Resolving Pronominal Anaphora in Hindi Using Hobbs' Algorithm

Resolving Pronominal Anaphora in Hindi Using Hobbs' Algorithm

KAMLESH DUTTA
National Institute of Technology, Hamirpur (HP), INDIA

NUPUR PRAKASH
GuruGobind Singh IndraPrastha University, New Delhi, INDIA

SAROJ KAUSHIK
Indian Institute of Technology, Delhi, INDIA

ABSTRACT

This paper presents the application of Hobbs algorithm for pronominal resolution in Hindi. Hobb's algorithm makes use of syntactic information rather than semantic information can thus be used as baseline algorithm. The algorithm has been adapted for Hindi language taking into account the roles of subject, object and its impact on anaphora resolution for reflexive and possessive pronouns. In case, if subject is dative, then possessive and reflexive pronouns loose the complimentary distribution, where reflexive pronoun binds with preceding nominative as well as dative noun phrase within a sentence, possessive pronoun extends the binding to the previous sentence as well.

INTRODUCTION

An extensive research work in the development of computational solutions to automatic anaphora resolution (Hirst 1981) in English, and many other languages is going on (Mitkov 2002). Hobb's algorithm (Hobbs 1976) does not make use of semantic information and is based on syntactic information. It is considered to be computationally an economical algorithm (Mitkov 2002). Though Hirst has reported limitations of this algorithm (Hirst 1981) the algorithm can be used as a baseline for comparison with other approaches. The purpose of this paper is to describe anaphoric resolution for pronominal anaphoric reference for Hindi language texts. The suitability of the application of Hobb's algorithm for Hindi is explored.

Anaphor resolution studies for Hindi language has been presented in (Sobha and Patnaik 2002) using rule based approach, (Sinha 2002) corpus based studies, and (Prasad 2000; Prasad 2003) using centering theory. Prasad's thesis work is based on the principle that the grammatical function is important for discourse salience in Hindi Language (Prasad 2003).

The aim of this paper is to implement a system that is based on Hobb's algorithm for pronominal anaphora in Hindi. A syntactic rule based algorithm is run on manually parsed sentences. Hobbs tested his algorithm for the pronouns *he, she, it*, and *they*, successfully 81.8% of the time. The algorithm is adapted successfully for those languages (e.g. Chinese), which have similar Subject-Verb-Object (SVO) structure and follow a fixed word order. The accuracy of the algorithm on overt, third person pronouns at the matrix level in Chinese (Converse 2005) has been reported to be 77.6%, and the accuracy for resolving matrix-level zero pronouns as 73.3%.

Spanish and Turkish languages being free-word order languages poses inherent difficulties for the application of Hobb's algorithm. Success of the algorithm in its original form as reported in case of these languages is limited and the algorithm has been used after modifications. Implementation for Spanish (Palomar et. al, 2001) does not uses full parsing and the specifications are adjusted searching Noun Phrases (NPs) from left to right, and testing those included in an NP and interrupting when an NP agreed in gender and number with the anaphor. The success rate is reported as 62.7%. The algorithm proposes full parsing and incorporates modifications to deal with null pronoun and pronominal possessives (Tüfekçi and Kılıçaslan 2005) for Turkish language. Hindi language is also a free word order and reflexive and possessive pronouns find their antecedents somewhat in a similar manner. However the role of subject and object become prominent in Hindi.

We first give a brief description of some key grammatical aspects from the pronominal perspective of the language, and then describe the application of Hobb's algorithm for Hindi.

CLASSIFICATION OF ANAPHORA AND PRONOUNS IN HINDI

Pronouns in Hindi exhibit a great deal of ambiguity. Pronouns in the first, second and third person do not convey any information about the gender. In Hindi there is no difference between "he" and "she". "veh" is used for both and gender is decided by verb form. With respect to number marking, while some forms, like "usne"(he), "usko" (him), are unambiguously singular but some forms can be both singular and plural, like "unhone" (he) (honorific)/they, or "unko" (him) (honorific)/them. The summary of comparison of pronominal anaphora for third person paradigm in English and Hindi is shown in table 1.

Honorificity is marked on animate nouns. However, all the nouns have semantic gender (masculine or feminine), though there are no overt morphological reflexes of this. First and second person pronouns are systematically excluded from this study.

Pronominal Anaphora in English	**Pronominal Anaphora in Hindi**
He(masculine third person, personal pronoun) She (feminine third person, personal pronoun) It (neutral third person, personal pronoun)	veh (third person masculine, feminine and neutral gender, singular personal pronoun)
He, she (honorific, singular, personal pronoun)	ve (honorific singular personal pronoun, personal pronoun)
He (proximate masculine, singular, personal pronoun) She (proximate feminine, singular, personal pronoun)	is-ne (proximate masculine, feminine, and neutral, singular)
He, she (honorific, singular, personal pronoun)	inhon-ne (honorific singular)
His, Her, its (personal pronoun)	us –(case marker-)
Him Her (personal pronoun)	usko
This (Third person proximate singular, demonstrative pronoun)	yeh
That (Third person remote singular, demonstrative pronoun)	veh
They (third person remote plural, personal pronoun) Those (third person remote plural, demonstrative pronoun)	ve (third person remote plural)
They (third person proximate plural, personal pronoun) These (third person proximate plural, demonstrative pronoun)	ye (third person proximate plural)
Them (third person, remote plural, personal pronoun)	un-ko /unse
Their (third person, remote plural, personal pronoun)	unka/unki/unke
Himself (reflexive pronoun, masculine, singular Herself (reflexive pronoun, feminine, singular) Itself (reflexive pronoun, neutral, singular) Themselves (reflexive pronoun, plural)	apne-aap apna/apni/apnee swayam khud

Table 1. *Pronominal Features in English and Hindi for the 3rd Person Paradigm*

The verb, upon conjugation, usually indicates the difference in the gender and thus reduces some ambiguty introduced by the pronominal system. Hindi has verb agreement with the subject or the direct object. The cases in Hindi indicate the relation of a noun or pronoun with the neighboring word. The case marker is added separately and the pronoun modifies accordingly. The agreement inflection is marked for person, number, and gender. Various inflections after adding case marker to root word veh (he) is shown in table 2.

S.NO.	CASE		SINGULAR	PLURAL
1	Nominative Case	He/she	veh, vo	ve, vo
2	Ergative Case	He/she	us-ne	unhon-ne
3	Accusative Case	him	use, usko	unhein, unko
4	Instrumental Case	With (him)	usse, uske dwara	unse, unke dwara
5	Dative Case	Him/he	usko, use, uske liye	unko, unhein, unke liye
6	Ablative Case		usse	unse
7	Genative Case	his	uska, uski, uske	unka, unki, unke
8	Locative Case	In/at	usmein, us per	unmein, un per

Table 2 *Case marking for vah (he)*

Although natural languages are not purely context-free in nature, for the sake of correctness, following context free grammar for generating the surface structures of a fragment of Hindi is defined. :

<S> → <NP> <VP>
<NP> → <NP_nom> | <NP_erg> | <NP_acc> | <NP_instr>
| NP_dat> | <NP_abl> | <NP_gen>
| NP_loc | [(<PP>)*] <Nbar> [<postp>]
| <pronoun> [<postp>]
<VP> → <NP>* <VP> | [(<PP>)*] <VP> |
| [(<adverb>)*] <verb> [<conjugation>]
<Nbar> → [(<adj>)*] <noun>
<PP > → <NPj_case> <postp>
| [<number>] <noun>* <postp>
<verb> → jaa | uttar| chal | …
<adj> → sunder| lambaa| acchhaa | …
<adverb> → tez | dheere | …
<number> → ek| do| teen|…
<conjugation>→ hai | hun |ho| hain | thaa| the| thii | thiin
<postp> → ka | ke| ki | ko| mein| par |
<pronoun> → veh | us-ne | use | usko | us-se | uske dwara
| uske liye | uska | uski | unke | us-mein | uspar | us|..

where

S	-	sentence
NP	-	Noun phrase
NP_nom	-	Nominative Noun phrase
NP_erg	-	Ergative Noun phrase
NP_acc	-	Accusative Noun phrase
NP_instr	-	Instrumental Noun phrase
NP_dat	-	Dative Noun phrase
NP_abl	-	Ablative Noun phrase
NP_gen	-	Genitive Noun phrase
NP_loc	-	Locative Noun phrase
VP	-	Verb phrase
PP	-	Postpositional phrase
Nbar	-	noun forms
postp	-	Postposition
adj	-	adjective
[]	-	Optional
*	-	Occurrence may be more than one or empty
|	-	Alternatives

Hobbs Algorithm for Pronouns in Hindi

Hobbs algorithm (Hobbs 1978) for the resolution of pronominal coreference in English is applied to the surface parse tree of sentences in a text. Leaves of the parse tree, in left to right order represent the original sentence and the surface parse tree exhibits the grammatical structure of the sentence – subject, object verb, adverb, etc. The resolution procedure traverses full parse tree starting from the pronoun looking for noun phrases that could be possible antecedents and adding them to a list of candidates left to right breadth-first in the sub tree subject to the constraints defined by the algorithm. The algorithm chooses as the antecedent of a pronoun P the first NP in the tree obtained by left-to-right breadth-first traversal of the subtree to the left of the path T such that

- T is the path from the NP dominating P to the first NP or S dominating this NP,
- T contains an NP or S node N-bar which contains the NP dominating P, and
- N does not contain NP. If an antecedent satisfying this condition is not found in the sentence containing P, the algorithm selects the first NP obtained by a left-to-right breadth first search of the surface structures of preceding sentences in the text.

As proposed by (Chomsky 1995), NP node is assumed to have an N-bar node below it, to which prepositional phrase containing an argument of the head noun may be attached. In English truly adjunctive prepositional phrases are attached to the NP node. On the other hand in Hindi post positional phrases are attached to the NP node as postposition follows the noun. More over Hindi has rich case system, though case marking for noun phrases is not obligatory. In Hindi, both animate and inanimate objects can be case-marked, but case marking of inanimates is possible only for

definite, while case marking of humans (and some animates) is possible for indefinites as well. Further, case marking for human (and some animates) is generally obligatorily, while case marking of inanimates is generally optional. Characterizing these systems then requires reference both to degree of animacy and degree of definiteness. Hindi marks some direct objects, but not all, with the postposition *ko* (again, the same postposition used for indirect objects). In general, Hindi permits more overt case marking of objects. Hindi requires extensive case marking for human-referring objects as demonstrated in table 2.
We take example of (Hobbs 1976) and also represent their equivalent forms in Hindi.

E1 Mr. Smith$_i$ saw a driver$_j$ in his$_{i/j}$ truck.
E2 Mr. Smith$_i$ saw a driver of his$_i$ truck.

In sentence E1, “his” may refer to Mr. Smith or the driver but in sentence E2, “his” may not refer to the driver as shown in fig. 1 and fig. 2 respectively.

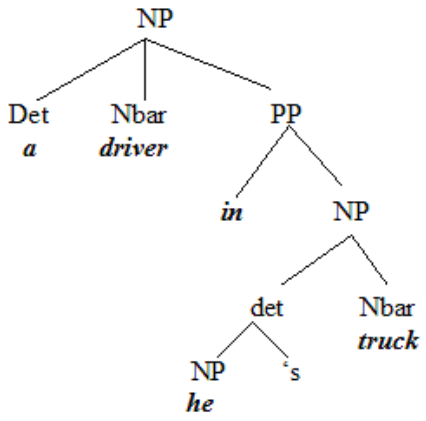

Fig. 1 *Structure for NP of sentence E1*

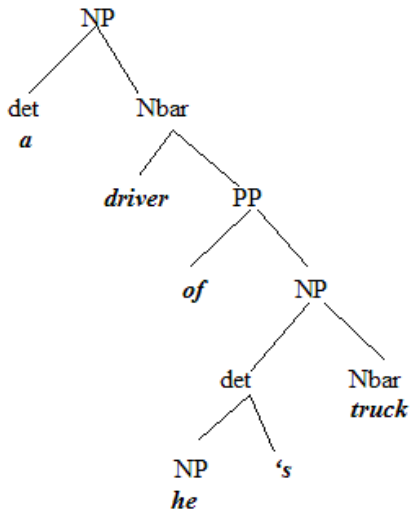

Fig.2 *Structure for NPs of sentence E2*

We consider eight possible translations in Hindi of sentence E1, though by scrambling the word order, many more permutations are possible.

H1: Mr. Smith$_i$ ne driver$_j$ ko uske$_{*\ i/j/k}$
Mr. Smith –ERG driver-ACC his/her-GEN-3-SG

truck mein dekhaa.
truck-POSS-3-LOC see-PAST

Pronominal possessors in Hindi have anti subject orientation. The pronoun ‘uske” occurring in indirect object cannot be co-referential with “Mr. Smith” which is an ergative subject in the sentence H1. It can be co referential with “driver” in the present sentence or another person in the previous sentence. The syntactic tree structure for the sentence H1 is shown in Fig. 3.

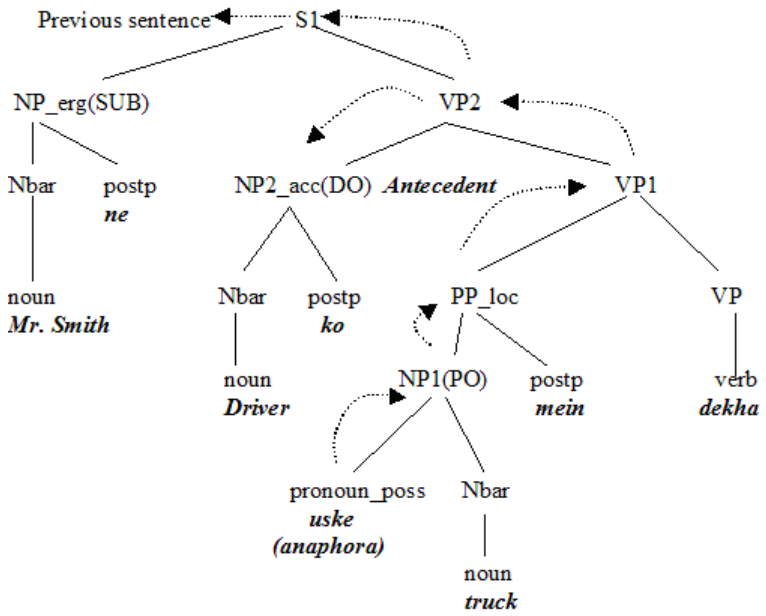

Fig. 3 *The structure of sentence H1 and co reference resolution*

H2: Mr. Smith$_i$ ne driver$_j$ ko apne $_{i/*j/*k}$
Mr. Smith –ERG driver-ACC his/her-GEN-3-SG

truck mein dekhaa.
truck-POSS-3-LOC see-PAST

Reflexives or possessive anaphors in Hindi are subject oriented and tend to bind with the matrix subject. Reflexives will never have a binding beyond the present sentence. Hence possibility of binding of driver beyond the sentence is overruled. In sentence H2, reflexive pronoun “apne” is coreferential with “Mr. Smith” as shown in Fig. 4.

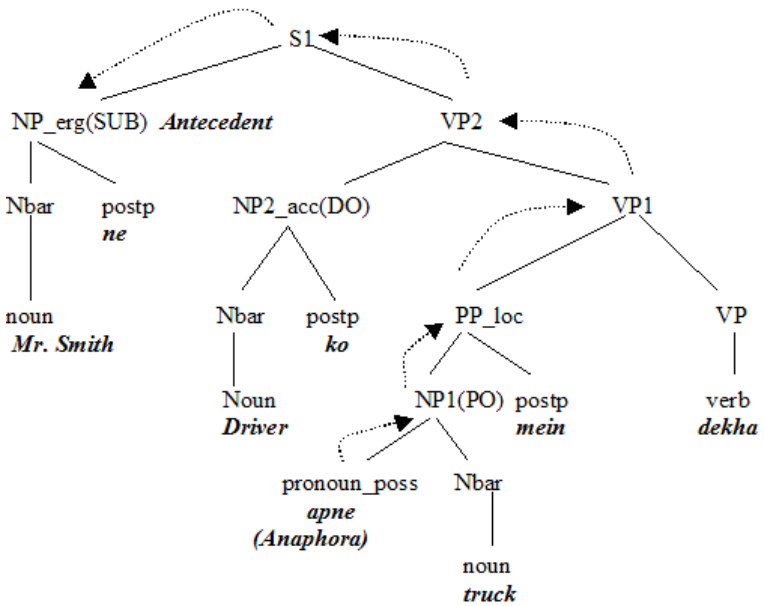

Fig. 4 *The structure of sentence H2 and co reference resolution*

H3 : Mr. Smith$_i$ ko driver$_j$ uskee $_{i/j/k}$
Mr. Smith –DAT driver-NOM his/her/self

truck mein dikhaa
truck-POSS-3-LOC see-PAST

With dative subject the obviation of pronominal possessor of object is lost. All the previous occurrences of possible antecedents are possible. Hence "uske" may be coreferential with Mr. Smith or "driver" are some noun phrase in previous sentence as shown in Fig. 5 for sentence S3.

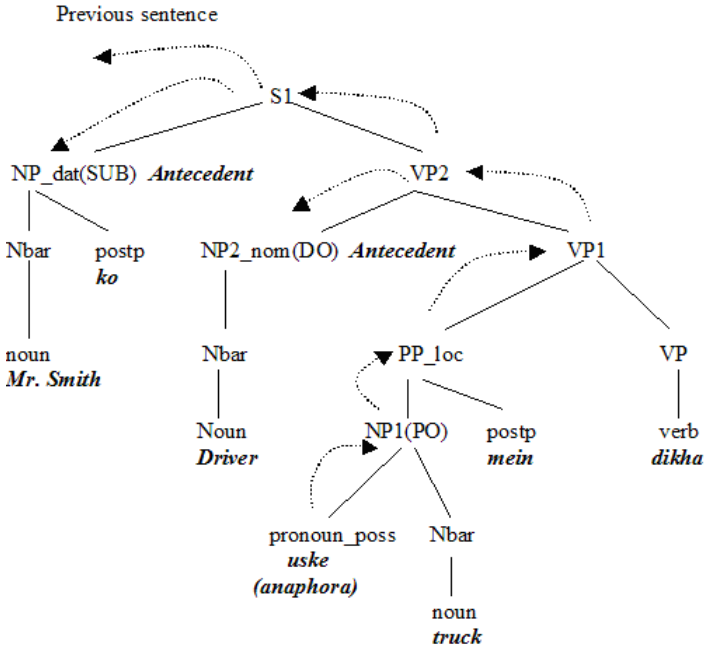

Fig. 5. *The structure of sentence H3 and corefeence resolution*

H4: Mr. Smith$_i$ ko driver$_j$ apnee$_{i/j/*k}$
Mr. Smith –DAT driver-NOM his/her/self

truck mein dikhaa
truck-POSS-3-LOC see-PAST

If subject is dative and indirect object is nominative, reflexive pronoun "apnee" can be coreferential with "Mr. Smith" or "driver" as shown in Fig. 6 for sentence H4.

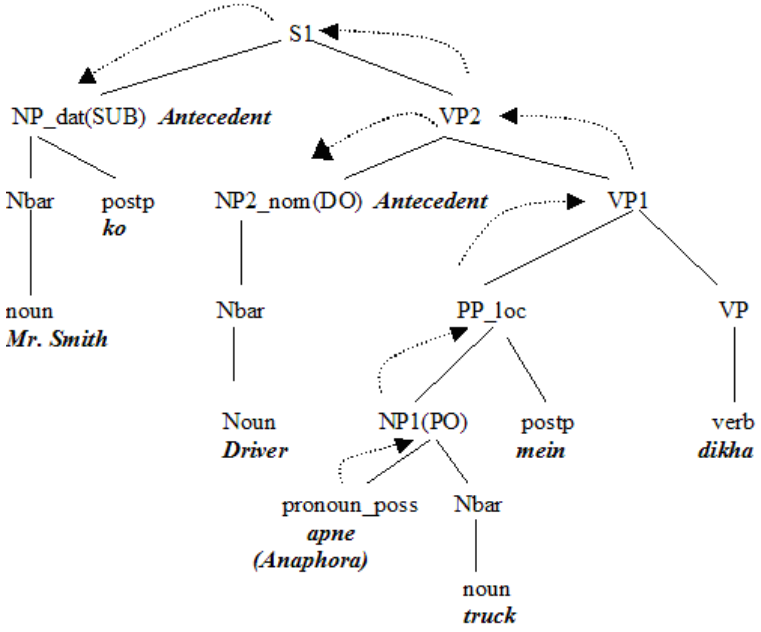

Fig. 6. *The structure of sentence H4 and corefeence resolution*

H5: Mr. Smith$_i$ ne uske$_{*i/*j/k}$ truck mein
Mr. Smith –ERG his/her-GEN-3-SG truck-POSS-3-LOC

driver ko dekhaa.
driver-ACC see-PAST

In sentence H5 “uske” may be co referential with another phrase in the previous sentences. The ergative subject does not bind “uske” in the present sentence but look for a binding in previous sentence as shown in Fig. 7

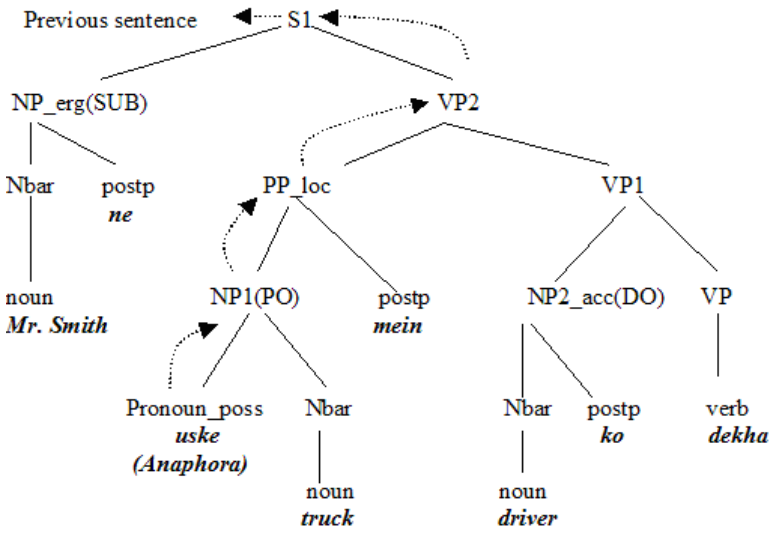

Fig. 7. *The structure of sentence H5and corefeence resolution*

H6: Mr. Smith$_i$ ne apne$_i$ truck mein
Mr. Smith –ERG his/her-GEN-3-SG truck-POSS-3-LOC

driver ko dekhaa.
driver-ACC see-PAST

In sentence H6 “apne” is a reflexive pronoun and binds with immediate subject and refer to “Mr. Smith” as demonstrated in fig. 8.

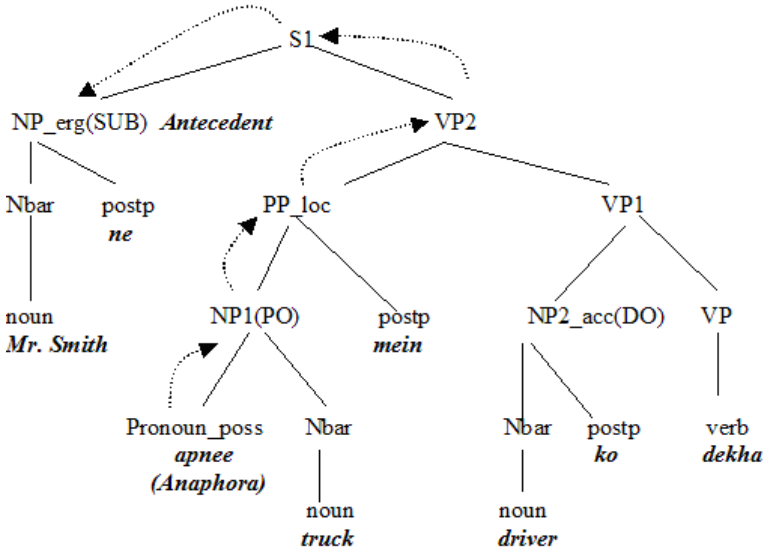

Fig. 8. *The structure of sentence H6 and corefeence resolution*

H7: Mr. Smith$_i$ ko uske$_{i/*j/k}$ truck$_j$ mein
Mr. Smith –DAT his/her-GEN-3-SG truck-POSS-3-LOC

driver dikhaa.
driver-NOM see-PAST

In sentence H7 “uske” may be co referential with Mr. Smith or another phrase in the previous sentence as demonstrated in Fig. 9.

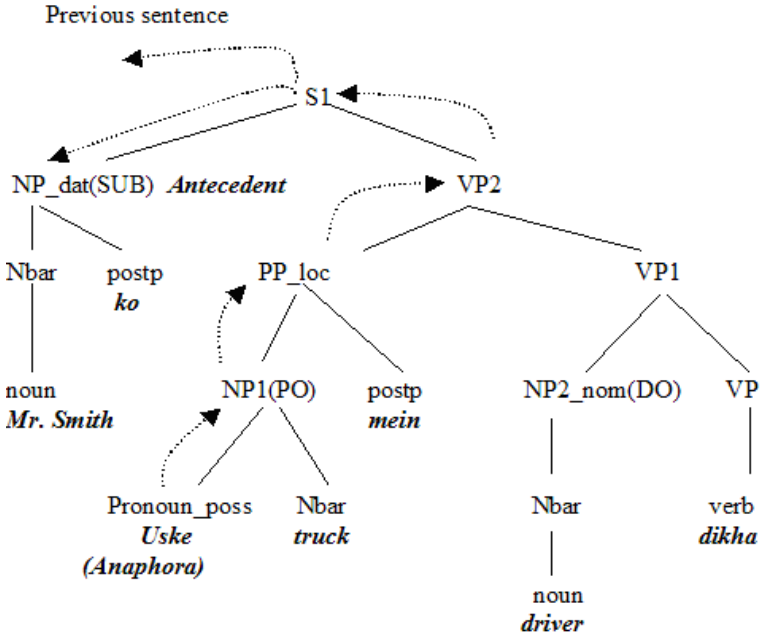

Fig. 9. *The structure of sentence H7 and corefeence resolution*

H8: Mr. Smith$_i$ ko apne$_{i/*j/*k}$ truck$_j$ mein
Mr. Smith – DAT his/her-GEN-3-SG truck-POSS-3-LOC

driver dikhaa.
driver-NOM see-PAST

In sentence H8 “apne” may be co referential with “Mr. Smith” but with no other entity as demonstrated in fig. 10.

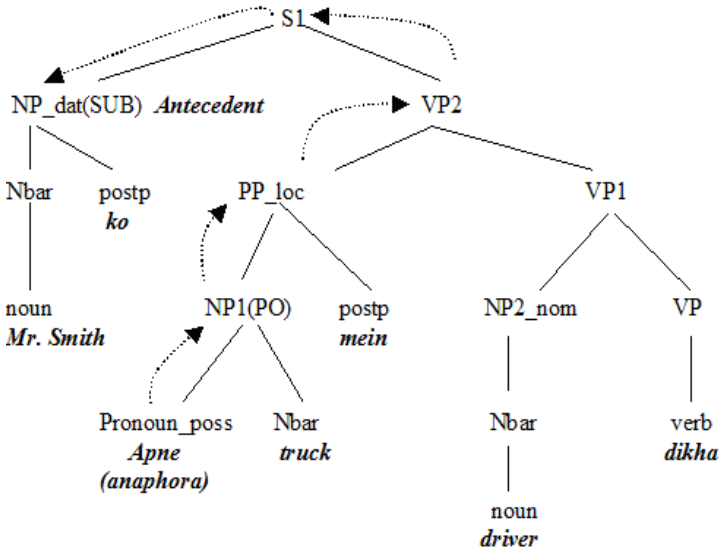

Fig.10 *The structure of sentence H8 and coreference resolution*

Possessive pronouns/anaphors in Hindi show gender agreement with the head nouns, which arises from the inherent gender in Hindi nouns (Prasad, 2003). However, the noun agreement does not provide any grammatical information for the interpretation of the pronoun/anaphors. Sentences H9-13 give examples of possessive- head gender agreement.

H9 uskii/apnii maa
PRO-POSS-3.sg.fem. mother-3.sg.fem.
"his/her/self mother"

H10 uskii/apnii KitaabeN
PRO-POSS-3.pl.fem books-3.pl.fem
"his/her/self books"

H11 uske/apne pitaa
PRO-POSS-3.sg.masc.hon. father-3.sg.masc.hon.
"his/her/self father"

H12 uske/apne kapRe
PRO-POSS-3.pl.masc. clothes-3.pl.masc.
"his/her clothes"

H13 uskaa/apnaa truck
PRO-POSS-3.sg.masc. truck-3.sg.masc.
"his/her/self head"

From these examples we observe that in Hindi , for an antecedent to bind with "apne", it must be either the logical or the grammatical subject (Bhatt & Anagnostopoulou 1976; Butt, 1993; Davison 2000, 2001a, 2001b; Dayal 1994; Kachru, Kachru,, & Bhatia,1979). Nominative case is an absolute criterion for subject status in English. But in Hindi with other case possibilities for subject, nominative case does not uniquely mark grammatical subjects. With ergative or nominative subjects reflexive pronouns ("anaphors") and possessive pronouns ("pronominals") are in complimentary distribution when it comes to expressing coreference relation. Anaphors are able to find the antecedent in a local domain. Possessive pronouns look for antecedent farther. If there is more than one c-commanding antecedent in the local binding domain, the subject is chosen as the antecedent for the reflexive. In simplex clauses, there is only one possible interpretation for the reflexive, even if the sentence has more arguments. In complex sentences, there can be multiple antecedents. Each subject is a possible antecedent provided that the intervening clause boundaries are non-finite. Even when there are multiple antecedents, which would make sense as antecedents, only a subject is the controller.

Hence in Hindi, in case subject is dative, possessive pronouns and reflexive pronouns loose the complimentary distribution. Where reflexive pronoun binds with preceding nominative as well as dative noun phrase within a sentence, possessive pronoun extends the binding to the previous sentence as well.

Hobb's Naïve Algorithm For Hindi

Based on the observations in previous section, we have adapted Hobbs Naïve Algorithm so that it can be applied to Hindi.

1. Begin at the NP node, which immediately dominates a pronoun 'veh', 'yeh', 'uska/uski/uske' or 'apna/apne/apnii'. If NP node immediately dominates a pronoun, continue to step 3.

2. Go up the tree to the first NP or VP node encountered. Call this node *X* and call the path used to reach it *p*.

3. If the pronoun is 'apna/apnii/apne', continue to step 6.

4. If *X* is an NP node of the Direct Object, traverse all branches below node *X* to the left of path *p* in a left-to-right, breadth-first fashion. Propose as the antecedent any accusative NP node or nominative NP node, which is immediately dominated by X, or propose as the antecedent any accusative NP node or nominative NP node that is encountered which has an NP, VP or S node between it and *X.*

5. If *X* is an NP node of the Subject, traverse all branches below node *X* to the left of path *p* in a left-to-right, breadth-first fashion. Propose as the antecedent any dative NP node, which is immediately dominated by X, or propose as the antecedent any dative NP node that is encountered which has an NP, VP or S node between it and *X.*

6. From node *X* go up the tree to the first NP, VP or S node encountered. Call this new node *X,* and the path traversed to reach it *p.* If *X* is an NP node or a VP node, continue to step 3. If *X* is an S node, continue to step 7.

7. If the pronoun is "apna/apne/apnii", the antecedent is a dative, ergative or nominative case-marked NP node preceding it or used later on in the sentence.

If the pronoun is not "apna/apne/apnii", continue to step 8.

8. If node *X* is the highest S node in the sentence, traverse the surface parse trees of previous sentences in the text in order of recency, the most recent first; each tree is traversed in a left-to-right, breadth-first manner, and when an NP node is encountered, it is proposed as the antecedent. If *X* is not the highest S node in the sentence, continue to step 9.

9. From node *X*, go up the tree to the first NP, VP or S node encountered. Call this new node *X*, and call the path traversed to reach it *p.*

10. If *X* is an NP node and if the path *p* to *X* did not pass through the Nbar node that *X* immediately dominates, propose *X* as the antecedent.

11. If *X* is an NP node and if the path *p* passed through the N-bar node that *X* immediately dominates, traverse all branches below node *X* to the *left* of path *p* in a left-to-right, breadth-first manner. Propose any NP node encountered as the antecedent.

12. If *X* is a VP or S node, traverse all branches of node *X* to the *right* of path *p* in left-to-right, breadth-first manner, but do not go below any NP or VP or S node encountered. Propose any NP node encountered as the antecedent.

13. Go to step 8.

A tree is traversed in breadth-first search manner where every node of depth n is visited before any node of depth n-1. Steps 3 and 7 of the algorithm take care of the level in the tree where a reflexive pronoun would be used. Step 9-13 cycle up the tree through S and NP nodes. Step 8 searches the previous sentences in the text.

DATA EXTRACTION AND ANALYSIS

Several sentences from (Hobbs 1976) were taken and translated in Hindi. The modified algorithm was applied on these translated sentences. The algorithm worked successfully on some of the sentences.
For example in the sample set SE3 Hobb argued that "himself" is acceptable in place of "him". But in compound sentences the occurrence of non-reflexive pronoun is more acceptable. In the translated sample SE3 we observed the same behaviour. In another sample set SE4 the Hobbs algorithm found incorrect antecedent "the man" for the pronoun "him". For Hindi translated example SH4 the algorithm resolved "uski" with "maa" (mother) which is again an incorrect resolution. Since "uski" can be used for both the genders, the resolution process could not disambiguate the gender and instead of "john" "maa"(mother) becomes an antecedent.

Further it has been observed that though in some cases verb phrase as a whole or the auxiliary verb convey information about the gender, in others it does not as demonstrated in table no.

English sentence	**Hindi Sentence**	**Observation**
He is angry	veh gussa he.	Gender differentiation is not available
She is angry	veh gussa he.	
He was angry	veh gussa **tha**.	Gender information is obtained from the auxiliary verb tha (male) and thii (female)
She was angry	veh gussa **thii**.	

Table 3 *Role of verb phrase(auxiliary verb) in gender disambiguation for pronoun "veh"(he/she)*

CONCLUSION

The purpose of the present work is to implement a syntactic anaphor resolution system that could be used as a baseline against which to measure other automated approaches to Hindi anaphora resolution. We have implemented a modified version of Hobbs' Naïve Algorithm for Hindi by taking into account the free word order and grammatical role in pronoun resolution in Hindi. The role of subject and object in Hindi are found to have significant impact on anaphora resolution for reflexive and possessive pronouns. Role of grammatical features play important role in word sense disambiguation, which can be used in resolving the pronouns. Addition of feature set will further improve our algorithm.

At present, we have tested the algorithm for limited set of sentences defined by our grammar. Our experiments have shown that the proposed algorithm has some limitations, which can be overcome by including the semantic information. Also the performance of the algorithm needs to be tested on fully parsed corpus, which as on today is not available. Once a sizeable corpus is available, the performance of the proposed algorithm can be ascertained. We are working toward the development of anaphoric corpus, which can be used in our future work of anaphor resolution.

REFERENCES

Bhatt R., and Anagnostopoulou,E., (1996). Object Shift and Specificity:Evidence from ko-phrases in Hindi, in L. M. Dobrin, K. Singer, and L. McNair, eds., *Papers from the 32nd Regional Meeting of the Chicago Linguistics Society*, Chicago Linguistics Society, Chicago, 11–22, 1996

Butt, M., (1993). Object Specificity and Agreement in Hindi/Urdu, *29th Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago, pp. 89-103,

Chomsky,N.,(1995). The Minimalist Program, *Cambridge,MIT Press*.

Converse,S.P.,(2005). Resolving Pronominal References in Chinese with the Hobbs Algorithm, *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing* 2005

Davison, A. (2000). Lexical anaphors in Hindi/Urdu, In: K. Wali, K.V. Subbarao, B. Lust and J. Gair (eds.) *Lexical anaphors and pronouns in some South Asian languages: a principled typology*. Berlin: Mouton de Gruyter, 397-470.

——. (2001a). the VP structure and case checking properties of sentences with non-nominative subjects: Hindi/Urdu, University of Iowa April 23, 2001

——. (2001b). Long-distance anaphora in Hindi/Urdu. In P. Cole, G. Hermon and J. Huang (eds), *Long distance anaphors*. Irvine, Calif.: Academic Press, 47-82,

Dayal,V.S. (1994). Binding facts in Hindi and the Scrambling Phenomenon. in M. Butt, T. H. King, and G. Ramchand, eds., *Theoretical Perspectives on Word Order in South Asian Languages*, CSLI Lecture Notes 50, CSLI, Stanford, California, 237–262.,

Hirst, G. (1981). Anaphora in Natural Language Understanding: A Survey, Number 119 in *Lecture notes in Computer Science*. Springer-Verlag,

Hobbs,J. R. (1976). Pronoun Resolution, *Research Report* 76-1, Department of Computer Science, City College, New York., August

——. (1977). 38 Examples of Elusive Antecedents from Published Texts, *Research Report* 77-2, Department of Computer Science, City College, City University of New York, August

——. (1978). Resolving pronoun references, *Lingua,* 44:311–338. Also in Readings in Natural Language Processing, B. Grosz, K. Sparck-Jones, and B. Webber, editors, pp. 339-352, Morgan Kaufmann Publishers, Los Altos, California,

Kachru,Y.B.,Kachru, B., and Bhatia,K., (1979). The notion "Subject", a note on Hindi- Urdu, In M. K. Verma (Ed.), *The Notion of Subject in South Asian languages*, 79–108. University of Wisconsion,

Mitkov, R., (2002). Anaphora Resolution, *Pearson Education*

Palomar, M., Ferrandez, A., Morenoy, L., Martinez-Barco, P., Peral, J., Saiz-Noeda, M., Mu˜noz, R., (2001). An Algorithm for Anaphora Resolution in Spanish Texts, *Computational Linguistic*s, Vol 27, No. 24, pp-545 - 567

Pandharipande,R., and. Kachru, Y., (1977).Relational Grammar, Ergativity, and Hindi-Urdu. *Lingua 41*: 217–238,

Prasad R., and Strube,M.,(2000). Discourse Salience and Pronoun Resolution in Hindi, *Penn Working Papers in Linguistics, Vol 6.3,* pp. 189-208,

——. (2003). Constraints on The Generation of Referring Expressions, With Special Reference To Hindi, *A PhD dissertation* University of Pennsylvania

Sinha,S.,(2002). ACorpus-based Account of Anaphor Resolution in Hindi, *MA thesis*, University of Lancaster, UK.

Sobha, L. and Patnaik,B.N., (2002). Vasisht" – An anaphora resolution system for Malayalam and Hindi, *Symposium on Translation Support Systems*, IIT Kanpur 15-17 March

Tüfekçi,P., and Kılıçaslan,Y., (2005) .A Computational Model for Resolving Pronominal Anaphora in Turkish Using Hobbs' Naïve Algorithm, *International Journal of Computational Intelligence* vol. 2 Number 2 pp-71-75

APPENDIX A:
Some abbreviations used in Figures 1-10

S1- sentence
NP – Noun phrase
PP – postpositional phrase
VP – verb phrase
postp- postposition
SUB- subject
DO – direct object
PO - positional object
DAT – dative case
ERG - ergative case
NOM – nominative case
ACC – accusative case
LOC – locational case
PRO- pronoun

APPENDIX B
Sample Sentences from Hobbs and their translations in Hindi

SE1	The castle in Camelot remained the residence$_i$ of the king$_j$ until 536 when he$_j$ moved it$_i$ to London
SH1	Kamelot kaa kilaa [raja ka nivas –sthan]$_i$ san 536 tak raha, jab

	tak ki vah_i $usko_j$ london nahin le gaya
SE2	$John_i$ shaved $himself_i$ $John_j$ shaved him_i $John_j$ shaved $John_i$.
SH2	$John_i$ ne $apni_{i/}khud_i$ ddadi banaai. $John_j$ ne $uski_i$ daadi banaai. $John_j$ ne $John_i$ kii daadi banaai.
SE3	$John_j$ saw him_i. $John_j$ saw a picture of him_i. $John_j$ saw a picture of him_i hanging in the post office. $John_j$ saw a picture of him_i was hanging in the post office. $John_i$ claimed that the picture of him_i hanging in the post office was a fraud.
SH3	$John_j$ ne $usko_i$ dekha $John_i$ ne $uski_j$ /$apnii_i$ tasveer dekhi. $John_i$ ne $uski_j$ /$apnii_i$ tasveer post office mein latki hui dekhi $John_i$ ne $uski_j$ /$apnii_i$ tasveer dekhi, jo post office mein latki hui thi. $John_i$ ne dava kia ki $uski_i$ /$apni_{*i}$ tasveer jo post office mein latki hui thii, nakli thii.
SE4	John said his mother would pay the man who shaved him_m. (hobbs algorithm associate him with “the man”)
SH4	$[John_i]_m$ ne kaha ki $[uski_i\ maa_j]_{*m}$ $[us\ aadmi]_k$ ko paise degi $jisne_k$ $uski_m$ hajaamat banai. Our algorithm associates “uski” with “ [uski maa]”
SE5	$John_i$ told Bill that he_i had been lucky.
SH5	$John_i$ ne kahaa ki veh_i khushkismat thaa.
SE6	John’s $mother_i$ told Bill she_i was angry. She_i wanted to know where he_j had been.
SH6	$John_j$ ki maa_i ne Bill se kaha ki veh_i naaraaz thi. Veh_i jananaa chaahti thii ki vo_i kahan thaa.