

Right Attachment of Prepositional Phrases by Using Reference Strategy

3R-7

Haodong Wu Teiji Furugori
University of Electro-Communications

1. Introduction

This paper describes a method for disambiguation of post verbal prepositional phrase (PP) modifiers. It employs commonsense knowledge about inherent features and ontological classifications about verbs, nouns and prepositions, as well as inter-concept relationships among them. This knowledge is used in preference strategies which define a set of patterns to attach preposition phrases. According to this method, a text understanding system can not only discard spurious syntactic ambiguities produced by a syntax-driven parser, but also resolve semantic ambiguities by employing preference strategies.

2. Syntactic and Semantical Ambiguities

Syntax-driven parsers produce ambiguities for prepositional phrases. For instance, consider the sentences (1)-(5), all of which take the form of NP-VP-NP-PP. A parser will produce three distinct syntactic structures for (1): the NP constituency (in which the with-phrase is the constituent of the noun phrase headed by /it lock), the VP constituency (in which the with-phrase is the constituent of the verb phrase) and the S-modification (in which the with-phrase modifies the sentence). In above syntactic possibilities, only the VP constituency is semantically plausible. Similarly, only the NP constituency is semantically possible for (2), and only the S-modification for (3).

- (1) Sam opened the door with the key.
- (2) Sam took the lock with the key.
- (3) Sam played tennis in the afternoon.
- (4) He killed a girl with a gun.
- (5) Sam bought the book on pets at that bookstore for three dollars.

In (4), both the NP and VP constituency are semantically possible. Anyhow, most of people tend to interpret it as "He used a gun to murder a girl" (the VP constituency) because there is a strong co-occurrence relation between the verb *killed* and *gun*.

Syntax-driven parses produce many semantical impossible parses. the number goes worse when multiple PPs appear in a sentence such as (5). To solve this problem and to produce a semantically acceptable parse, we need to take into account commonsense knowledge as we human beings employ in the disambiguation process.

3. Using Commonsenses in Preference Strategies

We use commonsense knowledge in attaching PPs in text. In (1), for example, by knowing that *key* is an instrument for opening *door*, we assign the VP constituency for the with-phrase. In (2), the with-phrase is a NP constituency, because *locks* typically have *keys*.

Our disambiguation process of PPs makes use of word level commonsense knowledge. We draw it from three sources: ontological information from EDR concept dictionary about the verb (VERB), direct object (OBJ), preposition (PERP) and object of the preposition (POBJ); inter-concept information from EDR concept dictionary about relationships between the VERB and the POBJ, or between the OBJ and the POBJ; and information about reference strategies in table 1 and 2.

3.1 Ontological Classes in Verb and Noun

The attachment of PP can simply be determined for VERB or OBJ or POBJ in certain ontological class. If the verb is *be*, for example, PP is VP constituency; If the object of the preposition is time, the PP modifies the sentence.

Most of information about ontological class can be found in concept classification in the concept dictionary of EDR, in which a concept is placed somewhere in a hierarchy. A concept inherits some features from its upper level concepts. this means in an example that the concept *girl* belongs to its parent concept *female* (encoded as *female(girl)*), and *female* belongs to *human being*, etc.

3.2 Concept Relations between Verb and Object of Preposition

Semantic or co-occurrence relationship between VERB and POBJ sometimes restricts the attachment of PP to certain constituency. In (6) below, *wrench* is typically an instrument for *fix*, and the PP is inevitably attached to the VP constituency.

- (6) Tom fixed the car with a wrench.

In EDR, the relationship like this is described by a set of concept relators, each relator representing a possible connection. For example, the relation between *fix* and *wrench* is defined as:

{*object.implement.source.goal.scope.a-object.basis*}

Here, the concept relator *implement* means that *wrench* is an instrument for *fix* (see [2] for details).

3.3 Concept Relations between Direct Object and Object of Preposition

The concept relation between OBJ and POBJ plays an important role for PP-attachment also. In (2), the

PP is assigned the NP constituency because a *lock* typically has *key*.

The relation between *lock* and *key* is defined as a-object in EDR, that means that *key* is a part of *lock*.

3.4 Preference Rules for Disambiguation

Preference rules relevant to five prepositions are listed in table 1. On the left hand of each rule, a one-atom predicate on the left hand presents a subclass of concept in the concept hierarchy, and a two-atom predicate describes the concept relation between two atoms. Time(POBJ) means that the object of preposition is a time such as three o'clock. Both construct(POBJ) and place(POBJ) mean that the POBJ is a place where the former is a construction and the later is a natural place.

at-rules:

abstract(POBJ) OR time(POBJ) - >
s_attach(PP)
construct(POBJ) OR place(POBJ) - >
s_attach(PP)
Elsewhere - > np_attach(PP)

of-rules:

mental(VERB) - > vp_attach(PP)
Elsewhere - > np_attach(PP)

in-rules:

abstract(POBJ) OR motion(POBJ) OR
time(POBJ) - > s_attach(PP)
(place(POBJ) & !a-object(OBJ,POBJ)) - >
s_attach(PP)
Elsewhere - > np_attach(PP)

with or without-rules:

implement(VERB,POBJ) - > vp_attach(PP)
a-object(OBJ,POBJ) - > np_attach(PP)
Elsewhere - > s_attach(PP)

Table 1: Preference Rules for PP-attachment

Two global rules are listed in table 2. Here, intransitive(VERB) means that the VERB is an intransitive verb. Null(OBJ) means that the OBJ doesn't exist. For example, Stative(VERB) means that the VERB is a stative verb such as *be*, *situate*, *exist*, *stand*.

- (1) intransitive(VERB) OR null(OBJ) - >
vp_attach(PP)
- (2) stative(VERB) - > vp_attach(PP)

Table 2: Global Rules

In resolving ambiguities, two global rules are attempted in the first place. If they fail, then preference rules relevant to the preposition are attempted. If two semantical possible parses are produced, the one with more relators in its relation set is selected.

4. Right Attachment of PP by Using Preference Strategy

We take a few examples to illustrate how the preference strategy works. In (4), with-rules are accessed as neither of two global rules applies, with-rules are accessed. Two concept relations of a-object

(girl,gun) and implement(kill,gun) are verified in with-rules. The related concept relator sets in the concept dictionary are listed below:

[kill] ↔ [gun]: {implement, goal, scope, a-object, basis}
[girl] ↔ [gun]: {scope, a-object}

Both implement(kill,gun) and a-object(girl,gun) are true. According to with-rules, both the VP and NP constituency are semantically possible. However, relator set between the verb *kill* and *gun* has more relators than relator set between *girl* and *gun*, and therefore the system adopts VP constituency for the PP.

A little more complex examples are shown below.

- (7) Peter called up a friend in Japan.
- (8) Peter called up a friend at home.
- (9) Peter called up a friend in a call box.

Here, each sentence has two PPs. The first PP (begun with *up*) is assigned the VP constituency by using global rule. The second PP is attached according to concept relations,

[friend] ↔ [Japan]: {place, a-object, possessor}
[friend] ↔ [call box]: {place}

The features or ontological classes of POBJs are: place(Japan), construct(home), construct(call box) Place(Japan) and a-object(friend,Japan) are true, so the elsewhere rule applies in (7), and the PP is assigned NP constituency. at-rule checks whether home is a construction in (8). It succeeds, so the PP modifies the sentence. For (9), construct(call box) and !a-object(friend,home) force the PP to modify the sentence.

5. Conclusion

Semantic information (include ontological and generic information, inter-concept relationship, commonsense knowledge about the pragmatic justification) is much more important than syntactic information to attach PP correctly. The preference strategy we presented here can be used as post-processor to select a semantically possible interpretation from the output of any type of parser, or as an embedded part of a parser to attach PPs during the parse process. Because it is not limited to a special domain of text, the method is widely applicable for any text-understanding system.

References

- [1] Dr.K.Dahigren, J.McDowell. "Using Commonsense Knowledge to Disambiguate Prepositional Phrase Modifiers", Proc. AAAI. Philadelphia (1986).
- [2] Japan Electronic Dictionary Research Institute, Ltd. "EDR Electronic Dictionary Specifications Guide", 1993.