Resolving Pattern Ambiguity for English to Hindi Machine Translation Using WordNet

Niladri Chatterjee **Shailly Goyal** **Anjali Naithani**
Department of Mathematics
Indian Institute of Technology Delhi
Hauz Khas, New Delhi - 110 016, India
{niladri_iitd, shailly_goyal}@yahoo.com

Abstract

A common belief about natural language translation is that sentences of similar structure in the source language have translations that are similar in structure in the target language too. However, with respect to English to Hindi translation, this assumption does not hold well always. At least eleven different patterns can be found in the Hindi translation of English sentences in which the main verb is "have" or any of its declensions. This poses a serious problem for designing any English to Hindi translation system. Traditionally such variations are termed as "translation divergence". Typically a study of divergence considers some standard translation pattern for a given input sentence structure. A translation is said to be a *divergence* if it deviates from this standard pattern. However, this is not the case with the above-mentioned sentence structures. We term this ambiguity as "pattern ambiguity". In this ongoing work we propose a rule-based scheme to resolve the ambiguity using word senses given by WordNet.

1 Introduction

Natural language translation between any two languages almost inevitably suffers from ambiguities of various types, such as, lexical ambiguity, semantic ambiguity, syntactic ambiguity (Dorr *et al.* 99). Typically, all these ambiguities are related to deciphering the inherent meaning of the source language sentence. Normally these ambiguities can be resolved by considering the part-of-speech of the word concerned, or from other words of the sentence, or from the context of the sentence. Once the ambiguity is resolved, obtaining the correct translation in the target language becomes simpler.

However, with respect to English to Hindi translation a different type of ambiguity is observed (Goyal *et al.* 04). The problem here is not in understanding the sense of the sentence, rather, the difficulty is in deciding the correct structure of the Hindi translation. The following sentences and their Hindi translations illustrate this point:

Ram has a pen ∼ *ram* (Ram) *ke pass* (near to) *ek* (one) *kalam* (pen) *hai* (is).

Ram has fever ∼ *ram* (Ram) *ko* (to) *bukhaar* (fever) *hai* (is).

Although the structures of the above two English sentences are very similar, the structures of their Hindi translations are visibly very different. This creates a different type of ambiguity to the translator, which we term as "pattern ambiguity". Typically, such variations in translations are considered under the study of "translation divergence" (Dorr 93), (Gupta & Chatterjee 03). However, a subtle difference between pattern ambiguity and divergence can be observed easily. Study of divergence assumes some typical translation pattern (P, say) for a given source language sentence structure S. A translation divergence is said to occur if a source language sentence having the structure S assumes a pattern P1 that is different from P, upon translation into the target language. On the other hand, pattern ambiguity does not assume any standard translation pattern. Rather, corresponding to different input sentences of the same structure different translation patterns are observed, leading to "pattern ambiguity". Handling this ambiguity requires deep semantic analysis of source language sentences to find answers to:

(a) How serious is pattern ambiguity in English to Hindi translation?

(b) How to find ways to resolve this ambiguity while translating from English to Hindi?

With respect to (a) we notice that the presence of pattern ambiguity is most prominent in dealing with English verbs. In particular, we observe that as many as eleven different translation patterns may be obtained in the translation of English sentences where the main verb is "have", or some of its declensions.

To provide an answer to (b), we suggest a rule based scheme that takes into account the senses of the underlying English verbs, and other constituent words of a sentence to resolve the ambiguity.

In framing the above-mentioned rules we make significant use of WordNet 2.0[1]. In WordNet, English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept. In the proposed scheme semantic information about the constituents of the sentence under consideration is extracted using WordNet, and this information is then processed to resolve the ambiguity.

2 Translation Patterns of Different English Verbs to Hindi

One interesting aspect of English is that here a single verb is used to convey different senses. However, almost for each of these senses, a specific verb exists in Hindi. Table 1 shows some of the Hindi equivalents for the verb "`run`" when used in different senses.

Sentences	**Translation of Verb**
They run an N.G.O.	*chalaanaa*
The army runs from one end to another.	*failnaa*
The river ran into the sea.	*milnaa*
He runs for treasurer.	*khadaa honaa*
Wax runs in sun.	*galnaa*
We ran the ad three times.	*prakaashit karnaa*

Table 1: Different translations of "`run`"

The same observations have been made with respect to different English verbs, such as, `be`, `go`, `take`, `let`, `give`. All these English verbs can be used to convey different senses in different contexts. WordNet 2.0 provides different senses in which the above-mentioned verbs can be used. For example, the verb "`run`" has 41 senses, "`call`" has 28 senses, "`take`" has 42 senses. Since the use of the appropriate Hindi verb can be determined by identifying the sense in which the English verb is used, resolving pattern ambiguity for these verbs is relatively simple.

Most interesting observation in this regard can be made with respect to the English verb "`have`". Although the number of possible senses for "`have`" is relatively less (only 19, as per WordNet 2.0), we have obtained as many as 11 translation patterns for sentences where "`have`" (or its declensions) is the main verb of the sentence. Further, depending upon the situation, there are variations in the verb used, or the case-ending used, or sometimes even in the overall sentence structure. This makes pattern ambiguity to be a serious problem for English to Hindi translation while translating sentences of this type. Below we describe the different translation patterns that we observed in dealing with the English verb "`have`".

Translation Pattern P1: Here, genitive case ending (*kaa, kii, ke*) is used to convey the sense of the "`have`" verb. For example,

`The school has good name` ∼ *vidyaalay* (`school`) *kaa* (`of`) *achchhaa* (`good`) *naam* (`name`) *hai* (`is`).

Which of the genitive case endings (i.e. *kaa, kii, ke*) will be used in a given case depends upon the number and gender of the object. It is "*kaa*" if the object is masculine singular, "*kii*" if the object is feminine (irrespective of the number of the object), and "*ke*" for masculine plural.

Translation Pattern P2: In this pattern the object and its pre-modifying adjective in the English sentence are realized as the subject and subjective complement (SC), respectively, in the Hindi translation. The subject of English sentence is realized as possessive case of the subject of the Hindi translation. For example,

`Gita has beautiful hair`[2] ∼ *Gita* (`Gita`) *ke* (`of`) *baal* (`hair`) *sundar* (`beautiful`) *hain* (`are`).

Translation Pattern P3: Here a locative case ending "*ke paas*" is used instead of genitive postposition. For illustration, consider the following,

`Mohan has a book` ∼ *Mohan* (`Mohan`) *ke paas* (`near to`) *ek* (`a`) *kitaab* (`book`) *hai* (`is`).

Translation Pattern P4: In this pattern a postposition "*ko*" is used in the Hindi translation of the given sentence. For example,

`My uncle has asthama` ∼ *mere* (`my`) *chaachaa* (`uncle`) *ko* (`to`) *asthamaa* (`asthama`) *hai* (`is`).

[1] http://wordnet.princeton.edu/

[2] Note that according to P1 it should have been *Gita ke sundar baal hain.*

Translation Pattern P5: Here the postposition "*mein*" is used for conveying the sense of the verb "`have`". For example,
`This city has a museum` ∼ *iss* (`This`) *shahar* (`city`) *mein* (`in`) *ek* (`a`) *sangrahaalay* (`museum`) *hai* (`is`).

Translation Pattern P6: This translation pattern is similar to the pattern P5, except for the fact that postposition "*mein*" is replaced with another postposition "*par*". For example consider the following:
`The tiger has stripes` ∼ *baagh* (`tiger`) *par* (`on`) *dhaariyan* (`stripes`) *hain* (`are`).

Translation Pattern P7: Here, upon translation in Hindi, the object of the English sentence is realized as an SC which is an adjective. The following translations illustrate this pattern.
`She has grace` ∼ *wah* (`She`) *aakarshak* (`graceful`) *hai* (`is`).

Despite the obvious differences all the above-mentioned patterns have one common feature: the main verb of the Hindi sentence is "*hai*", which means "`to be`", or any of its declension (*hain, thaa, the, thii, thiin*). But patterns P8 and P9, given below, illustrate cases when some other verb is used as the main verb instead of "*hai*" (or its declension).

Translation Pattern P8: This pattern occurs if the main verb of the Hindi translation is obtained from the object of the English sentence. For illustration, consider the following example:
`Gita has regards for old men` ∼ *Gita* (`Gita`) *buzurgon* (`old men`) *kii* (`of`) *izzat* (`respect`) *kartii hai* (`does`).

The main verb of the Hindi sentence is *izzat karnaa*, which comes from the object "`regards`". In this respect one may note that Hindi verbs are often made of a noun followed by a commonly-used verb. The verb "*izzat karnaa*" is an example of this type.

Translation Pattern P9: This pattern is similar to the translation pattern P8, but here the verb is not obtained from the object. Rather, a completely new verb is introduced in the Hindi translation. For example,
`I had tea` ∼ *maine* (`I`) *chai* (`tea`) *pee* (`drank`).
But,
`I had rice` ∼ *maine* (`I`) *chaawal* (`rice`) *khaaye* (`ate`).

Evidently, the verb of the translated sentence is obtained from the "sense" in which the verb "`have`" is used in the English sentence.

Translation Pattern P10: In all the above cases the structure of the English sentences considered has been <SVO>. But, if the sentence has an additional component in the form of adjunct, then a variation in the translation may be noticed. For illustration, consider the two sentences:

(a) `Ram has two rupees`

(b) `Ram has two rupees in his pocket.`

While the translation of the first one is "*Ram ke pass do rupayaa hain*", the translation of the second one is "*Ram ki* (`Ram's`) *zeb* (`pocket`) *mein* (`in`) *do* (`two`) *rupay* (`Rupees`) *hain* (`are`)".

Under this pattern the following changes take place:

(a) The object and the adjunct (PP) in the English sentence are realized as the subject and the predicative adjunct, respectively, in the Hindi translation.

(b) The subject of the English sentence contributes as the possessive case to the predicative adjunct.

Translation Pattern P11: This pattern is observed if, along with the subject, verb and object, the sentence has an infinitive verb phrase. For example,
`My children had me buy the car` ∼ *mere* (`my`) *bachchon ne* (`children`) *mujhse* (`me`) *gaadi* (`car`) *kharidvaayai* (`buy`).

Further, we have found instances where the Hindi translation follows pattern pertaining to two or more classes. We term them as "mixed patterns". Due to page limitation we keep mixed patterns out of the present discussion.

Such a large variety of translation patterns pose great difficulty for any MT system, as the system needs to take a decision regarding the pattern that will be most suitable for a given input sentence. In this work we study whether a rule-based scheme can be developed to resolve this ambiguity.

3 How to Design Rules?

We first attempted to frame rules based on sentence structures. We observed that translation patterns P10 and P11 are associated with specific sentence structures. The sentence structure for rest of the patterns is <SVO>. The rules for P10 and P11 that we could frame on the basis of studying translations of sentences of these structures are given below:

Rule for P11: If the input sentence structure is such that the object of the verb (which is typically noun or pronoun) is followed by another verb, then Translation Pattern P11 is observed.
`I had Rama write a letter` ∼ *maine* (`I`) *rama* (`Rama`) *se* (`by`) *patr* (`letter`) *likhvaayaa* (`write`).

Rule for P10: If the given sentence structure is of the type <Subject Verb Object Adjunct (PP)>, and the PP satisfies the following two conditions, then the translation of the concerned sentence will have pattern P10:

(a) The head noun of PP is not animate.

(b) Head of the PP has a genitive pre-modifier that refers to the subject of the sentence.

For example, consider the following sentences:

1. `The table has dust on its surface` ∼ *mej ki* (`table's`) *satah* (`surface`) *par* (`on`) *dhool* (`dust`) *hai* (`is`).
2. `Sita has vermillion on her forehead` ∼ *Sita ke* (`Sita's`) *maathe* (`forehead`) *par* (`on`) *sindoor* (`vermillion`) *hai* (`is`).

However, the pattern *may* not be appropriate if one of the two conditions given above is not satisfied. Consider, for instance, the following translations:

1. `She has regards for her uncle` ∼ *wah* (`she`) *apne* (`her`) *chaachaa* (`uncle`) *ki izzat kartii hai* (`respects`). Note that the head noun of the sentence is animate. Thus it violates the condition (a) and one can observe that the translation pattern is P8, i.e. it is different from P10.
2. `Sita has degree from IIT` ∼ *Sita* (`Sita`) *ke paas* (`near to`) *IIT* (`IIT`) *ki* (`from`) *degree* (`degree`) *hai* (`is`). This sentence violates the condition (b) above and the translation pattern is P3.
3. `I have two dogs at home` ∼ *mere* (`my`) *ghar* (`home`) *par* (`at`) *do* (`two`) *kutte* (`dogs`) *hain* (`are`). Although this sentence also violates condition (b), still the translation pattern in P10.

Thus we notice that if the input sentence violates any of the above two conditions, then a variety of translation patterns may be obtained.

The above rules, however, exclude the majority of the sentences, as these are relevant to some special structures only. The majority of the patterns are related to sentences having the simple <SVO> structures. Hence we needed to investigate them further. In this respect the following is observed.

3.1 Inadequacy of Subject/Object

Our first attempt has been to design rules on the basis of the subject and/or object of the sentence. However, we found that the subject of the sentence alone is not sufficient to determine the translation pattern of the sentence. For illustration, all the sentences given in Table 2 have the same subject, yet they differ in their translation patterns.

English sentence	**Hindi Translation**	**Pattern**
`Mohan has a good brain`	*Mohan kaa dimaag achchhaa hai*	P1
`Mohan has a good pen`	*Mohan ke paas ek achchhii kalam hai*	P3
`Mohan has high fever`	*Mohan ko tej bukhaar hai*	P4
`Mohan had a sweet apple`	*Mohan ne meethaa seb khaayaa*	P9

Table 2: Translation patterns for same subject

In a similar vein, one can see that the translation pattern does not depend on the object too. The sentences given in Table 3 have the same object, yet their translation patterns are different.

These examples highlight the inadequacy of the subject/object in determining the translation pattern. In the next step we considered the senses of the nouns used as subject/object as given in WordNet 2.0. We have been able to frame a few rules in this way. For illustration:

English sentence	Hindi Translation	Pattern
`Sita has flowers`	*Sita ke paas phool hain*	P3
`The tree has flowers`	*ped par phool hain*	P6
`The vase has flowers`	*phooldaan mein phool hain*	P5
`Meera has flowers in her home`	*Meera ke ghar mein phool hain*	P10

Table 3: Translation patterns for same object

Rule (a) If the object of the given sentence is *body part* and object has a pre-modifier adjective that is not a quantifier, then the translation of that sentence will have pattern P2. For example, `Meera has swollen fingers` ∼ *Meera* (`Meera's`) *kii anguliyaan* (`fingers`) *soozii hui* (`swollen`) *hain* (`are`).

But, in above case if the pre-modifier of object is absent, or it is a quantifier, then the translation pattern P1 is observed. For illustration:

`The elephant has a trunk` ∼ *haathi kii* (`Elephant's`) *ek* (`a`) *soond* (`trunk`) *hai* (`is`).

Obviously, obtaining rules, their exceptions etc. in this way is not practicable. Further, it is very difficult to take care of all the possible cases in this way. Hence in the next stage we attempted to frame rules on the basis of the senses of the verb "`have`" itself.

3.2 Rules Based on Senses of "Have"

WordNet 2.0 has been used to decide upon the senses of the "`have`" verb. Our observations in this regard are as follows.

(a) Use of the verb "`have`" to convey senses numbered 5 (cause to move), 10 (be confronted with), 11 (experience), 13 (cause to do) and 19 (have sex with) is very rare.

(b) Of the remaining fourteen senses, identification of translation patterns for eight senses (viz., 6, 8, 9, 12, 14, 15, 17 and 18) can be done using their senses, as in all these cases only a single translation pattern can be observed (which in some cases is a mixed pattern!).

(c) For sense numbers 1, 2, 3, 4, 7 and 16 more than one translation pattern is observed. Hence, in these cases the sense of "`have`" is not sufficient, and finer rules are required to determine the possible translation pattern of the given sentence.

Table 4 summarizes our findings in this regard. This observation was made on the basis of our manual analysis of about 6000 sentences with "`have`" as the main verb. We first worked on 2000 sentences, and corroborated our findings on the basis of the remaining. All the patterns obtained so far are given in Table 4. However, it is too early to claim that no other pattern exists in some of the cases. Further studies are required in this regard.

The above observation suggests that even the sense of the verb is not enough to resolve the pattern ambiguity. For further investigation we took the help of Lexicographer files of WordNet 2.0. The lexicographer file information helps one in identifying the selectional restriction (Allen 95) of subject's/ object's semantics of a sentence.

3.3 Rules Based on Lexicographer Files

Lexicographer files in WordNet 2.0 are the files containing all the synonyms logically grouped on the basis of syntactic category. For example, the file *noun.act* contains nouns that describe any act or action, *noun.animal* is a file containing nouns that are animals. According to WordNet, noun has 26 different logical groupings. Corresponding to these groupings there are 26 lexicographer files. Pronouns can be taken care of under these categories primarily as *noun.person*, or some other categories depending upon the context. We used these lexicographer files for designing rules for translation patterns. Further, there can be imperative sentences where the subject "`you`" is silent (e.g. `Have this book.`). Thus we have 27 possibilities for subjects, and 26 possibilities for objects for dealing with word sense disambiguation of "`have`".

On studying subject and object of our database sentences, a 27×26 matrix has been constructed. The matrix suggests the translation patterns obtained in different combination of subject and object. However, in our example base we found no sentences in which the subjects are one of *noun.motive*, *noun.phenomenon*, *noun.process*, *noun.feeling*, *noun.possession* and *noun.relation*. Similarly, there are no sentences in which the ob-

Sense Number	Definition (As given by WordNet 2.0)	Translation Pattern	Example sentence	Translated sentence
1	have or possess, either in a concrete or abstract sense	P1	Rita has two daughters.	*Rita ki do betiyaan hain.*
		P3	She has a degree from IIT.	*us ke paas IIT kii degree hai.*
2	have as a feature	P1	This dog has three legs.	*iss kutte kii teen taangen hain.*
		P2	She has beautiful eyes	*uskii aankhen sundar hain.*
		P5	This car has an airbag.	*iss gaadi mein ek airbag hai.*
		P6	The tree has flowers.	*ped par phool hain.*
		Mixed P1 and P8	Ravi has a good grasp of subject.	*Ravi kii vishay par achchhii pakad hai.*
3	of mental or physical states or experiences	P1	Ram has many dreams.	*Ram ke bahut sapnay hain.*
		P2	Mita has an idea.	*Mita ke paas ek upaay hai.*
		P3	Ram has sympathy for the poor.	*Ram ko gariibon ki liye shaanubhutti hai.*
		P8	She has regards for her father.	*vah apne pitaa kii izzat kartii hai.*
		P9	She had a difficult time.	*usne mushkil samay bitaayaa.*
4	have ownership or possession of	P1	Hemu has three houses.	*Hemu ke teen ghar hain.*
		P3	Mohan has a car.	*Mohan ke paas ek gaadii hai.*
6	serve oneself to, or consume regularly	P9 ("*khaanaa*" or "*peenaa*")	I had an apple.	*maine ek seb khaayaa.*
7	have a personal or business relationship with someone	P1	He has an assistant.	*us kaa ek sahaayak hai.*
		P3	This professor has a research scholar.	*iss professor ke paas ek gaveshi hai.*
8	organize or be responsible for	P1	John has a meeting.	*John kii ek meeting hai.*
9	have left	P3	Meera has two years left.	*Meera ke paas do saal bache hain.*
12	suffer from; be ill with	P4	Paul has fever.	*Paul ko bukhaar hai.*
14	receive willingly something given or offered	P9 ("*lenaa*" or "*sweekaar karnaa*")	Please have this gift.	*kripayaa yeh uphaar lein.*
15	get something; come into possession of	P9 ("*milnaa*" or "*prapt honaa*")	I have a letter from a friend.	*mujhe ek mitr kaa patr milaa.*
16	undergo (as of injuries and illnesses)	Mixed P1 and P8	Rama had a fracture	*Ram kii haddii tootii.*
		Mixed P4 and P8	His father had a heart attack.	*uske pitaa ko hradayaaghaat huaa.*
17	achieve a point or goal	P9 ("*banaanaa*")	Sachin had a century.	*Sachin ne shatak banaayaa.*
18	give birth (to a newborn)	P9 ("*janam denaa*")	My wife had a baby boy yesterday.	*kal meri patnii ne ladkee ko janam diyaa.*

Table 4: Rules for translation patterns for different senses of "`have`"

jects are *noun.motive* or *noun.relation*. Hence we discarded these columns and rows from the matrix. Therefore, the final matrix has $21 \times 24 = 504$ cells. A thorough scrutiny of the matrix reveals the following:

Case 1. Out of the 504 cells, 297 cells are empty i.e. no example has been found for corresponding combinations of subject and object. For example, when the subject is *noun.attribute* and object is *noun.animal*, then the cell is empty implying that our database contains no valid English sentence in which the above combination is observed. For these 297 situations no translation rules need to be formed.

Case 2. The simplest case is when there is only one entry in a cell. There are 85 (out of 504) cells which have only one entry. This implies that for these 85 combinations of subject and object, pattern ambiguity can be resolved directly. Some of these combinations are given in Table 5.

Subject Sense	**Object Sense**	**Pattern**
noun.act	noun.state	P1
noun.act	noun.substance	P5
noun.animal	noun.cognition	P2
noun.animal	noun.substance	P6
noun.group	noun.quantity	P1
noun.group	noun.substance	P3
noun.plant	noun.phenomenon	P8
noun.plant	noun.state	P5

Table 5: Singly occupied cells

Case 3. We further observe that for some columns and rows there are only two or three patterns occurring, i.e. for a given subject there are only two or three possible translation patterns, irrespective of the object used. For example, if the subject is *noun.act*, then the patterns observed are P1 or P5. Similarly, for some object senses only a limited number of patterns are possible. For example, if object is *noun.shape*, then possible translation patterns are P5 or P6.

The advantage of the above observation is that to resolve pattern ambiguity the system need not explore all the 11 possibilities. Rather, it may furnish two or three translations of the sentence and obtain user feedback. There is also scope of learning by the MT system, as it handles more cases of a particular type.

Case 4. There exist some subject-object combinations with only two or three entries. For instance,

1. If the subject is *noun.artifact*, and object is *noun.communication*, then the patterns observed are P5 or P6.
2. If the subject is *noun.act*, and object is *noun.cognition*, then possible translation patterns are P1 or P5.
3. If the subject is *noun.group*, and object is *noun.cognition*, the translation pattern is one of P3 or P5.

As in Case 3, here too the pattern ambiguity can be resolved through user feedback.

Case 5. However, there are 15 cells that are very dense, i.e. for these combinations of subject and object, the number of possible translation patterns is quite large. Table 6 shows these subject/object combinations, the possible translation patterns, and the number of observations. Pattern ambiguity cannot be resolved for these sentences, since for each of the 15 cases a large number translation patterns are possible.

The question therefore arises whether pattern ambiguity in translating English sentences with "`have`" as its main verb is completely resolvable. We tried to capitalize on all possible sentential information, yet we have not been able to find a foolproof solution. So far, we could resolve pattern ambiguity for about 75% of cases, out of about 4000 sentences (these are the sentences on which the rules designed have been testified (See Section 3.2)) using the above scheme. We feel that the only way it may be resolvable is by analyzing the context. But creating a large database containing appropriate context information as well as having "`have`" sentences is not an easy task. Currently we are looking into this aspect.

4 Concluding Remarks

This paper first defines the term "pattern ambiguity" that is observed in translation from English to Hindi. It has been observed that this ambiguity can occur during the translation of English sentences. Although the ambiguity exists with respect to translation of different English verbs, this is particularly prominent and not yet fully resolvable for sentences whose main verb is "`have`" or its declensions.

Subject	Object	Pattern Observed
noun.artifact	noun.artifact	P1 - 67, P2 - 35, P5 - 36, P6 - 45
noun.group	noun.act	P1 - 34, P2 - 9, P4 - 8, P5 - 18
noun.group	noun.attribute	P1 - 18, P2 - 7, P3 - 8, P5 - 17
noun.person	noun.act	P1 - 51, P2 - 34, P3 - 22, P4 - 8, P6 - 6, P8 - 16, P9 - 25
noun.person	noun.artifact	P1 - 25, P2 - 10, P3 - 35, P5 - 10, P10 - 24
noun.person	noun.attribute	P1 - 56, P2 - 34, P3 - 12, P4 - 4, P5 - 56, P6 - 23, P7 - 59, P8 - 13, P10 - 6,
noun.person	noun.body	P1 - 15, P2 - 6, P3 - 6, P5 - 10, P8 - 14, P9 - 7
noun.person	noun.cognition	P1 - 35, P2 - 24, P3 - 35, P4 - 23, P5 - 25, P7 - 12, P9 - 8
noun.person	noun.communication	P1 - 24, P2 - 34, P3 - 29, P4 - 4, P5 - 15
noun.person	noun.feeling	P1 - 16, P3 - 6, P4 - 35, P5 - 25, P7 - 27
noun.location	noun.group	P1 - 7, P2 - 5, P5 - 24, P6 - 7
noun.person	noun.person	P1 - 17, P2 - 3, P3 - 4, P9 - 2
noun.person	noun.possession	P1 - 40, P3 - 16, P8 - 16, P9 -6, P10 - 13
noun.person	noun.state	P1 - 24, P2 - 35, P3 - 18, P4 - 16, P5 - 26, P6 - 8, P7 - 17, P8-25, P9 - 16
noun.person	noun.time	P1 - 7, P2 - 7, P3 - 13, P8 - 13

Table 6: Densely occupied cells

The primary reason behind this ambiguity is that Hindi does not have a verb that is equivalent in sense to the English "`have`" verb. However, not only Hindi, many other languages (e.g. Bengali, Hausa[3]) do not have any possessive verb. We hope that this study will be helpful for studying translation patterns into such languages as well.

"Pattern ambiguity" is a serious problem for machine translation. It is more serious than "divergence" as it is possible to identify divergence by noting the structural changes in the source language and target language sentence (Gupta & Chatterjee 03). Also, it is more serious than typical WSD problem (Ide & Veronis 98), as WSD is not concerned with the translation structure. We feel that statistical techniques need to be applied to determine the translation pattern for a given input, when the subject and object senses lead to several possible ways of translation. However, this needs a large volume of appropriate database that is not available at present.

In this work we have used verb senses and subject-object senses separately. We feel that the problem may be dealt with at a more granular level by considering these two senses together for a given input sentence. Presently we are focusing our investigations to that direction.

References

(Allen 95) James Allen. *Natural Language Understanding.* Benjamin Cummings, California, 2nd edition, 1995.

(Dorr 93) Bonnie J. Dorr. *Machine Translation: A View from the Lexicon.* MIT Press, Cambridge, MA, 1993.

(Dorr *et al.* 99) Bonnie J. Dorr, Pamela A Jordan, and John W Benoit. A survey of current research in machine translation. In M. Zelkowitz, editor, *Advances in Computers*, volume 49, pages 1–68. Academic Press, London, 1999.

(Goyal *et al.* 04) Shailly Goyal, Deepa Gupta, and Niladri Chatterjee. A study of Hindi translation pattern for English sentences with "have" as the main verb. In *ISTRANS-2004*, pages 46–51, New Delhi, India, 2004.

(Gupta & Chatterjee 03) Deepa Gupta and Niladri Chatterjee. Identification of divergence for English to Hindi EBMT. In *MT Summit IX*, pages 141 – 148, New Orleans, USA, 2003.

(Ide & Veronis 98) Nancy Ide and Jean Veronis. Word sense disambiguation: The state of the art. *Computational Linguistics*, 24(1):1 – 40, 1998.

[3] http://www.humnet.ucla.edu/humnet/aflang/Hausa/Hausa_online_grammar/grammar_frame.html