Chapter 7
DISCUSSIONS AND CONCLUSIONS

The present work terminates with the design of LFG grammar for Hindi and development of a parser. The output of the parser is c-structures and corresponding f-structures. An f-structure is a pseudo-semantic representation of a sentence in the language. This form is language independent and thus can be used for mapping the information in one language to the notations in another language.

7.1 Discussions : Morphology and Lexicon

Since Hindi is a non-configurational language, rules have been devised for almost all possible combinations of noun phrases and adverb phrases in a simple sentence to meet this characteristic of the language. Different grammatical functions in a sentence are generally distinguished by the case markers following nouns. But, the case markers do not have one-to-one mapping with the grammatical functions. That is, some case markers may denote more than one grammatical function, and one grammatical function may be denoted by more than one case marker. To resolve this ambiguity, case markers having one-to-one mapping are separated from other. The case disambiguation is achieved by the rule, by providing the information with the verb regarding the case markers the verb expects with different grammatical functions, and regarding the semantic information a grammatical function requires in a sentence. For example, the case marker 'ko' may denote either direct object or indirect object. This is solved by rule combinations. In a rule, if both the grammatical functions are present, it will denote indirect object, otherwise it will denote direct object. This is shown in the following sentences:

(i) \textit{raam mohan ko pectataa hai.}
(ii) \textit{raam ne mohan ko pustak dee.}

In the following sentences, the second ambiguity type is resolved by supplying the appropriate case information regarding the grammatical functions along with the verb.
For example, the case marker 'se' denotes location as well as instrumentality, but its semantic interpretation varies according to the verb type.

(i) *raam ghar se nikala*.  
(ii) *rasoiye ne chhaakoo se sabjee kaaTee*.  

In the following sentences, the third ambiguity type is resolved by supplying the semantic information along with the verb for each grammatical function it requires. For example, both the subject and object are denoted by the same case marker 'ϕ'. But the verb requires that the noun of the noun phrase denoting subject must have the property 'human=+'. This property is tested by LFG module.

(i) *raakhee sabjee kaaTatee hai*.  
(ii) *sabjee raakhee kaaTatee hai*.  

The sophistication of morphological analysis finds justification in the fact that all the variants of a word may not be part of lexicon-format. Therefore, morphological module containing word attributes along with corresponding f-equations must back-up the system.

7.2 Conclusions: Limitations and Future Scope

Since the f-structure is not the ultimate semantic representation of knowledge, one may need to translate it into other semantic representations such as predicate calculus, semantic nets, conceptual graphs etc. The work therefore is extendable.

The parser designed here is language and grammar independent, and thus can work for other languages and grammatical frameworks as well. The modifications needed are in terms of providing required grammar and the lexicon, and in the representation of the constraints. This is not a major modification. The parser is also independent of the size of the grammar. It allows one to change the form or the number of rules in the grammar. Thus, the work carried out takes care of the portability problem which is crucial in the domain of NLP.
The functional structures are limited to sentence validation only, as the c-structures themselves are sufficient to account for the structural description of sentences. However, since the LFG formalism ignores the conventional TG kind of movement rules, the c-structures obtained as parses have to be combined with the advantages from the X-bar analysis to yield a unified mechanism applicable to cross-linguistic machine translation.