Resolving Bridging References in Unrestricted Text

Massimo Poesio, Renata Vieira and **Simone Teufel**
Centre for Cognitive Science - University of Edinburgh
2, Buccleuch Place EH8 9LW Edinburgh UK
{poesio,renata,simone}@cogsci.ed.ac.uk

Abstract

Our goal is to develop a system capable of treating the largest possible subset of definite descriptions in unrestricted written texts. A previous prototype resolved anaphoric uses of definite descriptions and identified some types of first-mention uses, achieving a recall of 56%. In this paper we present the latest version of our system, which handles some types of bridging references, uses WordNet as a source of lexical knowledge, and achieves a recall of 65%.

1 Previous Work

We are in the process of developing a system for interpreting definite descriptions (DDs) in written text without restrictions of domain. The implementation work has been supported by an analysis of definite description use in corpora of written language (Poesio and Vieira, 1997). In one of our experiments, we asked 2 subjects to classify the uses of definite descriptions in a corpus of English texts[1] using a taxonomy derived from the proposals of (Clark, 1977; Hawkins, 1978; Prince, 1981; Fraurud, 1990; Prince, 1992). In the taxonomy used in that study, we defined bridging references as those uses of definite descriptions based on previous discourse which require some reasoning in the identification of their textual antecedent (rather than just matching identical nouns). These definite descriptions may be co-referential with an entity already introduced in the discourse, but be characterized by a different head noun (as in *a car ... the vehicle*); or may be simply semantically related to it (in the sense that *the door* is related to *house*). Of the 1040 DDs in that corpus, 204 (20%) were identified as bridging descriptions, 312 (30%) as anaphoric (DDs and antecedents which co-refer and have the same head noun), and 492 (47%) as larger situation/unfamiliar (Hawkins, 1978) (Prince's discourse new (Prince, 1992)); the remaining definite descriptions were classified as idiomatic or doubtful cases.

[1] A set of randomly selected parsed articles from the Wall Street Journal contained in the ACL/DCI CD-ROM.

These results led us to concentrate initially on resolving same-head anaphoric DDs and on recognising larger situation/unfamiliar uses. Our analysis of the corpus suggested that many of the latter could be recognised using syntactic heuristics: e.g., on the basis of the presence of restrictive pre- and post-modification, of the presence of special predicates (such as the superlatives *first, best*), or because the DD occurred in a copula or appositive construction.

A first prototype with these capabilities (Vieira and Poesio, 1997) achieved an overall recall of 56% and precision of 84% when tested on our corpus. Of all anaphoric DDs 72% were resolved, and 74% of all larger situation and unfamiliar uses were identified. The definite descriptions not handled by this first prototype were typically larger situation uses based on common knowledge (such as *the government*) and bridging descriptions. In this paper we present our subsequent work devoted to handling some of these remaining cases.

2 Bridging Descriptions

Linguistic and computational theories of bridging descriptions identify two main subtasks involved in their resolution: first, finding the element in the text to which the bridging description is related (ANCHOR) and second, finding the relation (LINK) holding between the bridging description and its anchor (Clark, 1977; Sidner, 1979; Heim, 1982; Carter, 1987; Fraurud, 1990; Strand, 1997).

A speaker is licensed to use a bridging DD when he/she can assume that the common-sense knowledge required to identify the relation is shared by the listener (Hawkins, 1978; Clark and Marshall, 1981; Prince, 1981). This reliance on commonsense knowledge means that, in general, a system could only resolve bridging references when supplied with an adequate knowledge base; for this reason, the typical way of implementing a system for resolving bridging references has been to restrict the domain and feeding the system with hand-tailored world knowledge. (This approach is discussed in detail in (Carter, 1987)). In order to get a system capable of performing on unrestricted text, we decided to use WordNet (WN) (Miller, 1993) as an approximation of a knowledge base containing generic information, and to supplement it with heuristics to handle those cases which WN couldn't handle.

Vieira and Teufel (1997) analyzed the corpus to identify the cases of bridging descriptions that could be resolved using WordNet, those for which we could use heuristics, and those that couldn't be interpreted at the moment. Six classes of bridging descriptions were identified.[2]

Synonymy/Hyponymy/Meronymy This class (henceforth, Syn/Hyp/Mer) includes those DDs which are in a synonymy/hyponymy/meronymy relation with their anchors, i.e., the kind of semantic relation that is currently encoded in WordNet. Examples[3] are:

(1) a. Synonymy: *new album — the record*; *three bills — the legislation.*
  b. Hypernymy-Hyponymy: *rice — the plant*; *the daily television show — the program.*
  c. Meronymy (part-of relation): *plants — the pollen*; *house — the chimney.*

Names This class includes definite descriptions that refer back to proper names such as people's and company names, as in:

(2) *Bach — the composer*; *Pinkerton's Inc — the company.*

Compound Nouns This class includes bridging descriptions whose LINGUISTIC ANCHOR (i.e., the element in the text to which they are related) is a noun occurring as part of a compound noun other than the head. Examples include:

(3) *stock market crash — the markets*; *discount packages — the discounts.*

Events These are cases where the linguistic anchors of DDs are not NPs but VPs or sentences. Examples are:

(4) Individual investors and professional money managers *contend.* — They make *the argument* ...; Kadane Oil Co. *is currently drilling two wells and putting money into three others.* — *The activity* ...

Class	Total	%
Syn/Hyp/Mer	12/14/12	19%
Names	49	24%
Compound Nouns	25	12%
Events	40	20%
Discourse Topic	15	7%
Inference	37	18%
Total	204	100%

Table 1: Distribution of types of bridging DDs

Discourse Topic There are some cases of DDs which are related to the (often implicit) discourse topic (in the sense of (Reinhart, 1981)) of a text, rather than to some specific NP or VP. For instance,

(5) *the industry* (in a text whose discourse topic is oil companies); *the first half* (in a text whose discourse topic is a concert).

Inference We collect in this class all the cases of bridging descriptions whose relation with their NP anchor was based on more complex inferential relations: for example, cases in which the relation between the anchor and the DD was of reason, cause, consequence, or set-membership:

(6) *last week's earthquake — the suffering people are going through*; *Democratics/Republicans — the two sides.*

The relative importance of these classes in our corpus is shown in Table 1. This classification is based on what we took to be the *main* linking relation for each of the 204 bridging DDs in the corpus[4].

3 Resolution of Bridging Descriptions

We used Vieira and Teufel's analysis as the basis for the implementation of a second prototype. In this section we discuss how this prototype handles the different types of bridging descriptions.

3.1 Bridging DDs and WordNet

We implemented a WordNet interface that reports a possible semantic link between two nouns when one

[2] We should stress that this classification is primarily motivated by processing considerations.

[3] Note that the examples in (1a) and (1b) are classified as bridging even though the relation is of co-reference.

[4] One problem with bridging references is that they are often related to more than one antecedent in the discourse (Poesio and Vieira, 1997).

of the following is true:

- The nouns are in the same synset (= synonyms of each other), as in *suit* — *lawsuit*.
- The nouns are in direct hyponymy relation with each other, for instance, *dollar* — *currency*.
- There is a direct or indirect meronymy between them. Indirect meronymy holds when a concept inherits parts from its hypernyms, like *car* inherits the part *wheel* from its hypernym *wheeled_vehicle*.
- Due to WordNet's idiosyncratic encoding, it is often necessary to look for a semantic relation between *coordinate sisters*, i.e. hyponyms of the same hypernym, such as *home* — *house* which are hyponyms of *housing*, *lodging*.

Sometimes, if a relation between two head nouns is not encoded in WN directly, the semantic closeness might be found through the compound nouns made up of them. Thus, for a pair such as *record,album* we find synonymy between *record_album* and *album*.

We ran a test in which WordNet was used to identify the DD's anchors. For each of the 204 bridging DDs in our corpus, we considered the NPs in the previous five sentences as a potential anchor[5], and queried WN with the DD and potential anchor pair.

WordNet reported a possible relation for 107 of the 204 DDs. Often, more than one link was found between a DD and a potential anchor,[6] and sometimes links could be found with more than one potential anchor in the previous five sentences. Only in 34 of these 107 cases we found at least one appropriate anchor linking relation, and only in 21 cases we found only appropriate anchors (for 13 there was a mixture of suitable and unsuitable anchors).

Of these 34 DDs for which a correct anchor was found, only 18 were among those we had classified as Syn/Hyp/Mer. In 8 of these 18 cases, WordNet found a link with an anchor that was not the one we had identified manually, but which was still valid; for instance, we identified the link *pound* — *the currency* which was not found in WordNet, whereas our automatic search found *sterling* — *the currency*).

The 16 remaining relations were found for DDs that we had not classified as Syn/Hyp/Mer: for instance, whereas we had classified the DD *the population* as belonging to the class of DDs based on names having as anchor *Koreans*, our implementation found a WordNet relation for the pair *nation* — *the population*; the system also found a few relations with proper names, such as *Bach* — *the composer*.

In the following tests we have considered only direct meronymy, as indirect meronymy presented extreme low recall and precision at a very expensive computational cost.

In order to reduce the number of false positives (86 out of 107)[7], we tried using a stack-based approach towards finding potential anchors in the previous sentences, as suggested in (Sidner, 1979); i.e., the system would go back one sentence at a time, and stop as soon as a relation with a potential anchor was found, rather than trying to find all possible links. As a result, we found exactly one correct anchor for 30 DDs, slightly improving our results (76 false positives).

Class	Total	Right	Wrong	P
Syn	11	4	7	36%
Hyp	59	18	42	30%
Mer	6	2	4	33%
Sister	30	6	24	20%
Total	106	30	76	28%

Table 2: Analysis of the anchors found in WN

Table 2 shows the distribution of the different semantic relations between DDs and the anchors found by our stack-based search. It presents precision figures (P) related to each type of relation[8]. Sister relations are the least satisfactory ones.

We tested, in particular, whether WordNet encoded a semantic link between the 38 syn/hyp/mer relations in our corpus (just described) plus other 32 relations extracted from a second corpus study (Poesio and Vieira, 1997). The results for the 70 DDs are summarised in Table 3. Overall recall (R) was 46%. We could have expected 100% precision, since we had manually identified the anchors, but the precision figures (P) report an error when a sister relation is found instead of the expected (syn/hyp/mer) re-

[5]Our system does not currently include a proper segmentation algorithm. Instead, we use a simple recency heuristic—we only consider the antecedents in the n previous sentences, where n is a constant determined empirically. In our previous work (Vieira and Poesio, 1997), we observed that 5 was the value of n which gave the best tradeoff between precision and recall.

[6]For instance, *transaction* and *trade* were reported to be in a hypernym relation and were also reported as coordinate sisters having as common hypernymy *commerce*.

[7]The anchors found for 73 of the DDs were incorrect, for the remaining 13 DDs the resulting anchors were mixed (some right/some wrong).

[8]We cannot estimate recall since we do not have a precise number of syn/hyp/mer anchors that should be found (as different types of anchors may allow resolution of bridging descriptions).

lation. The poorest recall was surprisingly obtained for synonymy relations, followed by meronymy relations, as expected, since these are only partially implemented in WordNet.

Class	Total	In WN	Out	R	P
Syn	20	7	13	35%	71%
Hyp	32	18	14	56%	94%
Mer	18	7	11	38%	71%
Total	70	32	38	46%	84%

Table 3: Search for semantic relations in WN

The low recall for synonymy relations may be due to the context dependent, specialized senses of sublanguage terminology (for instance, *crash, bust* and *slump* in Economics terminology). Some nouns were not even encoded in WN (such as *newsweekly, spinoff*). Other relations were missed due to the unexpected way in which knowledge is organised in WordNet. For example, no association was found between *house* and *walls*, because *house* is not encoded in WordNet as a hyponym of *building* but of *housing*, and *housing* does not have a meronymy link to *wall* whereas *building* does. Another example of counter-intuitive position in the hierarchy is that of *catastrophe*, not listed as a hypernym of *earthquake*, but as its coordinate sister. This example demonstrates the problems that WordNet lexicographers faced when they had to coerce real-world concepts into a hierarchy with single, no-default inheritance: catastrophes clearly subclassify into earthquakes and floods, but minor earthquakes don't have to be catastrophic.

3.2 Proper names

Definite descriptions which refer back to proper names are very common in Wall Street Journal articles. Processing such DDs involves, first, determining an entity type for each name in the text, then searching for semantic relations. If we get the entity type *person* for the a name such as *Mrs. Y.J. Park* we could, ideally, resolve the subsequent DD *the housewife* using WordNet. A few names are available in WordNet: typically, famous people, countries, states, cities and languages. Other entity types can be identified using appositive constructions and abbreviations like *Mr., Co., Inc.* etc. as cues.

The algorithm we developed, based on a mixture of access to WordNet and heuristics such as those we described, found the correct type for 66% of the names in our corpus (535/814). Including a back-tracking mechanism which re-processes a text looking for missing name types (with this mechanism we identify the type for the name *Morishita* in a textual sequence like *Morishita — Mr. Morishita*) increases our recall to 69% (562/814). We then used WordNet to match the types found with previous references in the text. This resulted in the resolution of 53% (26/49) of the cases based on names. We missed relations which are not found in WordNet (for instance, *Mr. Morishita — the 57 year-old*). But again we also found a large number of false positives.

3.3 Compound Nouns

Sometimes, a bridging description is linked to a non-head noun in a compound noun:

(7) *stock market crash — the markets*; *rule changes — the rules*; *discount packages — the discounts.*

One way of processing these definite descriptions would be to update the discourse model with discourse referents not only for the NP as a whole, but also for the embedded nouns: for example, after processing *stock market crash*, we could introduce a discourse referent for *stock market*, and another discourse referent for *stock market crash.* The DD *the markets* would be co-referring with the first of these referents (with identical head noun), and we could simply use our anaphora resolution algorithms. This solution, however, makes available discourse referents that are generally unaccessible for pronominal anaphora. We therefore followed a different route: our algorithm for identifying antecedents attempts to match not only heads with heads, but also the head of a DD with the pre-modifiers of a previous NP, the pre-modifiers of a DD with the pre-modifiers of its antecedents, and the pre-modifiers of the DD with the head of a previous NP. With this, we obtained a total of 54 resolutions of which 36 were correct, although we did not always find a main linking relation licensing the use of a DD.

Examples of correct resolutions are:

(8) a. Head of DD with pre-modifier of antecedent: *the stock market crash — the markets; rule changes — the rules*;

b. Pre-modifiers of DD with pre-modifiers of antecedent: *most oil companies — the oil fields; his art business — the art gallery*;

c. Pre-modifiers of DD with head of antecedent: *New York City — the city council district lines; a 15-acre plot and main home – the home site.*

There are also cases in which the pre-modifiers plus the head noun of a DD may indicate a bridging reference: we may find an antecedent with the same head noun for a DD but referring to a different entity, this being signalled by the pre-modification. Some examples:

(9) *the company's abrasive segment — the engineering materials segment*; *Italy's unemployment rate — the southern unemployment rate*; *Pinkerton — the new Pinkerton*; *increases of 3.9 % — the actual wage increases may have been bigger.*

Our previous heuristics for treatment of pre-modifiers in anaphoric resolution handled the first two examples correctly (Vieira and Poesio, 1997): as they present different pre-modifiers we did not treat them as anaphoric in the first version of our system. Such cases, as well as DDs modified by *new* and *actual* (last two examples), may now be treated as bridging references[9].

3.4 Events

To process DDs based on events (situations or propositions), we are trying, as a first approach, to transform verbs into their nominalizations, and then looking for a relation in WordNet. Some nominalizations can be generated by general procedures or learned by means of a stochastic method: e.g., we could use WordNet's morphology component as a stemmer, and augment the verbal stems with the most common suffixes for nominalizations which could be kept in a list, like *-ment, -ion*. In our corpus, 17% (7/40) of the bridging references based on events are direct nominalizations of this type (for instance, *changes were proposed - the proposals*).

Another 17% are cases in which knowledge of the semantics of the verb is necessary (as in *borrowed - the loan*). The remaining 66% (26 cases) of bridging DDs based on events require inference reasoning based on the compositional meaning of the phrases (as in *It went looking for a partner - pitching the prospect*); these cases are out of reach just now, as well as the cases listed under "discourse topic" and "inference".

4 Restrictive Post-modification as Anchors

Whereas the problem of finding the appropriate textual anchor for bridging descriptions requires knowledge inference and reasoning, DDs with restrictive post-modification give the reader both anchor and description in the same expression. As Hawkins points out (Hawkins, 1978):

> ... first mention uses of *the* with both referent establishing relative clauses and associative clauses are not essentially different from the other uses mentioned in the last section (Hawkin's associative uses/our bridging uses). The only difference is that in the latter uses set identification and the locatability of the referent were possible on account of previous triggers, whereas it is now function of the modifier itself to provide the information which makes set identification and location possible— (parentheses ours).

This could be seen as an advantage, as we directly find the anchors of these DDs. They are quite common uses of DDs, but not much attention has been devoted to them as a special case of anchor linking relations. The main problem for these cases is to find out their links, which is also a remaining problem for our proposals of anchor identification throughout the paper.

The head noun of a DD and its modifier may be related in different ways, as shown by the examples:

(10) *the number of job seekers*; *the anthers of the plant*; *the ideal of a level playing field*; *the flip side of the Stoltzman personality.*

In some cases, the relation between noun and complement seem to be looser than the relations for bridging descriptions. Sequences such as *the laws of heredity; the cost of the plan* are acceptable, whereas *heredity — the laws; the plan — the cost* are unlikely to occur. On the other hand, bridging such as *the house — the kitchen; the firm — the owners* are as acceptable as *the kitchen of the house; the owners of the firm.* Some proposals of a systematic treatment for the identification of anchor linking relations for bridging DDs are (Heim, 1982; Barker, 1991; Poesio, 1994; Strand, 1997).

5 Discussion

We presented our most recent results concerning the resolution of bridging descriptions. We identified different types of bridging descriptions, and we proposed a treatment for each of them separately.

We presented some preliminary experiments with WordNet. Our experience was mixed. WordNet was useful in determining the type of entity for some of the proper names in our corpus, typically cities, states and countries. On the other hand, WordNet proved to be unreliable for the automatic identification of possible link relations, as too many false positives were found. A 'blind' WordNet search for semantic relations is also very expensive computationally. A mechanism for focus tracking (Grosz and

[9]This idea is not implemented yet.

Sidner, 1986) or a clustering algorithm should be applied first in order to minimise the costs.

In order to have proper names available for resolution of future references, it is useful to create discourse referents for them which contain their entity types. Up to now we have identified an entity type for 69% of the names in our corpus, and we resolved 53% of the DDs referring back to proper names with the help of WordNet.

Pre-modifiers in compound nouns may license the use of definite descriptions. We have presented some preliminary tests that should be further developed.

Cases of bridging references based on events or propositions usually involve common-sense reasoning; some of them (in our corpus, 34% of all cases based on events) can however be solved by transforming verbs into their nominalizations and then searching for a semantic relation.

We also claimed that the same problem of determining a linking relation for bridging descriptions holds for first mention uses of DD based on restrictive post-modification.

As an estimate, we could say that about 60% of the cases in the bridging class could be treated by developing the ideas proposed here. We combined the proposed heuristics with the first version of our system— we applied the heuristics for proper names, compound nouns and WordNet consult (in this order) to the cases which the system did not handle before. The impact on the overall system performance was an increase in recall from 56% to 65% (note that the bridging class is a small class compared to the others) but precision decreased from 84% to 82%. The heuristics should be further developed, and their integration into the system should be worked out in more detail—the heuristics could be implemented in parallel or through a decision tree.

Acknowledgements The authors would like to thank Kjetil Strand and the anonymous referees for their comments on earlier drafts of the paper. The authors are supported by an EPSRC Advanced Fellowship, a CNPq studentship and an EPSRC studentship, respectively.

References

Barker, C. 1991. *Possessive Descriptions.* Ph.D. thesis, University of California at Santa Cruz, Santa Cruz, CA.

Carter, D. M. 1987. *Interpreting Anaphors in Natural Language Texts.* Ellis Horwood.

Clark, H. H. 1977. Bridging. In P. N. Johnson-Laird and P.C. Wason, eds., *Thinking: Readings in Cognitive Science.* Cambridge University Press.

Clark, H. H. and C. R. Marshall. 1981. Definite reference and mutual knowledge. In *Elements of Discourse Understanding.* Cambridge University Press.

Fraurud, K. 1990. Definiteness and the processing of NPs in natural discourse. *Journal of Semantics*, 7:395–433.

Grosz, B. J. and C. L. Sidner. 1986. Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Hawkins, J. A. 1978. *Definiteness and Indefiniteness.* Croom Helm.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* Ph.D. thesis, University of Massachusetts at Amherst.

Miller, G. et al. 1993. Five papers on WordNet. *Technical Report CSL Report 43, Cognitive Science Laboratory, Princeton University.*

Poesio, M. 1994. Weak definites. In *Proceedings of the Fourth Conference on Semantics and Linguistic Theory, SALT-4.* Cornell University Press.

Poesio, M. and Vieira, R. 1997. A Corpus based investigation of definite description use. Manuscript, Centre for Cognitive Science, University of Edinburgh.

Prince, E. F. 1981. Toward a taxonomy of given-new information. In P. Cole, ed., *Radical Pragmatics.* Academic Press, pages 223–256.

Prince, E. F. 1992. The ZPG letter: subjects, definiteness, and information status. In S. Thompson and W. Mann, eds., *Discourse description: diverse analyses of a fund-raising text.* John Benjamins, pages 295–325.

Reinhart, T. 1981. Pragmatics and Linguistics: An Analysis of Sentence Topics *Philosophica*, 27(1).

Sidner, C. L. 1979. *Towards a computational theory of definite anaphora comprehension in English discourse.* Ph.D. thesis, MIT.

Strand, K. 1997. A Taxonomy of Linking Relations. *Journal of Semantics*, forthcoming.

Vieira, R. and M. Poesio. 1997. Processing definite descriptions in corpora. In S. Botley and M. McEnery, eds., *Corpus-based and Computational Approaches to Discourse Anaphora.* UCL Press.

Vieira, R. and S. Teufel. 1997. Towards Resolution of Bridging Descriptions. In *Proceedings of the ACL Student Session.*