10. See the example below.

```
<A>      1      Animal, umbrella tag
<AA>     5      Group of animals
<Adom>  10     Domestic animal
```

In this experiment the chosen candidate is the nearest one whose sum of equal tag values has higher weight. Table 3 shows just a small improvement in the correct cases. If we do not consider unsolved cases, mostly related to proper names, indirect anaphors were correctly identified in 46% of the cases (43/96).

Indirect anaphora		
Results	#	% of Total
Same	24	18%
In-chain	19	14%
Total Correct	43	32%
In-candidates	6	5%
Unsolved	40	30%
Error	44	33%
Total	133	100%

Table 3: Indirect anaphor - weighting schema

Since there is no semantic tagging for proper names as yet, the relationship between pairs such as *São Carlos - a cidade* [*São Carlos - the city*] could not be found. Regarding wrong antecedents, we have seen that some semantic relationships are weaker, having no semantic tags in common, for instance: *a proposta - o aumento* [*the proposal - the rise*]. In some cases the antecedent is not a previous noun phrase but a whole sentence, paragraph or disjoint parts of the text. As we consider only relations holding between noun phrases, these cases could not be resolved. Finally, there are cases of plain heuristic failure. For instance, establishing a relationship between *os professores*

[*the teachers*], with the semantic tags *H* and *Hprof*, and *os politicos* [*the politicians*], with the semantic tags *H* and *Hprof*, when the correct antecedent was *os docentes* [*the docents*], with the semantic tags *HH* (group of humans) and *Hprof*.

4 Final Remarks

Previous work on nominal anaphor resolution has used lexical knowledge in different ways. (Poesio et al., 1997) presented results concerning the resolution of bridging definitions, using the WordNet (Fellbaum, 1998), where bridging DDs enclose our *Indirect* and *Associative anaphora*. Poesio et al. reported 35% of recall for synonymy, 56% for hypernymy and 38% for meronymy. (Schulte im Walde, 1997) evaluated the bridging cases presented in (Poesio et al., 1997), on the basis of lexical acquisition from the British National Corpus. She reported a recall of 33% for synonymy, 15% for hypernymy and 18% for meronymy. (Poesio et al., 2002) considering syntactic patterns for lexical knowledge acquisition, obtained better results for resolving meronymy (66% of recall). (Gasperin and Vieira, 2004) tested the use of word similarity lists on resolving indirect anaphora, reporting 33% of recall. (Markert and Nissim, 2005) presented two ways (WordNet and Web) of obtaining lexical knowledge for antecedent selection in coreferent DDs (*Direct* and *Indirect anaphora*). Markert and Nissim achieved 71% of recall using Web-based method and 65% of recall using WordNet-based method. We can say that our results are very satisfactory, considering the related work. Note that usually evaluation of bridging anaphora is made on the basis of a limited number of cases, because the data is sparse. Our study was based on 133 examples, which is not much but surpasses some of the previous related work. Mainly, our results indicate that the semantic tagging provided by the parser is a good resource for dealing with the problem, if compared to other lexical resources such as WordNet and acquired similarity lists. We believe that the results will improve significantly once semantic tags for proper names are provided by the parser. This evaluation is planned as future work.

Acknowledgments

This work was partially funded by CNPq.

References

- Eckhard Bick. 2000. *The Parsing System PALAVRAS: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Ph.D. thesis, Arhus University, Arhus.
- Eckhard Bick. 2003. Multi-level ner for portuguese in a cg framework. In Nuno J. et al. Mamede, editor, *Computational Processing of the Portuguese Language (Proceedings of the 6th International Workshop, PROPOR 2003)*, number 2721 in Lecture Notes in Computer Science, pages 118–125, Faro, Portugal. Springer.
- Razvan Bunescu. 2003. Associative anaphora resolution: A web-based approach. In *Proceedings of the Workshop on The Computational Treatment of Anaphora - EACL 2003*, Budapest.
- Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.
- Caroline Gasperin and Renata Vieira. 2004. Using word similarity lists for resolving indirect anaphora. In *Proceedings of ACL Workshop on Reference Resolution and its Applications*, pages 40–46, Barcelona.
- Katja Markert and Malvina Nissim. 2005. Comparing knowledge sources for nominal anaphora resolution. *Computational Linguistics*, 31(3):367–401.
- Massimo Poesio, Renata Vieira, and Simone Teufel. 1997. Resolving bridging descriptions in unrestricted texts. In *Proceedings of the Workshop on Operational Factors In Practical, Robust, Anaphora Resolution for Unrestricted Texts*, pages 1–6, Madrid.
- Masimo Poesio, Ishikawa Tomonori, Sabine Shulte im Walde, and Renata Vieira. 2002. Acquiring lexical knowledge for anaphora resolution. In *Proceedings of 3rd Language resources and evaluation conference LREC 2002*, Las Palmas.
- Eleanor Rosch. 1978. Principles of categorization. In E. Rosch and B. Lloyd, editors, *Cognition and Categorization*, pages 27–48. Hillsdale, New Jersey: Lawrence Erlbaum Associate.
- Sabine Schulte im Walde. 1997. *Resolving Bridging Descriptions in High-Dimensional Space Resolving Bridging Descriptions in High-Dimensional Space*. Ph.D. thesis, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, and Center for Cognitive Science, University of Edinburgh, Edinburgh.