Implicit Role Reference

Joel R. Tetreault
University of Rochester
Department of Computer Science
Rochester. NY, 14627
tetreaul@cs.rochester.edu

Abstract

This paper presents a preliminary examination of a rarely studied form of reference called implicit role reference. Verbs have certain required roles, which refer to discourse entitites, that are necessary for comprehending the verb phrase. Roles that do not have explicit antecedents in the verb phrase are deemed implicit. We show that resolving implicit role references are important to natural language processing and conduct an automated corpus-based study that uses a focusing algorithm to resolve implicit roles.

1 Introduction

We claim that in addition to canonical reference types such as pronominal reference, VP ellipsis, and discourse deixis, verb phrases have certain required roles that can be viewed as anaphoric. These required roles refer to discourse entities and are necessary for the interpreter to understand the verb phrase, and thus the complete utterance. For example, in order to use the verb "take" one needs to understand that an entity is being moved, that it is being moved to one place from some other place, and that there is some entity that is responsible for moving it.

This paper describes preliminary work in the area of implicit role reference. The goal of this project is to show examples of implicit role reference and provide a basic approach to resolving this reference type that is supported with a small corpus study. In the following section we discuss in more depth implicit role reference. In section 3, we detail our annotation scheme, and in the subsequent sections, we describe a focus-based algorithm for resolving implicit roles and show how it performs on the annotated corpus.

2 Implicit Role Reference

Implicit role reference has been briefly studied as a side effect of bridging and discourse relations (Poesio, 1994) and (Asher and Lascarides, 1999) but no major theoretical work or empirical work has been done in the area.

Resolution of implicit roles occurs frequently in naturally occurring dialog. Consider the following, modified from (Asher and Lascarides, 1999) (p. 90):

> (1) Take engine E1 from Avon to Dansville.
> (2a) Pick up the boxcar and take it to Broxburn.
> (2b) Also take the boxcar.
> (3) Leave it there and go to Clarksville.

For the sake of simplicity, assume that the verb "take" has these roles: "Theme": the entity being moved; "To-Loc": the location we are taking the "theme"; and "From-Loc" the location we are leaving.

In utterance (2a) one needs to know the At-Loc of the boxcar, in order to send it to Broxburn. This role is implicit and is resolved to Dansville. In order to resolve "there" after utterance (2b) one must resolve the implicit "From-Loc" in "take" in the previous sentence.

In our study we wish to show that implicit roles can be resolved by using focusing mechanisms reminiscent of (Sidner, June 1979) and similar to (Strube, 1998) and (Tetreault, 2001). We believe that each role has its own unique way of being resolved and that a discourse should have a focus stack for each role to track it.

3 Annotation Scheme

We use a subset of the TRAINS-93 Corpus (Heeman and Allen, 1994) annotated with coreference information for pronouns (Byron and Allen, 1998) The corpus consists of one TRAINS-93 dialog (d92-10.3) of 86 sentences in which two human participants are given a task involving moving food and trains around a fictional world within a certain time limit. The dialogs typically consist of short sentences, usually under 10 words or less. Each sentence (which we call an utterance) is annotated with speaker and utterance number and each word of text is marked with a part of speech tag.

For our work here we extend the corpus by annotating implicit reference roles. The annotation scheme has three parts: first annotating all verbs with role information and marking the roles that are implicit and explicit; second, marking all NP's and; third, general discourse markup of identifying disfluencies, grounding utterances, and conjoined phrases. We use a sgml encoding for the annotation.

For NP's, we assign it a unique ID number as well as mark its semantic type (such as whether it is a city, train, foodstuff, etc.).

Verbs are annotated in much the same way as NP's except that the tags are different. All verbs have these tags: location (either From-Loc or To-Loc) and Theme. Other tags include agent (which usually defaults to "You" and so is rarely encoded) and instrument.

We tag a subset of the verbs in the corpus - only those that are related to the moving and loading of food and trains. The 14 verbs are listed in Figure 1with their respective roles and what class(es) each role should be resolved to.

The verbs not encoded were stative, such as "need," "have," and "is." The location tags (From-Loc, To-Loc, and At-Loc) represent where items are moved to and from. The From-Loc for most verbs is the place where the theme comes from. The To-Loc is the place where the theme is being sent to. For verbs of loading (load, put) the To-Loc represents the container where the theme is being placed.

The theme is what is acted on. If something is being loaded or moved, it's the theme. The instrument is the entity used to complete the verb and act on the theme. These are engines or boxcars. The agent is the entity that uses the instrument to act on the verb. In the dialogs considered so far the agent has always been understood to be "You."

The location, instrument, theme and agent tags are all set to unique id's of NP's already mentioned within the verb phrase or in past utterances. Implicit roles are denoted by adding an "i" to the end of the ID. A role is implicit if it is not mentioned explicitly within the verb phrase. If a role has more than one element, multiple entities are joined by an ampersand.

Past work such, as (**?**) and (Tetreault, 2001), has shown that segmenting discourse can improve performance of reference resolution algorithms and make for more efficient algorithms. So we also break apart conjoined utterances, that is, utterances that have clauses separated by a conjunction.

4 Algorithm for Resolving Implicit Roles

Figures 2-6 show the distribution of roles and their antecedents . Figures 3-6 show for a given role, at how many sentences back (depth) its antecedent is found and in what role focus list it is found in. So in Figure 3, there are 9 implicit role references that have their antecedent in an instrument role in the preceding sentence. Also, there are cases of NP's that are not placed in a role focus list because they are part of sentence fragments: sentences without a verb. To make sure they can be possible antecedents, we place these entities in an ID focus list.

The statistics presented above lay the foundation for the unique algorithm for each role. Looking at the tables, it is clear that each role favors a certain set of roles for its antecedents. For example, implicit From-Loc's have most of their antecedents split between previous From-Loc's and To-Loc's. An antecedent for an implicit Instrument role is most likely to be found in past utterance's Instrument focus lists before Theme focus lists. In the following sections we discuss the distribution for each role and develop a strategy for resolving them.

4.1 Locations: From-Loc and To-Loc

Figures 2 and 3 show that majority of the antecedents for both roles are found in past From-Loc or To-Loc focus lists and that there is roughly an even split between the two making it hard to favor one over the other. We tried two metrics for each role: parallel and narrative. The parallel method ranks the roles in a parallel order, that is, if the referent is an implicit

Verb	Roles
make	From-Loc, Theme
get, pick-up	From-Loc, Instrument, Theme
carry, transport, get-back, move, take, deliver, continue, come, go	To-Loc, From-loc, Instrument, Theme
load, put	To-Loc, Theme

Figure 1: Verbs and Roles

Role	Total	Explicit	Implicit
Instrument	25	6	19
Theme	40	30	10
From-Loc	39	13	26
To-Loc	27	20	7
Total	131	69	62

Figure 2: Distribution of Implicit and Explicit Roles

Depth	Instrument	Theme	ID
0	0	0	1
1	9	2	1
2	5	0	0
3	1	0	0
%	78.95%	10.5%	10.5%

Figure 3: Instrument

Depth	Instrument	Theme	ID
1	0	6	1
3	0	1	0
%	0%	87.5%	12.5%

Figure 4: Theme

Depth	From-Loc	To-Loc	ID
1	11	9	0
2	4	1	0
3	0	0	0
4	1	0	0
5+	0	0	0
%	61.5%	38.5%	0%

Figure 5: From-Loc

Depth	Theme	From-Loc	To-Loc
1	0	1	1
2	0	0	2
3	0	0	0
4	1	0	1
5+	0	1	0
%	14.3%	28.6%	57.1%

Figure 6: To-Loc

From-Loc, we'll first search through the previous utterance From-Loc, then To-Loc list, and then ID list. And continue this pattern through the history until a match is found. In the narrative order, we assume that a From-Loc would be resolved to the previous utterance's To-Loc since many utterances in the transcripts involve moving things from city to city to city. This linear movement implies that the place where "we" were just came to will be the place we leave from in the next utterance. For example:

> utt57 : u: so the second the third engine whatever which one was still there
> utt58 : u: will go has to go to Corning to Dansville
> utt59 : s: okay
> utt60 : u: to pick up another boxcar well another two boxcars
> utt61 : u: and then back down to Corning to load the oranges

The engine and its train move from Corning to Dansville and then to Corning again. In utt61, the From-Loc is implicit but because the rhetorical structure is a narrative, it is understood to be Dansville since that is where the engines were moved to.

The problem with the narrative approach is that it is only applicable to From-Loc's (where the order of search would be To-Loc, From-Loc and then ID). To-Loc's in a narrative will almost always refer to something new given the nature of a narrative, so that approach for To-Loc's is not intuitively the best, unless the verb is a "return" type verb such as "go back". If a preference should be given to a To-Loc it should be another To-Loc since the parallel relation happens more frequently than the "return" type verb occurs. In our study we tried out both approaches for both location roles.

In the case of loading verbs such as "load" and "put", the To-Loc role has a different resolution strategy. An antecedent is searched in the theme focus list then instrument then ID (in that order).

Return verbs such as "return" and "go back" should also have a unique strategy. Their To-Loc role should be resolved to a previous From-Loc since that is the most likely candidate for a source to return to. This preference has not yet been implemented.

4.2 Instrument and Theme Roles

With nearly 79% of all implicit Instrument roles referring to previous mentions of Instrument, it makes sense to first search the Instrument focus list before searching the Theme and ID lists. The same pattern holds for the Theme role which has 87.5% of its referents in the Theme role of previous utterances. The search preference in this case is Theme, then ID then Instrument.

4.3 Reference Resolution

We combined the above metrics into one parsing and resolution system. From the parsing phase we extract all discourse entities: verbs and NP's. The resolution phase works as follows: for each utterance put all entities, explicit and implicit, onto their respective focus list (so if an Instrument role were processed, its entity would be placed in the Instrument focus list for that utterance. If a role is implicit, we attempt to resolve it using the algorithms from above. After an utterance is processed (focus lists are updated) it is added to the history list and we process the next utterance. The algorithm is summarized below:

Parse and extract incrementally from each utterance all entities. For each entity do:

1. if role is explicit, add it to the appropriate focus list
2. else role is implicit, attempt to resolve given its type:
   (a) **Instrument** Search through current ID focus list. If one is not found then search through history list looking at instrument, then theme, then ID focus lists for each past utterance until an antecedent is found.
   (b) **Theme** Search through current ID focus list. If one is not found then search through history list looking at theme, then instrument, then ID focus lists for each past utterance until an antecedent is found.
   (c) **From-Loc** Search order is determined by strategy: narrative, parallel or most recent.
   (d) **To-Loc** Search order is determined by strategy: narrative, parallel or most recent.

   Next, check if the antecedent is correct and update reference statistics. Then add role to the appropriate focus list.

5 Results and Discussion

5.1 Results

We ran several instantiations of the reference engine; the results of which are displayed in Figure 7. The baseline algorithm eliminates all the individual role algorithms, and replaces them with a single search through all entities ranked in order of recency. The first entity that meets the constraints imposed by the verb is returned as the antecedent (for example, a implicit role reference to a location must be resolve to a NP that is a location).

The next two algorithms test the narrative and parallel strategies for the location roles. For the Instrument and Theme roles, we use the search order described in the preceding section. The final row lists the overall performance by combining the best metrics for each individual role. In this case, the Instrument and Theme ones use the focusing algorithm developed above; From-Loc uses the most-recent strategy and To-Loc's can be resolved using either narrative or parallel methods.

The results show that our focusing-based mechanisms perform much better than the baseline strategy. Most of this boost comes from the improvements in the Instrument and Theme roles. Any analysis of the To-Loc performance is tainted by the fact that there are only five instances in our corpus. So it is hard to make a decision on whether a narrative or parallel strategy is best, though intuitively a parallel one would seem to perform better. Also a most-recent strategy could perform better since it actually does so for the From-Loc case.

It should be noted that the utterances in this domain are fairly simple and that moving to a more complex domain (as well as larger) would allow performance differences to be more visible.

5.2 Error Analysis

In this section, we discuss some of the implicit roles that were resolved incorrectly. Some of them were due to difficulty in annotation, ambiguity in how to deal with the engines and boxcars, and set membership; all problems not related to the focusing algorithm.

5.2.1 Case 1

utt27-A: s: okay so we'll take a boxcar
utt27-B: s: we need one right
utt27-C: s: so we'll take one boxcar +
with us as we go +

Both utt27-A and utt27-C have an implicit Instrument role that refers to an engine from the previous utterance. Utt27-A gets resolved correctly, however utt27-C fails because it finds a suitable antecedent in utt27-B - the Theme, which is the boxcar mentioned explicitly in all three utterances. There are two ways of solving this problem: one is to make the constraint that the same entity cannot be found in both the Theme and Instrument roles of the same verb (unless it is an engine, because one can envision a sentence such as "take the engine to Bath" in which the engine is the Theme but also the Instrument since it is the only entity that can

Algorithm	Instrument	Theme	From-Loc	To-Loc	Overall
Baseline	42.1%	11.1%	88.5%	44.5%	57.1%
Narrative	84.2%	66.7%	84.6%	55.6%	74.6%
Parallel	84.2%	66.7%	61.5%	55.6%	68.3%
Best	84.2%	66.7%	88.5%	55.6%	79.4%

Figure 7: Reference Resolution Results

take anything." The second would be to view utt27-B as an aside and thus skip over it since it really has no content.

5.2.2 Case 2: Merging Antecedents from Different Roles

utt27-C s: so we'll take one boxcar + with us as we go +
...
utt29-B : s:uh to pick up the oranges one hour to load

We want to find the antecedent for the implicit Instrument role in utt29-B. In utt27-C there are two entities - the boxcar and an implicit Instrument (engine 2) that can serve as the antecedent. In our implementation, we only try to find one antecedent. In this case, the Instrument can be interpreted as being both the engine and boxcar and this what the annotation has. One way of handling this problem is to modify the search such that whenever an utterance is encountered where an engine and a container car are joined they become a set. Another option is to consider set membership correct. In this example, the Instrument role is resolved to engine 2. This can be viewed as correct since one can use the engine to pick up the oranges.

5.2.3 Case 3: Verbs with At-Loc and "Return" Type Verbs

utt29-A: s: okay so that'll be two hours to Corning
utt29-B: uh to pick up the oranges
utt29-C: one hour to load
utt29-D: plus two hours to get back

There are two issues here. First is that a preference for the locations for "Return" type verbs has not been encoded. Second, verbs with only one location role, such as "put" or "pick up" are currently annotated as having a From-Loc when a better representation would be At-Loc. In this example, we incorrectly resolve the To-Loc to Corning since it is the first location encountered in focus. One way of solving this is to make a constraint on "Return" type verbs (actually all verbs in general) that states the From-Loc and To-Loc can't be the same.

5.2.4 Other Cases

There are a few cases where we resolve roles incorrectly simply because there are more recent items in focus. Another issue is with backtracking: when the participants are reviewing a plan and time points may come up again and out of order.

5.3 Discussion

It should be noted that this is a preliminary study. We have a very small corpus consisting of simple sentences with a constrained set of verbs and therefore roles. Also, our algorithms are tested on the same corpus that we developed our focusing strategies from. In addition, there is a lot of information that is annotated that is being used to facilitate the reference process. First of all, since all roles are marked as implicit or explicit we simply pick out the implicit ones and try to resolve them. So our recall is 100%. In an advanced system one would expect part of the reference process to decide what the roles for each verb were and which ones are explicit or implicit. The first part is fairly simple since we have a database of all the verbs and their roles. The second one requires looking at the NP's in the verb phrase and deciding what role each NP fits into.

Currently, we are working on automatically annotating a much larger and richer domain to give more reliable results.

The limited conclusion we can draw from the empirical test is that a focusing-based approach appears to work for most of the roles except for From-Loc in which the most-recent strategy beats out the narrative one marginally. However, a more extensive corpus test is needed to confirm this claim.

I would like to thank James Allen for many discussions on implicit role reference. I am also grateful to Mark Core and Donne Byron for their help in revising the paper. Partial support for the research reported in this paper was provided by the DARPA research grant no. F30602-98-2-0133 to the University of Rochester.

References

Nicholas Asher and Alex Lascarides. 1999. Bridging. *Journal of Semantics*, 15:83–113.

D. Byron and J. Allen. 1998. Resolving demonstrative pronouns in the TRAINS93 corpus. pages 68 – 81.

P. Heeman and J. Allen. 1994. The TRAINS93 dialogues. Technical Report TRAINS TN 94-2, University of Rochester.

M. Poesio. 1994. Definite descriptions, focus shift and a theory of discourse interpretation. In *In Proceedings of the Conference on Focus in Natural Language Processing.*

Candace L. Sidner. June 1979. *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. thesis, MIT.

Michael Strube. 1998. Never look back: An alternative to centering. In *Association for Computational Lingusitics*, pages 1251–1257.

Joel R. Tetreault. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Lingusitics*, 27(4).