What's a Reference Resolution Module to do?
Redefining the Role of Reference in Language Understanding Systems

Donna K. Byron
Computer and Information Science Department
The Ohio State University
Columbus, Ohio, USA 43210
dbyron@cis.ohio-state.edu

James F. Allen
Department of Computer Science
University of Rochester
Rochester NY 14627, U.S.A.
james@cs.rochester.edu

Abstract

This paper addresses the role of reference resolution within language understanding software. To process the range of referring expressions found in natural language, we argue that reference resolution must be defined so as to completely rewrite the semantic representation of the referring expression and its referent.

1 Introduction

What is the proper function of reference resolution within the language understanding process? For anaphoric referring expressions (RE's), the problem has been framed in a variety of ways, depending on the goals of the larger process. For a given anaphor A, anaphora resolution has been modeled as a function that:

1. Determines the antecedent of A (the *coreference* model) (Hobbs, 1986; Brennan et al., 1987; Lappin and Leass, 1994; Tetreault, 2001)
2. Determines the referent of A (the *extension* model) (Allen et al., 2000)
3. Determines the logical description of A (the *sense* model) (Hobbs et al., 1988)

Our experience building the reference resolution component within the TRIPS spoken dialog system has shown that none of these models is sufficiently general to handle the range of referring expressions encountered in conversational language. Reference resolution in the TRIPS system is defined in a way that supports the needs of practical dialog systems:

> **Reference resolution** is a process that accepts a term description and rewrites it to account for information available in the discourse context, which may include a suggested extension for the expression.

This paper begins by discussing the formalizations of reference that have been used in the past and their shortcomings when faced with task-oriented conversational language. We then present details of the model that has evolved in the TRIPS system, which covers many different types of referring expressions not possible in other representations.

An extensive body of literature exists that describes different methods for choosing the correct binding of REs (e.g. which person a particular *she* refers to), however, there is little information about how to embed this process within a working system, and what the exact scope of responsibility of the reference module should be. The goal of this paper is to describe the TRIPS-REF module interface in sufficient detail to provide assistance to system builders needing either to create a reference resolution component or to improve an existing one.

2 Background

2.1 The components of a referring expression

A nominal referring expression is a noun phrase that is represented as a variable in the semantic representation of the discourse unit containing it. RE's are compact expressions containing many properties that must be taken into account in their interpretation. Table 1 shows the attributes generally agreed to be important in building a semantic representation of the RE, and Table 2 shows the values of these attributes for several example RE's. Note that the different RE attributes do not map one-to-one with parts of speech. For example, Givenness is indicated by the determiner in *b* and by the head noun in *c*. The modifier *both* indicates both definiteness and set size. For some expressions, such as *b* and *c*, some of the attributes cannot be determined from the surface expression.

2.2 Working with referring expressions

The outcome of reference resolution has been defined in a variety of ways in the past. Although each is appropriate for certain situations, they do not cover the full range of

Attribute Name	Impact on the RE's Interpretation
Specificity	Definite or Indefinite
Number	Singular or Plural, and if plural the set size
Level	Individual or Kind
Properties	Absolute properties such as color and also discourse properties such as comparative relation markers
Givenness	Grammatically marked accessibility (using categories from (Gundel et al., 1993))

Table 1: Elements of a Referring Expression

	RE phrase	Specificity	Number	Level	Properties	Givenness
a	a truck	Indefinite	singular	individual	isa truck	type identifiable
b	those ones from Avon	definite	plural	??	from Avon	familiar
c	them both	definite	plural (size=2)	individual	none	in focus
d	both red trucks	definite	plural (size=2)	individual	color=red isa truck	uniquely identifiable
e	two red trucks	indefinite	plural (size=2)	individual	red isa truck	type identifiable
f	another two trucks	indefinite	plural (size=2)	individual	isa truck other-than the salient truck	type identifiable
g	the three brown ponies at Bath	definite	plural (size=3)	individual	brown, at Bath isa pony	uniquely identifiable

Table 2: Analysis of some example Referring Expressions

phenomena found in conversational language. Here we review the main approaches adopted to date and identify their weaknesses in relation to natural discourse.

The coreference model defines resolution as a function from an RE to another linguistic object that is its antecedent, linking REs together in *co-reference chains* (e.g. (Corporation, 1997; Tetreault, 2001)). Applying this model to the phrase "President Clinton$_i$ and his$_i$ dog" would result in *President Clinton* being identified as the antecedent of *his*, without computing a deeper meaning for either expression.

This model is appropriate for applications that do not link the surface language to referents, such as open-domain information retrieval. It typically excludes REs that do not corefer with another noun phrase (about half of the pronouns in spoken dialog (Eckert and Strube, 2000; Byron, 2000)); therefore, anaphoric REs referring to, e.g. events or situations, or those that stand in some relationship other than coreference with their antecedents, cannot be processed correctly. For instance, a coreference model cannot represent the relationship between a member of a class of objects and a reference to the class itself, as in "It should be easy to travel with **the baby** this summer because **they** don't need much at that age."

Computing coreference relationships alone is inadequate for practical dialog systems because they must compute precisely which objects are referred to. Knowing simply that there are two co-referring phrases, in "Pick up **the other boxcars**$_i$ and take **them**$_i$ to Elmira" is not sufficient to allow the system to assist the user with the task.

The sense model defines resolution as a function from an RE to a logical proposition that becomes part of the representation of the entire utterance (e.g. (Hobbs et al., 1988)). This representation is fed to back-end reasoners for subsequent inference and, where appropriate, unification against a database. For example, the expression "President Clinton" might result in the logical expression:

```
the x: (and (isa x President)
            (name-of x Clinton))
```

This model does not provide adequate support for systems that require exact referents to be found within a discourse, because it ignores the issue of salience. There may be several database objects that unify with the semantics of the RE. Which of those objects is the correct binding will depend on the discourse state. A system that sends logical expressions to back-end reasoners to determine variable bindings must either ignore the discourse salience of candidate referents or require the reasoners to know about the discourse history. We believe discourse information

should be maintained in the language interpretation system and not in agents that are designed to find solutions to the task at hand.

The extension model defines resolution as a mapping from REs to extensions, which are objects from the domain of interpretation, typically provided by a database (e.g. (Allen et al., 1996)). In this model, the semantic representation of the discourse is computed in a context-independent manner, and reference resolution subsequently provides the interpretation function to link variables in the semantic form to their referents. For example, the result of resolving the expression "PRESIDENT PUTIN" is a token from the system's database (e.g. VLADIMIR-PUTIN).

The TRIPS system previously followed this model, as described in (Allen et al., 2000). However, we have found it to be inadequate for several types of anaphoric expressions that we encounter in practical dialog, namely metonymy and reference to semantic Kinds. Consider first an example of metonymy, in which the sentence "Send the ambulance to Pittsford" uses the location name "Pittsford" in the place of a contextually salient accident site in Pittsford. An initial semantic analysis of this sentence might be something like:

x, y
ambulance(x)
city(y)
name-of(y,Pittsford)
Send(x,y)

Under the extension model, reference resolution would be expected only to supply bindings for the variables `x` and `y`. However, if the variable `y` is interpreted as referring to a salient accident site in Pittsford such as Shopping Mall `MASSIVE_MALL24`, this results in a situation where the discourse representation incorrectly states that `MASSIVE_MALL24` is a city named Pittsford.

x, y
ambulance(x)
city(y)
name-of(y,Pittsford)
Send(x,y)

{<x, AMBULANCE1>, <y, MASSIVE_MALL24 >}

In order to make the discourse representation a true description of the objects selected as the variable interpretations, reference resolution must be allowed to rewrite all portions of the discourse representation related to variable `y`.

x, y
ambulance(x)
mall(y)
Send(x,y)

{<x, AMBULANCE1>, <y, MASSIVE_MALL24 >}

Pronominal reference to kinds and one-anaphora creates a slightly different problem in which the initial analysis of the referent's *level* might be incorrect, and needs to be updated after reference resolution is complete. This problem is described in more detail below in Section 3.3.

Our solution to these shortcomings is to model resolution as a mapping from *initial* RE descriptions to *final* logical term descriptions. Reference resolution is responsible for determining the entire logical form for the expression. This logical form includes a referent when one can be found.

3 The TRIPS-REF Model

The TRIPS system is a generic shell for developing practical dialog systems (Allen et al., 2001). A module called TRIPS-REF runs within the language understanding process to interpret REs, including anaphoric expressions. This section describes the TRIPS-REF interface definition. It is based around a logical form language along the lines of (Allen, 1995) that can represent important distinctions necessary for spoken language, such as:

- Givenness indicated by different RE forms, e.g. "the engine" *versus* "that engine" *versus* "that" alone
- Comparative relations such as same, other, different, identical
- Referring expressions that are modified by other referring expressions such as "both of them"

3.1 TRIPS-REF Input

For each RE in the discourse to be interpreted, an RE description structure is created using a format typical of quantified logical forms:

```
RE: RE-form
    variable
    <restrictions on the variable>
```

`RE-form` is one of the following:

1. `THE` - The RE is a definite instance NP
   (a) `PRO` - The RE is a pronoun
   (b) `*IM-PRO*` - The RE is ellipsed
2. `A` - The RE is an indefinite instance NP
3. `KIND` - The RE is a bare plural NP

PRO and *IM-PRO* are subtype of definite REs. Implicit referents are generated for functional terms such as *the other truck* which generates an implicit object for the truck that this truck is not.

The `restrictions` field contains a well formed proposition that includes all of the constraints on interpretation that were expressed in the surface form, using logical connectives `and` and `or` when necessary. Constraints in the TRIPS system are typically three-place predicates in prefix notation. The restrictions can contain any number of constraints such as color, location, height, etc.:

```
(color x RED)
(height x 60cm)
(at-loc x PARIS)
```

Constraints come not only from within the RE but also from adjectives, relative clauses, predicate compliments, prepositional phrase post-modifiers, etc. Special predicates within the restrictions have been developed for a variety of anaphoric constructions:

1. **Type** The `type` constraint specifies a semantic category from the TRIPS semantic hierarchy. For descriptive NPs, it is the semantic type associated with the word sense chosen for the head noun. For proper names, it is pulled from a database, and for pronouns, it might be derived from semantic restrictions.

   The `type` predicate also indicates number agreement. For singular RE's, the `type` is either an atomic semantic category or a disjunction. Disjunctions are common when the type has been derived from a semantic restriction. For example:

   ```
   (type v123 ambulance)
   (type v123 (or ambulance helicopter))
   ```

   The type for plural RE's contains a specification of the set members, using a new variable for the set members and the keyword `SET-OF` as the specifier for sets or `SEQUENCE-OF` for sequences. This allows properties to be attributed separately to the set as a whole and to its members. For the RE "The three red trucks" the restrictions would be:

   ```
   (and (type x (set-of y
                  (and (type y truck)
                       (color y red))))
        (size x 3))
   ```

2. **Name-of** If the referring expression is in the form of a proper name, this comes through as a `name-of` constraint. For example: "Bob"

   ```
   (the x (and (name-of x bob)
               (type x person)))
   ```

3. **Context-Rel** A special constraint `context-rel` contains accessibility markers to indicate the relationship between this RE and the current context. Context-Rel can come from a variety of syntactic positions, such as:

   (a) **Head noun** if the RE is a pronoun. "that"

   ```
   (pro x (and (type x truck)
               (context-rel x that))
   ```

   In this case, the type has been determined from semantic restrictions on the verb.

   (b) **Determiners** this/that
   "that dog"

   ```
   (the x (and (type x dog)
               (context-rel x that)))
   ```

   (c) **Noun modifier** in the case of possessive pronouns
   "my dog"

   ```
   (pro x (and (type x person)
               (context-rel x my)))
   (the y (and (type y dog)
               (owns x y)))
   ```

   The value of `context-rel` is simply the accessibility marker used, and this is converted to the appropriate givenness category within TRIPS-REF by the use of a simple case statement.

4. **Comparatives** Comparative anaphora such as "the other truck" or "another two trucks" generate two RE specifications. The implied contrast object uses the RE-form value `*IM-PRO*` which indicates ellipsis.
   "the other truck"

   ```
   (*im-pro* y (type y truck))
   (the x (and (type x truck)
               (other-than x y)))
   ```

3.2 TRIPS-REF Output

The output from TRIPS-REF follows this same format, but its contents change to reflect the interpretation of the RE rather than its surface attributes. The RE-FORM is replaced with a TERM-CONSTRUCTOR from the set {`THE` | `A` | `KIND` } and the restrictions are supplemented with a `REFERS-TO` assertion that indicates the referent chosen for the variable. Referents can have the following forms:

1. Singular referents have one of these two forms:

   (a) An atomic element of a conceptualization, such as BATH, TRUCK1

   (b) `(elem-of (elements))` used for indefinite expressions where elements contains a list of atomic objects from the referential world that satisfy the referring expression constraints.

2. Plural referents have one of these two forms:

   (a) `(enum-set (elements))` A collection composed of all of the listed elements, where elements contains atomic objects from a conceptualization

   (b) `(some-of (elements))` A collection composed of an undetermined subset of the listed elements, where elements contains atomic objects from a conceptualization. This form is used for plural indefinite descriptions.

Also, any restrictions that reflect surface properties of the expression are not appropriate to maintain in the final meaning representation of the discourse because they should not be sent to back-end reasoners, so they are removed from the restrictions in the final term description. Examples are the `context-rel` restriction, comparative anaphora constraints, and the `sequence` restriction in ordinal expressions.

Because the output from TRIPS-REF is a complete logical term expression, TRIPS-REF is allowed to change any part of the parser's original context-independent analysis of the referring expression. This capability will be demonstrated in the examples in the next section.

3.3 Examples

This section details several examples of TRIPS-REF processing, giving particular emphasis to how this model addresses the shortcomings of previous approaches described above in Section 2.2. Note in particular that the RE attributes in the original analysis reflect the surface properties of the RE, but the final term description output by TRIPS-REF should be a logical proposition that is true of the referent.

1. First, a straightforward example in which the initial RE is simply supplemented with a referent:
   Phrase: "a truck"

```
Initial RE:
(a x (type x truck))

TRIPS-REF output:
(a x
  (and (type x truck)
       (refers-to x (elem-of
          (TRUCK1 TRUCK2 TRUCK3)))))
```

2. The input format allows us to express subtle distinctions between different RE forms that are important for choosing the correct referent, while removing these features from the final meaning representation for the RE. For example:
   Phrase: "the truck"

```
(the y (type y truck))
```

   Phrase: "that truck"

```
(the y (and (type y truck)
            (context-rel y that)))
```

   Phrase: "that"

```
(pro y (and (type y truck)
            (context-rel y that)))
```

   All three of these inputs would produce the same logical form output, although the truck chosen as the referent might be different.
   TRIPS-REF output:

```
(the y (and (type y truck)
            (refers-to y TRUCK1)))
```

3. Next, consider the case in which a pronoun anaphorically refers to a KIND. This happens in habitual sentences about medication scheduling in the TRIPS-Medadvisor system, such as "I need help scheduling my Prinivil$_i$. Do I need to take **that**$_i$ with food?" The RE Prinivil is analyzed as a Kind, so the final term description of the co-refering pronoun should also represent it as a Kind.
   Phrase: "that"

```
Initial RE:
(pro x (and (type x medication)
            (context-rel x that)))

TRIPS-REF output:
(kind x (and (type x medication)
             (refers-to x PRINIVIL)))
```

   This also works when the initial reference is to an individual and the partially anaphoric pronoun refers to that individual's class, as the "the baby ... they".
   Phrase: "the baby"

```
Initial RE:
(the x (type x baby))

TRIPS-REF output:
(the x (and (type x baby)
            (refers-to x CHIP1)))
```

   Phrase: "they"

```
Initial RE:
(pro w
  (and (type w
         (set-of y (type y UNKNOWN))))
       (context-rel w they))
TRIPS-REF output:
(kind w (and (type w baby)
             (refers-to w BABY)))
```

	# Correct	% Correct Representations
Output the Antecedent	0	0%
Output the Sense	104	62%
Output the Extension	124	73%
TRIPS-REF (Logical Expression)	155	92%

Table 3: Example REs analyzed

4. Finally, it is easy to see how this interface can handle cases of metonymy, in which the `type` restriction is modified to be consistent with the referent chosen. Consider the sentence from the TRIPS medadvisor domain "I need help scheduling when to take my prescription". The phrase "my prescription" is a metonymy referring to the medications on the prescriptions. Not modifying the description of this entity will cause problems for back-end reasoners because the SCHEDULE action is defined for medications, but no SCHEDULE action is defined for prescriptions.

```
Initial RE:
(the x (and (type x prescription)
            (owns y x)))

TRIPS-REF output:
(KIND x (and (type x medication)
            (prescribed-
on x PRESCRIPT1)
            (refers-to x PRINIVIL)))
```

4 Evaluation

To illustrate the advantage of the TRIPS-REF interface, we manually analyzed the representations used by each reference model described above in Section 2.2 on a set of dialogs from a medication scheduling domain. The set contains 118 sentences, with 169 total REs (disregarding first/second person pronouns). Table 3 indicates whether each model would result in a correct semantic representation for the discourse unit containing the RE.

5 Acknowledgments

This material is based upon work supported by ONR research grant no. N00014-01-1-1015; DARPA research grant no. F30602-98-2-0133; and NSF grant no. EIA-0080124. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the above-mentioned organizations.

References

James Allen, Brad Miller, Eric Ringger, and Teresa Sikorski. 1996. A robust system for natural spoken dialogue. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL '96)*, June.

James Allen, Donna Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2000. An architecture for a generic dialogue shell. *Natural Language Engineering*, 6(3 and 4):213–228, September.

James F. Allen, Donna K. Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2001. Towards conversational human computer interaction. *A.I. Magazine*, 22(4), Winter.

James Allen. 1995. *Natural Language Understanding, 2nd edition*. Benjamin/Cummings, Menlo Park.

Susan E. Brennan, Marilyn W. Friedman, and Carl J. Pollard. 1987. A centering approach to pronouns. In *Proceedings of the 25th Meeting of the Association for Computational Linguistics (ACL '87)*, pages 155–162.

Donna K. Byron. 2000. Analysis of pronominal reference in two spoken language collections. Technical Report 703, University of Rochester Computer Science Dept.

Science Applications International Corporation. 1997. Muc-7 coreference task definition, July. Available from ftp://ftp.muc.saic.com/pub/MUC/MUC7-guidelines/.

Miriam Eckert and Michael Strube. 2000. Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics*, 17(1):51–89.

Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.

J. R. Hobbs, M. Stickel, P. Martin, and D. Edwards. 1988. Interpretation as abduction. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics (ACL '88)*, pages 95–103, Buffalo.

Jerry Hobbs. 1986. Resolving pronoun reference. In *Readings in Natural Language Processing*, pages 339–352. Morgan Kaufmann.

Shalom Lappin and Herbert J. Leass. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, pages 535–561.

Joel R. Tetreault. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*, 27(4):507–520.