

VASISTH: An Anaphora Resolution System

Sobha,L
M.G.University
Technology
Kottayam, Kerala, India
sobha@rocketmail.com

B.N.Patnaik
Indian Institute of
Kanpur,UP, India
patnaik@iitk.ac.in

Introduction

It is well known that natural languages contain anaphoric expressions, gaps and elliptical constructions of various kinds and that understanding of natural languages involves assignment of interpretations to these elements. Therefore it is only to be expected that natural language understanding systems must have the necessary mechanism for the same. Most of the anaphora resolution systems developed so far seem, to the best of our knowledge, to be monolingual, i.e., the systems are language specific and it has not always been shown that these are easily extendable to deal with other languages. VASISTH, in contrast, is a multilingual system, which presently handles two languages from two different language families: Malayalam, from, Indo-Dravidian and Hindi from Indo-Aryan. It can easily be extended to handle other Indian languages as well, more generally, other morphologically rich languages. What further distinguishes VASISTH from other similar systems is that exploiting the morphological richness of the Indian languages, it makes limited use of grammatical rules and uses only morphological markings to identify subject, object, clause etc. It uses limited parsing: the information required from the parser is limited to parts of speech tagging, clause identification, subject of the clauses and person-number-gender of the NPs. Initially VASISTH was developed and tested for Malayalam, and then modified for Hindi. It is well known that pronouns often have more than one possible antecedent: the pronoun resolution mechanism of this system captures the ambiguity but does not resolve it. The system aims to resolve (and achieves considerable success in doing so --- complete success as far as reflexives, reciprocals, and distributives --- are concerned) the antecedent problem of referentially dependent elements such as pronouns --- both anaphoric and cataphoric uses ---, including the so-called “one – pronouns”, reflexives, emphatic and non-emphatic, reciprocals, and distributives, gaps, and certain kinds of ellipsis. This paper confines itself to the handling of only one-pronouns, reflexives, distributives, reciprocals in both the languages (for gaps, ellipsis and more details regarding the pronouns, reflexives, etc., see Sobha, L.(1999)).

Anaphors in Malayalam and Hindi

The third person pronouns in Malayalam (M) and Hindi (H) are given in the first and the second boxes respectively.

Singular	Plural
avan (he) aval(she) atu (it)	avar (they)

Singular	Plural
voh (he/she) ve (he/she – hon)	ve (they)

In Malayalam there is a singular-plural distinction in all third person pronouns and a masculine-feminine distinction whereas in Hindi there is no such distinction. Both in Malayalam and Hindi, the pronouns take the entire range of cases. The pronoun – antecedent relation in Malayalam and Hindi is discussed below.

- 1.(M) mo:han_i avanRe_i kuttiye kantu.
mohan he-poss child-acc see-pst
(Mohan saw his child.)
- 2.(M) mo:han avane_i aticcu ennu kRisnan_i paRannu.
mohan he-acc beat-pst compl krishnan say-pst
(Krishnan said that Mohan beat him.)
- 3.(M)*mo:han_i avalute_i ammaye kantu.
mohan she-poss mother-acc see-pst
(Mohan saw her mother.)
- 4.(H) ra:m_i uski_i kita:b sya:m ko di.
ram he-poss book syam-acc give-pst
(Ram gave his book to Syam.)
- 5.(H) ra:m_i ko ma:lum hai ki mo:han uskeliye_i ka:m nahi kare ga.
ram-acc know-prst copula compl mohan he-ben work neg do-prst
(Ram said knows that Mohan does not work for him.)

In (1) pronoun is in the possessive form and its antecedent is “mo:han” “Mohan”, which is also the subject of the sentence. In (2) the antecedent of the accusative pronoun “avane” “him” is “kRisnan” “Krishnan”, which is the subject of the immediate clause IC (immediate clause is a clause which precedes or follows the one which contains the anaphor) with respect to the one in which the pronoun occurs. Sentence (3) is ungrammatical in the given reading because there is a gender disagreement between the

pronoun “avalute” “her”, which is feminine and the antecedent “mo:han” “Mohan”, which is masculine. From this we can arrive at the following:

- I. A pronoun P is coreferential with an NP iff the followings hold:
 - a. P and NP have compatible PNG (person, number, gender) features.
 - b. P does not precede NP.
 - c. If P is possessive, then NP is the subject of the clause which contains P.
 - d. If P is non-possessive, then NP is the subject of the immediate clause which does not contain P.

In (4), the antecedent of the possessive pronoun “uski” is the subject of the sentence “ra:m”, and in (5), the non possessive pronoun is “uskeliye” “his” and its antecedent is “ra:m” “Ram”, which is the subject of the IC. From this the following emerges:

- II. A pronoun P is coreferential with an NP iff the following hold:
 - a. Pand NP have compatible PNG features.
 - b. P does not precede the NP.
 - c. If P is non-possessive, then NP is the subject of the IC in which P does not occur.
 - d. If P is possessive, then NP is the subject of the IC in multiple clause constructions or the NP immediately preceding the P.

The one-pronouns in Malayalam and Hindi can be classified on the basis of countability. There are pronouns, which are [+C], and there are others which are [-C]. The two realizations of one-pronoun, which are homophonous, are “kuch” [+C] and “kuch” [-C]. The [+C] pronoun takes [+/-animate, +count] NPs as its antecedent and [-C] pronoun takes [-animate, -count] NPs. In the case of the homophonous pronouns the antecedent is [+/-animate, +/-count] NPs. The antecedent NP is the non-subject NP in the immediate clause in which one-pronoun does not occur. Here the use of the world knowledge becomes necessary, which the system under reference does not make use of. The table given below provides information about the one-pronouns:

One-Pronoun in Malayalam and Hindi	Inherent feature of One-Pronoun	NP which can be the antecedent
“orannam” (M) “ek” (H) “one”	+C	+/-animate, +count
“alpan” (M) “thoda”(H) “little”	-C	-animate, +count

“kuRe”(M) (H)	“kuch” “some”	+C, -C	+/-animate, +count
------------------	------------------	--------	--------------------

Given below are examples which demonstrate the above claim regarding [+C] pronouns along with two other claims that hold for [-C] pronouns, namely that (a) the antecedent must precede the pronoun and (b) the one-pronoun must have an explicit antecedent.

- 6.(M) na:n rantu paksikale; ku:ttil kantu. ra:man orannatte; marattil kantu.
I two bird-pl-acc nest-loc see-pst raman one tree-loc see-pst
(I saw two birds in the nest. Raman saw one on the tree.)
- 7.(M) ra:man pa:lu; vanniccu pu:cca alpam; kuticcu.
raman milk buy-pst cat little drink-pst
(Raman bought milk. The cat drank a little.)
- 8.(M) kRisnan ra:manu a:ppl; koduttu avan kuRe; kalyiccu.
krishnan raman-acc apple give-pst he some eat-pst
(Krishnan gave Raman an apple. He ate some (part of it).)
- 9.(M) ammama:r kuttikalkku vellam; koduttu avar kuRe; kuticcu.
mother-pl children-pl water give-pst they some drink-pst
(Mothers gave water to the children. They drank some.)
- 10.(H) mai ne do cidia; ghosle me dekhi, mo:han ne ek; sakha me dekhi.
I-erg two birds nest-loc see-pst mohan-erg one branch-loc see-pst
(I saw two birds in the nest, Mohan saw two on the branch.)
- 11.(H) mo:han ne du:dh; kharida, billi ne tho:da; pi:liya.
mohan -erg milk buy-pst cat -erg little drink-pst
(Mohan bought milk, the cat drank a little.)
- 12.(H) aurato:n ne bacco:n ko pa:ni; diya unhone kuch; pi:ya.
woman-pl-erg children- acc water give-pst they-erg some drink-pst
(Women gave some water to the children, they drank some.)
- 13.(H) mo:han ne ra:m ko se:v; diya usne kuch; kha:ya.
mohan-erg ram-acc apple give-pst he-erg some eat-pst
(Mohan gave Ram some apples, he ate some (of them).)

In (6) "orannatte" "one" refers to the birds in the previous sentence and in (17) "orannum" refers to the books. The one-pronoun "alpam" in (7) refers to "pa:lu" "milk". In sentence (6) the one-pronoun is [+C] and in (7) it is [-C]. In (8) the one-pronoun "kure" "some" refers to "a:ppl" "apple" which is [+C] and (9) it refers to "vellam" "water" which is [-C]. The sentences in Hindi also show the distribution of one-pronouns. In (10) "ek" "one" refers to the "birds" in the previous sentence. In (11) the one-pronoun "tho:da" "little" refers to "du:dh" "milk" In (12) the one-pronoun "kuch" "some" refers to "pa:ni" "water", which is [-C] and in (13) "kuch" "some" refers to apple which is [+C]. From the above we arrive at the following:

II. A one-pronoun corefers with an NP iff

- e. non-subject NP is in the IC of one-pronoun.
- f. NP precedes the one-pronoun.
- g. the one-pronoun and NP agree with respect to C features.

Consider the non-pronominals in both the languages. There are at least three types of non-pronominals in these languages: emphatic, non-emphatic and the possessive. The non-emphatic reflexive in Malayalam is “ta:n”, which can take all the case forms. In Hindi there are four non-emphatic reflexives: “apna”, “apnea:p”, “khud”, “svayam”. The following examples show the non-emphatic reflexive-antecedent relation in both Malayalam and Hindi

- 14.(M) ra:man_i paRannu tanne_i na:ttuka:r aticcu ennu.
raman say-pst self-acc people hit-pst compl
(Raman said that people hit him.)
- 15.(H) mo:han_i ne apnea:p_i ghar cala gaya.
mohan-erg refl house go-pst
(Mohan went home by himself.)
- 16.(H) mo:han ne apne ko aine me dekha
mohan-erg refl-acc mirror-loc see-pst
(Mohan saw himself in the mirror.)
- 17.(H) mai apni ladki ki intazar me hun
I refl girl-acc wait-loc copula-prst
(I am waiting for my daughter.)
- 18.(H) manisa ne apnekeliye ghar kharida
manisha-erg refl-ben house buy-pst
(Manisha bought a house for herself.)

In sentence (14) the non-emphatic reflexive “tanne”, which is in the accusative form, has its antecedent “ra:man” “Raman” in the matrix clause. In the next two sentences (15) and (16) too, the antecedent is the subject of the matrix clause. In (16) the antecedent of “apne ko” is the subject of the clause “mo:han” “Mohan”. The same is the case with (17), where antecedent of “apni” is “mai”. Sentence (18) has “manisa” “Manisha” as the antecedent for the non-emphatic reflexive “apnekeliye”, which is the subject of the clause. From the above examples we arrive at the following.

III. A non-emphatic reflexive R₁ corefers

- h. with an NP iff the NP is the subject of the clause in which the R₁ occurs, or with the head noun N, if the NP is possessive, if the same clause contains both NP and R₁

Turning to the other type of reflexive, the emphatic reflexive in both the languages. In Malayalam the emphatic reflexives are “tannata:n” and “svayam”, the latter borrowed

from Sanskrit and it is morphologically invariant in that it does not take any oblique case forms. In the case of “tannata:n” it takes only two cases namely, nominative and accusative. Both “svayam” and “tannata:n” have free distribution. The emphatic reflexives in Hindi are “apnea:p”, “khud” and “svayam”. The following examples show the occurrence of the emphatic reflexives in both the languages.

- 19.(M) ra:man_i tannata:n_i sku:lil po:yi.
 raman e-refl school-loc go-pst
 (Raman went to school himself.)
- 20.(M) ra:man_i svayam_i sku:lil po:yi.
 raman e-refl school-loc go-pst
 (Raman went to school himself.)
- 21.(H) bacce; ne apnea:p_i khana khaya.
 children-erg e-refl food eat-pst
 (Children ate their food by themselves.)
- 22.(H) manisha; khud_i a:yi thi.
 manisha e-refl come-pst
 (Manisha came by herself.)
- 23.(H) bhagava:n_i svayam_i prakat ho gaye.
 God e-refl manifest be-pst
 (God Himself became manifest.)

In the above sentences (19) and (20), the antecedent of the emphatic reflexives “tannata:n” is the subject of the clause in which it occurs, i.e., “ra:man” “Raman”. The same is the case with the emphatic reflexive “svayam”. Both “svayam” and “tannata:n” are in free distribution. In (21) the antecedent of the emphatic reflexive apnea:p is “bacce” “child” and it is the subject of the clause in which it occurs. The same relation hold in the case of (22) where “manisha” “Manisha”, the subject of the clause is the antecedent of “khud”. In (23) the antecedent of “svayam” is “bhagava:n” which is the subject of the clause in which the emphatic reflexive occurs. Thus:

III An emphatic reflexive R₂

- i. corefers with an NP if the NP is the subject of the clause in which R₂ occurs or
- j. corefers with an N, the head of the possessive NP, iff the NP is the subject of the clause that contains R₂.

The third type of reflexive is the possessive reflexive, which in Malayalam is “svantam”. It is also borrowed from Sanskrit. It behaves like the emphatic reflexive “tannata:n” in the relevant respects. Consider the following examples:

- 24.(M) si:ta_i svantam_i kuttiye aticcu ennu amma paRannu.
 sita refl child-acc beat-pst compl mother say-pst
 (Mother said that Sita beat her (=Sita's) child.)
- 25.(M) ra:man_i svantam_i vanti kRisnanu o:tikka:n kotuttu.
 raman refl vehicle krishnan-dat drive-purposive give-pst

(Raman gave his (=Raman's) vehicle to Krishnan for driving.)

In both the cases the antecedent of the possessive reflexive “svantam” is the subject of the clause in which “svantam” occurs. In (24) “si:ta” “Sita” is the antecedent of “svantam” and in (25), “ra:man” “Raman”. From the above one can conclude that:

IV. The possessive reflexive R₃ corefers with an NP iff NP is the subject of the clause containing R₃

Now consider the reciprocals and the distributives in both the languages. There are several types of reciprocals in Malayalam. The ones which have the most frequent distribution are “ora:l-ora:l” and “ora:l-matte-a:l”. The others are “anyo:nyam”, “tammiltammil”, “parasparam” and “anno:ttum-inno:ttum”. These do not take any oblique case markers. The reciprocals in Hindi are “paraspar” and “ek dusre”. Consider the following:

26.(M) ii kuttikal_i ora:l-ora:le/ora:l-matte-a:le_i atikkilla.
these children each other beat-neg
(These children do not beat each other.)

27.(M) avarkku_i ora:l-ora:le/ora:l-matte-a:le_i istamilla.
they-dat each other like-neg
(They do not like each other.)

28.(M) avar_i anyo:nyam/ tammil tammil/ parasparam/ anno:ttum inno:ttum_i
they each other
sne:hiccu.
like-pst
(They liked each other.)

29.(H) ye bacce_i ek dusare_i se ba:t nahi karenge.
These children each other-acc talk neg do-fut-pl
(These children will not talk to each other.)

30.(H) ye_i ek dusere_i ko pasand nahim karte hain.
These each other-acc like neg do-pl copula-prst
(They do not like each other.)

31.(H) un_i dono me paraspar_i ladhai hui.
These two-loc each other war be-pst
(There was war between them.)

The antecedents of the reciprocal anaphors in (26) and (27) are “kuttikal” “children” and “avarkku” “they” respectively. In (26) and (27) each, the antecedent precedes the reciprocal. The sentence (28) shows the distribution of other reciprocals in Malayalam. In (29) the antecedent of “ek dusre” is “bacce” “children”, which is plural and which precedes the reciprocal. In (30) the antecedent is “ye” “these”, which again is plural and which precedes the reciprocal. Unlike Malayalam, reciprocals in Hindi take all case suffixes. Consider the other reciprocal “paraspar”. In (31) the antecedent of the reciprocal

“paraspar” is “un dono”. Here too the antecedent is plural and it precedes the reciprocal. The reciprocal “paraspar” does not take any case suffix. The coindexing in the above examples shows that in each case the antecedent of the reciprocal anaphor is the subject of the clause in which it occurs. The antecedent has to be plural for a reciprocal and must precede the reciprocal.

- VI. A reciprocal anaphor RA is said to corefer with an NP iff
- k. NP is the subject of the clause, which contains RA.
 - l. NP is plural.
 - m. NP precedes RA.

Turning to the distributive anaphors, in Malayalam distributives are “avar-avar” and “avan-avan”. These distributive anaphors are reduplications of the pronouns “avar” and “avan” and with respect to antecedent choice, they behave like reciprocals. The distributive reflexive “apna apna” is the reduplicated form of the non-emphatic reflexive “apna” and behaves like a reciprocal in the relevant respect.

- 32.(M) ammama:r_i avar-avarute_i kuttikale raksiccu.
 mother-pl their children save-pst
 (Mothers saved their own children.)
- 33.(M) ellavarum_i avan-avanRe_i saukaryam ma:tram no:kki.
 all their convenience alone see-pst
 (All looked at their own convenience.)
- 34.(H) mo:han aur ra:m_i ko apne-apne_i ghar pasand hai.
 mohan and ram-acc each other house like copula-prst
 (Mohan and Ram like their respective houses.)
- 35.(H) mo;han aur ra:m_i ne si:ta aur gi:ta ko apni-apni_i kita:b di.
 mohan and ram-erg sita and geeta-acc each other book give-pst
 (Mohan and Ram gave their books to Sita and Geeta respectively.)

In (32) and (33) the antecedent of the distributive anaphor is the subject of the clause in which it occurs. In (32) the antecedent is “ammama:r” “mothers” and in (33) it is “ella:varum” “all”. In both the antecedent is plural and precedes the distributive anaphor. In (34) the antecedent of the distributive reflexive is “ra:m aur mo:han” “Ram and Mohan”, which is the subject of the sentence and is plural. It precedes the distributive anaphor. In (35) the antecedent is “ra:m aur mo:han” “Ram and Mohan” which is plural and it precedes the distributive element and is the subject of the clause in which the distributive element occurs. The distributive reflexive does not take any case suffixes in Hindi. It takes the gender of the subject NP in a non-ergative sentence and that of the object NP in an ergative sentence. Thus (VI):

VI. A distributive reflexive D is coreferential with an NP iff the following holds:

- o. NP is plural.
- p. NP precedes D.
- q. NP is the subject of the clause in which D occurs.

The Parser.

The parser described here is based on rule-based, not principle based grammars, and in any case, is not based on any of the contemporary models of generative grammars, such as GB, Minimalist, LFG, etc. The hypothesis is that for configurational languages like Malayalam and Hindi, rule-based parsing is easier and faster than principle-based parsing. Keeping in view the requirement of simplicity for parsing, some freedom has been exercised in the construction of the grammar that is used for parsing. For example, although copula is really unrelated to the “dativeness” of the subject, it has been used as a subject identifier for the purpose of parsing. Similarly, possessive NP has been used as a subject-identifier, although a theoretically more correct statement would not characterize “possessive NP” as a subject identifier. Again, all NPs, other than the subject are classified, for parsing purposes, as object NPs. It is a well-known fact that the complementizer occurs in the (embedded) clause final position in Malayalam, however, the rules of parsing here describe the position of the complementizer in different sentences with respect to where it occurs in the whole sentence.

The lexicon, which supports the system, is described below. It contains three thousand words. The lexicon described here is a data structure, which represents a lexeme. Each lexeme has four zones, where each zone corresponds to a slot in the entry zone. The four zones are as follows:

1. Grammatical zone
2. Morphological zone
3. Semantic zone
4. Pragmatic zone

The grammatical zone contains the parts of speech of the word. Consider the following examples:

avan <Noun><Pronoun>
vanu <Verb>
karutta <adj>

In the morphological zone the lexemes has a relevant morphological tag with respect to the grammatical zone. For example, if the grammatical zone is verb then the morphological zone contains the tense and the subcategory transitive/intransitive. If the grammatical zone is noun then the morphological zone contains the case inflexions. The example is as follows.

po:yi <Verb><trans><pst>
avan <Noun><Pronoun><NOM>

The semantic zone gives the information regarding meaning, animate-inanimate etc.

avan <Noun><Pronoun><NOM><+animate><+human>

In the last zone the pragmatic zone the other informations like number, gender, person are given.

avan <Noun><Pronoun><NOM><m><s><third><+human><+animate>

The lexicon is alphabetically ordered with respect to the root and tagged with parts of speech.

The parser implemented here does only partial parsing and the anaphora resolution module works on this parser. The lexicon is very limited. Given below is a sample sentence and the parsed output.

kutti sku:lil vanna ka:ryam adhyapakan
child school-loc come-RP matter teacher

polisino:tu paRannu pakse po:lisu visvasiccilla ennu patranal
police-inst say-pst but police believe-neg compl newspaper

paRannu.
say-pst

(The teacher told the police the fact of the child's coming to the school, which the newspapers say that the police did not believe.)

The parser output of is as follows:

kutti <N><NOM><subj><n><s><third><+human><+animate><+count>
sku:lil <N><LOC><obj><n><s><third><-human><-animate><+count>
vanna <RP>
karyam <N><NOM><obj><n><s><third><-human><-animate><-count>
adhyapakan <N><NOM><subj><m><s><third><+human><+animate><+count>
poli:sinotu <N><INST><obj><n><s><third><+human><+animate><+count>
paRannu <V><pst><tran>
po:lisu <N><NOM><subj><n><s><third><+human><+animate><+count>
visvasiccilla.<V><neg><pst><intran>
ennu <compl>
patRannal <N><NOM><subj><n><pl><third><-human><-animate><+count>
paRannu <V><pst><tran>

The clauses are:

kutti sku:lil vanna ka:ryam
adhyapakan poli:sino:tu paRannu
poli:su visvasiccilla
patRannal paRannu

Anaphora resolution system

The algorithm for identifying the antecedent for each anaphor is given below. The input to the anaphora resolution component is the parsed output from the parser.

- I
 1. Create a list of words of S with NPs, clause and immediate clause identified
(From the parser)
 2. For each pronoun P in S:
 - (a) if (P is poss) then
 - (i) Select the NP which is the subject of the clause in which P occurs.
 - (ii) If P precedes NP in S
then STOP and RETURN.
 - (b) else/*(case(p) != poss)*/
 - (i) identify the immediate clause(s) of the clause in which the pronoun occurs.
 - (ii) if one immediate clause is identified then
 - (a) select the NP which is the subject of this clause.
 - (b) if P precedes NP in S then STOP and RETURN
 - (iii) if more than one immediate clause then
 - (a) select the NP which is the subject of the clause preceding the clause containing P
else/*no immediate clause*/
antecedent lies outside the S
 3. If P and NP do not agree in number, gender, person, then STOP and RETURN
 4. The NP is the antecedent of P.
- II.
 1. Create a list of words of S with NPs, clause and immediate clause identified
(From the parser)
 2. For each reflexive R in S
 - (a) if R is a non-emphatic reflexive then
 - (i) identify the immediate clause of the clause in which R occurs.
 - (ii) select the NP which is the subject of the immediate clause.
 - (b) if R₁ is emphatic or possessive then
 - (i) select the NP which is the subject of the clause in which R₁ occurs.
 3. The NP is the antecedent of R.
 4. For each reciprocal anaphor R' in S
 - (i) select the NP which is the subject of the clause in which R' occurs.
 - (ii) if R' precedes the NP.
The STOP and RETURN
 - (i) if NP is not plural
then STOP and RETURN.
 5. The NP is the antecedent of R'.
 6. For each distributive D in S
 - (i) select the NP, which is the subject of the clause in which R occurs.
 - (ii) if D precede NP
then STOP and RETURN
 - (iii) if NP is not plural
then STOP and RETURN.

7. The NP is the antecedent of D.
- III. 1. Create a list of words of S with NPs, clause and immediate clause identified
(From the parser)
2. Identify the one-pronoun in the list.
 3. Identify the NPs in the IC
 - (i) Identify the non-subject NP identified
 - (a) if one-pronoun is [+c]
 - (i) if not [+count]
 - mark the NP as NON-ANT
 - (b) if one-pronoun is [-C]
 - (i) if not [-animate]
 - (ii) if not [-count]
 - mark the NP as NON-ANT.
 - © if one-pronoun is [+/-C]
 - (i) if not SUBJ
 - mark the NP as ANT
 4. The NP marked as ANT in (3) is the antecedent of one-pronoun.

Conclusion

VASISTH is a system for resolving anaphora, ellipsis and gaps in two Indian languages, namely, Malayalam and Hindi. Computational grammars of both the languages for this specific purpose have been developed. (For details, see Sobha L. (1999)). Only a part of the system has been presented here. Although at this stage this system can handle only two languages, we have reasons to believe that it can be used without major modification for anaphora resolution in all Indo-Aryan, Indo-Dravidian and Indic family of languages, in general all morphologically rich languages.

The limitation of this system is that it uses only syntactic knowledge and does not use any world knowledge for purposes of resolution. This is a serious handicap when it comes to pronominal resolution. However, what is encouraging is that even with this limitation, its performance may not hopefully be termed as discouraging; it shows 93% success in pronoun and one-pronoun resolution and interpretation of gaps, and 96% success in the case of reflexive and ellipsis both. This is for the purpose of information, although there has been no discussion of the ellipses and the gaps in this paper. The present system, which basically handles sentences, and very short discourses, can hopefully be easily extended to cover longer discourses and corpora, but this requires to be examined.

Bibliography

Allen, J. (1987). *Natural Language Understanding*. Benjamin/cumming Publishing company, Inc. California.

Baldwin, Breck. (1997). "CogNIAC: High Precision Coreference with Limited Knowledge and Linguistic Resources", *Proceedings of the ACL'97/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*, 38-45, Spain.

Carbonell, J and R Brown. (1988). "Anaphora Resolution: A Multistrategy Approach", *Proceedings of the 12th International Conference on Computational Linguistics*, 96-101.

Hobbs, Jerry. (1978). "Resolving pronoun references", *Lingua*, 44. 311-338.

Kennedy, Christopher and Branimir Boguraev. (1996). "Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser", *Proceedings of the 16th International Conference on Computational Linguistics (COLING'96)*, 113-118, Denmark.

Lappin, Shalom and M Mccord. (1990). "Anaphora Resolution in Slot grammar", *Computational Linguistics*, 16, 4, 197-210.

Lappin, Shalom and Herbert Leass. (1994). "An Algorithm for Pronominal Anaphora Resolution", *Computational Linguistics*, 20, 4, 535-561.

Mitkov, Ruslan. (1997). "Factors in Anaphora Resolution: They are not the only Things That Matter. A Case Study Based on Two Different Approaches", *Proceedings of the ACL'97/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*, 14-21, Spain.

Mitkov, Ruslan. (1997). "How Far are We from (semi-) Automatic Annotation of Anaphoric Links in Corpora?", *Proceedings of the ACL'97/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*, 82-87, Spain.

Mitkov, Ruslan. (1998). "Robust Pronoun Resolution with Limited Knowledge", *Proceedings of the 18th International Conference on Computational Linguistics (COLING'98)/ACL'98 Conference, (forthcoming)*, Canada.

Mitkov, Ruslan. ".(Forthcoming),"Evaluating Anaphora Resolution Approaches".

Mitkov, Ruslan and Malgorzata Stys. (1997). "Robust Reference Resolution with Limited Knowledge: High Precision Genre-Specific Approach for English and Polish", *Proceedings of the International Conference "Recent Advances in Natural Language Processing"(RANLP'97)*, 74-81, Bulgaria.

Mitkov, Ruslan, Lamia Belguith and Malgorzata Stys. (1998). "Multilingual Robust Anaphora Resolution", *Proceedings of the Third International Conference on Empirical Methods in Natural Language Processing (EMNLP-3)*, 7-16, Spain.

Sobha L, B.N. Patnaik. (1998). “An Algorithm for Pronoun and Reflexive Resolution in Malayalam”, *Proceedings of the International Conference on Computational Linguistics, Speech and Document processing*, C63-66.

Sobha L, B.N.Patnaik. (1999). “One-pronoun resolution in Malayalam”, *Indian Linguistics*, 60, 1-4, 137-144.

Sobha L, (1999) “Anaphora Resolution In Malayalam and Hindi” Doctoral dissertation submitted to Mahatma Gandhi University, Kottayam , Kerala.