Government and Binding Theory for Hindi Language

Lalita Kumari
Department of Computer Science, National Institute of Technology, Agartala, India

Radhey Shyam
National Informatics Center, Tripura State Center, Agartala, India

Swapan Debbarma, Nirmalya Kar, Smita Das
Department of Computer Science, National Institute of Technology, Agartala, India

ABSTRACT
Government and Binding theory is used to analyze phase structure rules in Natural Language Processing. It is used to filter out grammatically incorrect sentence. Government and Biding (GB) theory is useful and well applicable in English language. This paper shows way to apply GB theory on Hindi Language. For Applying on Hindi, GB theory is modified little bit, but its flavor is intact. Using this, all components and attributes of Government and Binding can be easily described for Hindi language**.**

General Terms
Natural Language Processing, Government and Binding, Grammar and Parser.

Keywords
Lexicon, D-structure, F-structure, S-structure. Logical form, C-command, M-command..

1. INTRODUCTION
Government and binding is used in Natural Language Processing for syntax transformation. It has great expressive power and is easy to understand. GB theory helps to design computational model in natural language processing. It is developed principally by Noam Chomsky in the 1980s [1, 2, 3]. This theory is a subsequent modification and revision of Chomsky's earlier theories. It was modified latter in 1950 by The Minimalist Program and several subsequent research papers. The latest revision was Three Factors in Language Design (2005) [4, 5, 6, 7, 8].

Government and Binding (GB) theory of Chomsky introduced in 1981 and improved in 1986. Government and Binding approach to linguistic theory. GB theory is based on Extended Standard Theory in transformational grammar. GB theory consist set of theories that interact Government and Binding being two of them. It gives emphasis on principles rather than rules. Each theory implies principles of rules and representations that are a subsystem of UG. GB theory operates always on syntactic structure, whatever level of language (d-structure, s-structure or LF) are they affecting. The interactions between the theories becomes complex. If we keep principle simple than interactions between simple principles may lead to complex properties. But we can understand why language is complex but we can easily learn.

A sentence in the theory of Government and Binding is represented at following four levels: (1) D-structure, (2) S-structure, (3) Phonetic Structure and (4) Logical Form. We can represent it in the simplest form as follow: D-structure (DS) captures the argument structure of lexical categories: Verb, Noun, Pre (or Postposition and Adjective, present in the sentence. Logical Form (LF) comes nearest to the meaning representation of the sentence. in terms of quantifier scoping, anaphoric reference indexing etc. Quite often PS and SS are the same, but they could also be different as in "I wanna go" (PS) and "I want to go" (SS).

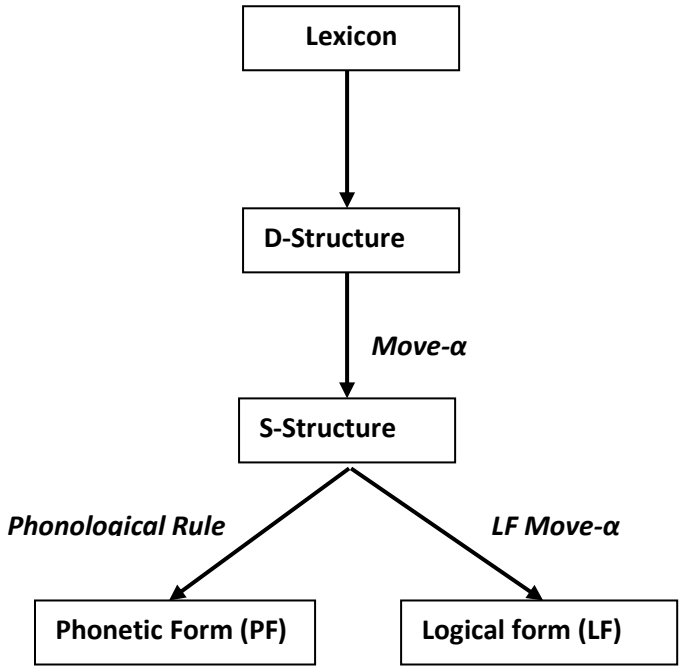

Figure 1 Representation of D-structure and S-structure.

Figure 1 shows D-structure and S-structure representation of lexicon. Facts about them are as follows:

- D-structure combined all lexical items.
- D-structure is mapped into S-structure that syntactically represents surface order of sentence.
- S-structure is factorized into Logical Form and Phonological Form.
- Phonological Form (PF) is directly concerned with phonology i.e. related to sounds of items .
- Logical Form is the interface with the Semantics. Predication relationships and the scope of operators and quantifiers of various kinds are represented explicitly in the phrase structure at Logical Form.
- Every word is of lexical category which is called head. Examples are noun, verb, adjective, preposition

Meaningful grouping of same category of lexical words is called Phrases (e.g. AP, NP, PP, VP, etc.).

2. GOVERNMENT AND BINDING FOR HINDI

2.1 X-Bar Theory

X-bar theory, developed by Chomsky is mostly applicable in English language. Little modification in algorithm is needed to adopt in other language. Similarly, for Hindi language modified algorithm is presented here.

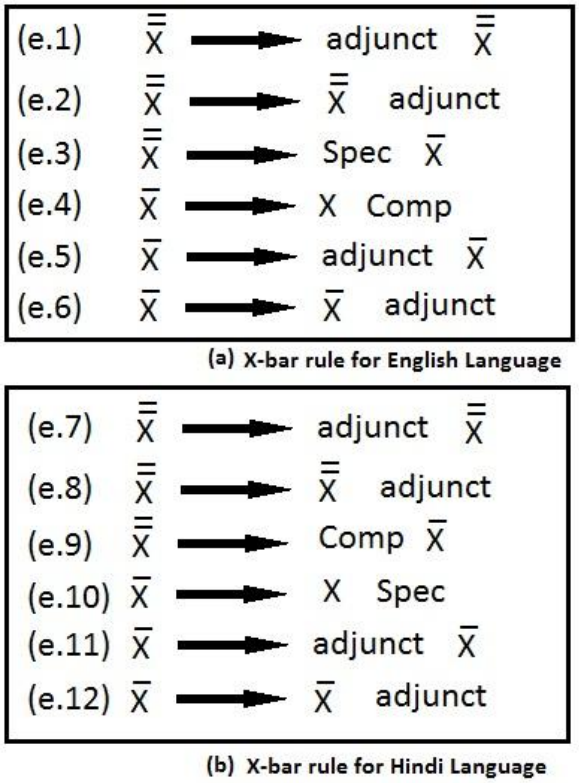

Figure 2 Rules in X-bar theory for English and Hindi language shown in figure a. and b. respectively

Figure.2.a shows original algorithm given by Chomsky for English language. And figure.2.b. shows proposed algorithm for X-bar theory, which can be implemented for Hindi Language. Figure.3 shows example sentence satisfying tree structure of X-bar theory presented in figure 2.b. shown in figure a. and b. respectively

Comparison of equations from figure 2.a and figure 2.b shows that position of compliment and specifier are interchanged and it is reflected in equation e.9 and e.10. (reflected in figure 3).

Here in figure 3. Noun राजा act as X0, and its bar, X-bar is represented as N-bar in tree structure.

Here राजा is the head, मिथिला के is the specifier, and जनक is the complement in the NP.

The head is called the zero projection. The top node of tree structure is called the maximal projection of head (X). All other projections are called intermediate projections. The sibling of X is called the complements of the head, and the sibling of X' is called specifier.

2.2 C-Command

In Government and binding framework, both government and binding notions are defined upon c-commands. C-command represents binary relation between nodes of a tree structure. Chomsky defined c- command as:

A c-commands B iff :

- Neither A nor B dominate each other.
- First Branching node that dominate A also dominate B.

In a sentence, subject c-commands object. If there are two objects in sentence, one object c-command another object

Figure 4 shows example sentence structure, illustrating c-command and m-commands. C-commands in figure4 are:

- (NP) "किताब " c-commands (V) "खरीदी ".
- (V) "खरीदी"A c-commands (NP) "किताब".
- (Adj) "नयी" c-commands (NP) "किताब" and (V) "खरीदी"A.
- (Sub) "मनीष" c-commands (NP) "किताब", (V) "खरीदी", and . (Adj) "नयी "

Neither (NP) "किताब" nor (V) "खरीदी" c-commands (Adj) "नयी ".

None of (NP) "किताब", (V) "खरीदी", and (Adj) "नयी " commands (Sub) "मनीष ".

The c-command is used to explain the working of anaphora.

2.3 M-Command

A m-commands B if

- Neither A nor B dominate each other.
- The first maximal projection of A dominates B.

First maximal projection in tree structure is its grandparents i.e. two level up in tree. 2nd maximal projection in tree structure is 3 levels up etc.

First maximal projection in figure 4, are as follow:

First maximal projection of "मनीष" is root node of tree i.e.

VP.

First maximal projection of "किताब" and "खरीदी" are its grandparents VP.

First maximal projection of "नयी", NP, and V is VP (parent of Adj.)

First m-command in figure 2, are as follow:

"मनीष" NE m-commands "किताब", "खरीदी", and "नयी"

- "किताब" m-commands "खरीदी".
- "खरीदी"I m-commands "किताब".

- "नयी" m-commands "किताब", and "खरीदी" .

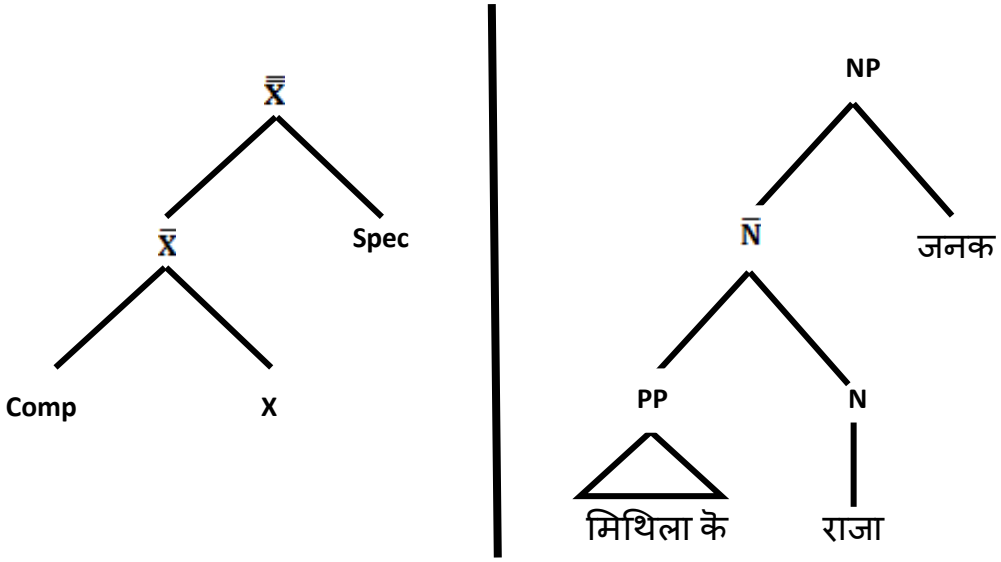

Figure 3 X-bar tree representation for Hindi NP

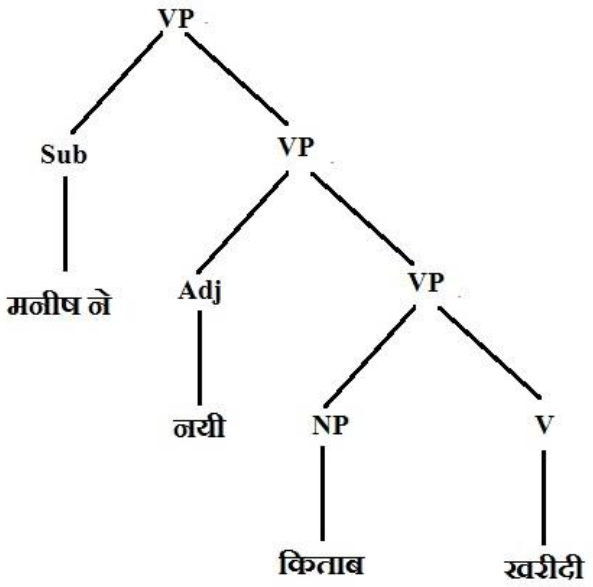

Figure 4 Tree Structure representing Representation of c-command and m-commands

2.4 Difference between m-command and c-command:

Difference between c-command and m-command is distinguished by definition of first branching node and first maximal projection respectively For determining first branching node we go up until we not get the node which has more than one children. But for determining first maximal projection we go only two level up without checking whether it has more than one children or not. Table 1 distinguishes these two commands for tree representation shown in Figure

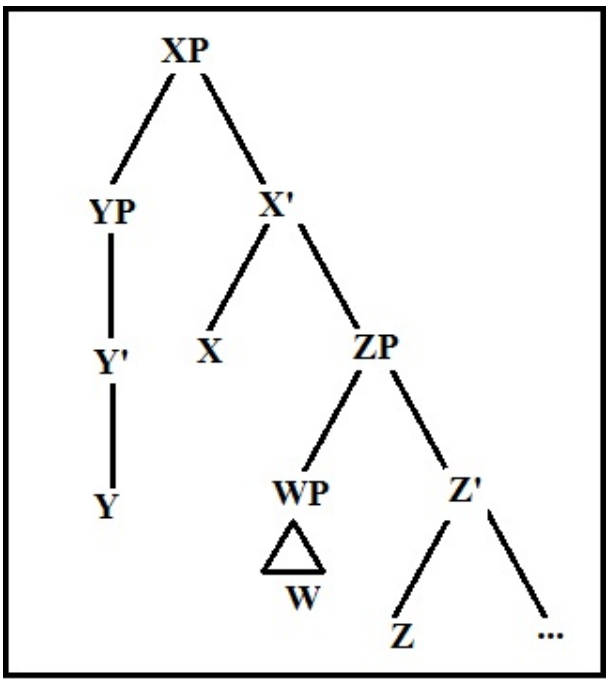

Figure 5 c-command and m-commands

Table 1 Representation of First maximal projection and First branching node.

Node	First Maximal Projection	First Branching Node
Y	YP	XP
Y'	XP	XP
X	XP	X'
WP	X'	ZP
Z	ZP	Z'

2.5 Government:

A governs B if and only if

- A is a governor and
- A m-commands B and
- no barrier intervenes between A and B.

Governor is head of a lexical categories. A barrier is any node Z such that Z is a potential governor for B and Z c-commands B and Z does not c-command A.

In figure 4, head is "खरीदी". Therefore "खरीदी" Governs "किताब" only. Whereas "खरीदी"I m-commands both "नयी" and "किताब". Because there is barrier intervenes between "खरीदी" and "नयी"

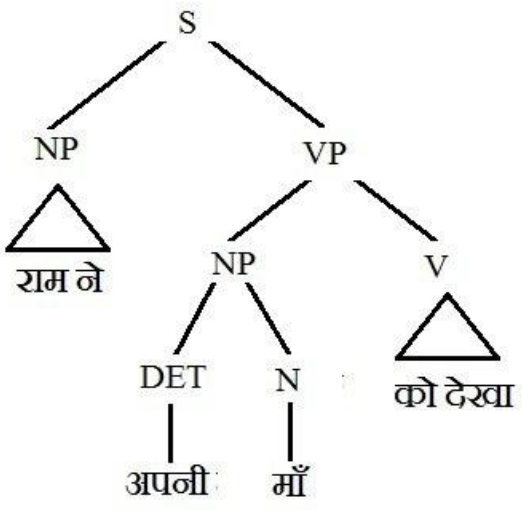

Figure 6 Representation of binding

2.6 Binding

An element α binds an element β if and only if:

- α c-commands β,
- α and β co-indexed

Here co-index means α and β should represent same thing. Figure 6 represent binding. Here in figure 6 (N) "राम " c-command (DET) "अपनी" . Both "राम" and "अपनी" are co-indexed. i.e. representing same thing (राम). Therefore (N) "राम" binds (DET) अपनी.

3. CONCLUSION

Here we presented concepts of Government and binding. This gives clear view of Government and Binding theory with example from Hindi sentences. Government and Binding theory focuses on principle rather than rule so its principle seems so complex but when we apply this can be easily understood by doing little modification for Hindi language. Only we have to change in some given rule for GB eg. X-bar theory (as shown in this paper). Using this concept of government and binding can be implemented in Indian languages.

4. REFERENCES

[1] Chomsky, Noam (1981). Lectures on Government and Binding. Berlin:

[2] Chomsky, Noam (1982). Some Concepts and Consequences of the Theory of Government and Binding.

[3] Chomsky, N. and Lasnik, H. (1993) Principles and Parameters Theory, in Syntax: An International Handbook of Contemporary Research, Berlin: de Gruyter.

[4] Haegeman, Liliane (1991). Introduction to Government and Binding Theory. Oxford: Blackwell Publishers.

[5] Jackendoff, Ray (1977). ¯X Syntax: A Study of Phrase Structure. Cambridge Massachusetts: MIT Press.

[6] Ouhalla, Jamal (1994). Introducing Transformational Grammar. New York: Oxford University Press.

[7] Sag, Ivan A., Thomas Wasow and Emily M. Bender (2003). Syntactic Theory. A Formal Introduction. Palo Alto: CSLI Publications.

[8] Kayne, Richard. 1981. Unambiguous paths. In Levels of syntactic representation, ed. by Robert May and Jan Koster, 143–183. Dordrecht: Reidel.

[9] Laughlin, Simon, and Terrence J. Sejnowski. 2003. Communication in neuronal networks. Science 301 (26 September 2003): 1870–1874.

[10] Haegeman, L. (1994). *Introduction to Government and Binding* (2nd Ed.). Oxford: Blackwell Publishers Ltd.

[11] Cowper, E. A. (1992). A concise introduction to syntactic theory: The Government and Binding approach. Chicago: The University of Chicago Press.

[12] Black, C. A. (1999). A step-by-step introduction to the Government and Binding theory of Syntax. Summer Institute of Linguistics.

[13] Pollock,Jean-Yves. 1989. Verb movement, UG, and the structure of IP. Linguistic Inquiry 20:365-424.

[14] Diesing, Molly. 1990. Verb movement and the subject position in Yiddish. Natural Language and Linguistic Theory 8:41-79.

[15] 1995. Bare phrase structure. Government and Binding theory and the Minimalist Program, edited by Gert Webelhuth, 71-132. Oxford: Blackwell.

[16] 1982. Some concepts and consequences of the theory of Government and Binding. Chicago: MIT Press.

1997. Introduction to Government and Binding theory: Semantic roles and Case theory. Notes on Linguistics 77:7-19.