Developing a UNL to Hindi Converter from a English Converter

A Project Report

Submitted by

Somani Abhishek

in partial fulfilment of the requirements
for the award of the degree of

Bachelor of Technology

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MADRAS, CHENNAI

MAY 2014

CERTIFICATE

This is to certify that the report titled **Developing a UNL to Hindi Converter from a English Converter**, submitted by **Somani Abhishek**, to the Indian Institute of Technology, Madras, for the award of the degree of **Bachelor of Technology**, is a bona fide record of the work done by him under our supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Mr. Ajit Narayanan

CEO

Inventions Labs

IIT Madras Research Park, 600036

Prof. S Umesh

Professor

Dept. of Electrical Engineering

IIT Madras, 600036

Place: Chennai

Date: 13th May 2014

ACKNWOLEDGEMENTS

I would like to express my heartfelt gratitude to my guide, Mr. Ajit Narayanan for his invaluable support and guidance throughout my project duration. I would also like to thank Mr. Mohan Ravi Chandran for being extremely helpful. During the project his patience and guidance helped understand concepts and provided greater insight into the topic. I thank Prof. S Umesh for all the support he has extended. I also take this opportunity to thank all the professors at IIT who have taught me or helped me otherwise.

ABSTRACT

In this report we start by discussing a Universal Networking language (UNL) which represents the sentence in the form of semantic relations like a hyper graph which is language independent. We later discuss the X bar theory which gives structure to the nodes of the UNL. We later formulate an algorithm which joins the elementary trees and movements in the trees to form a sentence form the UNL. We also discuss the changes required in the algorithm to shift from English language to Hindi Language

Contents

INTRODUCTION

Avaz has been successful in the area of Medical assistive technologies for Autism. Avaz is a picture based communication application. This applications helps Challenged children reinforce and model speech. Its builds the language. It makes children use motor planning for language development. It has been a great success but it is difficult to develop fairly long complex sentences using this application. To solve this issue, a new interface is being developed. This interface is based on the logical relation between two words. For example, verb would be connected to the agent using a relation arrow by asking you a question, "Who did it?". This way, the graph with key words as nodes and relations between them as edges would be built by the user. This graph would be converted to Universal Networking language (UNL) which is basically a list of relations in the sentence. This UNL would be used to develop a sentence in any language. I had a working UNL to English converter. Every language has a different structure. My project was to develop a UNL to Hindi converter. The main idea was to make the converter as language independent as possible.

The project report consists of the following part. The chapter on Theory presents a brief background on the theory of Universal networking language. X bar Theory gives an in-depth account of structure of language. The chapter UNL to English explains the implementation of the algorithm using the X bar theory. Changes from English to Hindi discusses the modifications in the algorithms and structure in of X bar theory due to change in language

The end of the project report has in-depth discussion regarding future development of this application

UNIVERSAL NETWORKING LANGUAGE

In this chapter, we look briefly at the back ground which is necessary to understand the forth coming chapters.

The UNL is an effort to achieve a simple basis for representing the most central aspects of information and meaning in a human-language-independent form. As a knowledge representation language, the UNL aims at coding, storing, disseminating and retrieving information independently of the original language in which it was expressed. In this sense, the UNL seeks to provide the tools for overcoming the language barrier in a systematic way. Its primary objective is to serve as an infrastructure for handling knowledge rather than individual languages.

In the UNL approach, there are two basic different movements: UNL-ization and NL-ization. UNL-ization is the process of representing/mapping/analysing the information conveyed by natural language utterances into UNL; NL-ization, conversely, is the process of realizing/manifesting/generating a natural language document out of a UNL graph. These processes are completely independent

Currently, the main goal of the UNL-ization process has been to map the information that is verbally elicited in the surface structure of written texts into a language-independent and machine-tractable database. This means that the UNL representation has not been committed to replicate the lexical and the syntactic choices of the original, but focuses in representing, in a non-ambiguous format, one of its possible readings, preferably the most conventional one. In this sense, the UNL representation has been an interpretation rather than a translation of a given text.

A computer in future needs a capability to make knowledge processing. Knowledge processing means a computer takes over thought and judgment of humans using knowledge of humans. It is necessary to make a processing based on contents. Computers need to have knowledge for knowledge processing. It is necessary for computers to have a language to have knowledge like human. It is also necessary to have a language to process contents like

human. The UNL is a language for computers to do so. The UNL can express knowledge like a natural language. The UNL can express contents like a natural language.

The main goal of the UNL is to represent, in a machine-tractable format, the information conveyed by natural language documents. In the UNL framework, this information is represented by a semantic network, i.e., a network which represents semantic relations between concepts. This semantic network, or UNL graph, is made of three different types of discrete semantic entities: Universal Words, Universal Relations and Universal Attributes. Universal Words, or simply UW's, are the nodes in the semantic network; Universal Relations are arcs linking UW's; and Universal Attributes are used to instantiate UW's.

1. **Universal Words:**

A Universal Word represents simple or compound concepts. UWs are made up of a character string (an English-language word) followed by a list of constraints. There are three kinds of UWs.

1. Basic UWs
2. Restricted UWs
3. Extra UWs.

<UW> :: = <Head Word> [<Constraint List>]

<Head Word> :: = <character>…

<Constraint List> :: = "(" <Constraint> ["," <Constraint>]… ")"

<Constraint> :: = <Relation Label> { ">" | "<" } <UW> [<Constraint List>] |

<Relation Label> { ">" | "<" } <UW> [<Constraint List>]

[{ ">" | "<" } <UW> [<Constraint List>]] …

<Relation Label> :: = "agt" | and" | "aoj" | "obj" | "icl" | ...

<character> :: = “A” | ... | “Z” | “a” | ... | “z” | 0 | 1 | 2 | ... | 9 | “_” | ” “ | “#” | “!” | “$” | “%” |

“=” | “^” | “~” | “|” | “@” | “+” | “-“ | “<” | “>” | “?” | “ ”

Head Word:

The Head Word is an English word/compound word/phrase/sentence that is interpreted as a label for a set of concepts: the set made up of all the concepts that may correspond to that in English. A Basic UW (with no restrictions or Constraint List) denotes this set. Each Restricted UW denotes a subset of this set that is defined by its Constraint List. Thus, the headword serves to organize concepts and make it easier to remember which is which.

Basic UWs

A Basic UW is expressed by an English word/compound word/@hrase/sentence. The concept that a basic UW represents is the same concept that corresponding to that in English.

Restricted UWs

The Constraint List restricts the range of the concept that a Basic UW represents.

The Basic UW “drink”, with no Constraint List, includes the concepts of “putting liquids in the mouth”, “liquids that are put in the mouth”, “liquids with alcohol”, “absorb” and others.

The Restricted UW “drink(icl>do(obj>liquid))” denotes the subset of these concepts that includes “putting liquids in the mouth”, which in turn corresponds to verbs such as “drink”, “gulp”, “chug” and “slurp” in English.

Consider again the examples of Restricted UWs given above:

state(icl>do(obj>thing) is more specific concept

(arbitrarily associated with the English word “state”) that denotes situations in which humans produce some information, or state something.

state(icl>nation) is more specific sense of “state” that denotes a nation.

state(icl>situation) is more specific sense of “state” that denotes a kind of situation.

state(icl>government) is more specific sense of “state” that denotes a kind of government.

The information in parentheses is the Constraint List and it describes some conceptual restrictions that is why these are called Restricted UWs. Informally, the restrictions mean “restrict your attention to this particular sense of the word”. Thus, the focus is clearly the idea and not the specific English word.

It often turns out that for a given language there is a wide variety of different words for these concepts and not, coincidentally, all the same word, as in English.

Notice that by organizing these senses around the English words, we can simplify the task of making a new UW/Specific Language dictionary: we can use a bilingual English/Specific Language dictionary and proceed from there, specifying the number different concepts necessary for each English word.

This of course does not mean that we’re translating English words; we’re just using the English dictionary to remind us of the concepts that we will want to deal with and thus to organize work more efficiently.

Extra UWs:

Extra UWs denote concepts that are not found in English and that have to be introduced as extra categories. Foreign-language labels are used as Head Words. Consider again the examples given above:

samba(icl>dance) “a kind of dance”

omelet (icl>food, pof>egg) “a kind of food made with eggs”

murano(icl>glass, aoj>colorful) “a kind of colorful glass”

To the extent that these concepts exist for English speakers, they are expressed with foreign-language loanwords and don't always appear in English dictionaries. So, they simply have to be added if we are going to be able to use these specific concepts in the UNL system. Notice that the Constraint List or restrictions already give some idea of what concept is associated with these Extra UWs and the Constraints binary relation this concept to other concepts already present (activity, flower, egg, food, etc.).

2. **Relations**

Binary relations are the building blocks of UNL sentences. They are made up of a relation and two UWs. The relations between UWs in binary relations have different labels according to the different roles they play. These Relation-Labels are listed and defined below. There are many factors to be considered in choosing an inventory of relations. The principles to choose relations as follows

Principle 1: Necessary Condition

When a UW has relations between more than two other UWs, each relation label should be set as to be able to identify each relation on the premise that we have enough knowledge about a concept of each UW express.

Principle 2: Sufficient Condition

When there are relations between UWs, each relation label, we should be set as to be able to understand each role of each UW only by referring a relation label.

Universal relations are represented as follows:

<Relation> : <scope> (<source>; <target>)

Where:

<Relation> is the name of the relation (two-character or three-character lower-case strings) (see the complete list of relations below)

<Scope> is the scope of the relation (two-character unique identifier for the scope). The scope may be omitted if the main scope, i.e., :00. See scope.

<Source> is the UW that assigns the relation <rel>

<Target> is the UW that receives the relation <rel>.

Agt	Agent	a thing which initiates an action
And	Conjunction	a conjunctive relation between concepts
Aoj	thing with attribute	a thing which is in a state or has an attribute
Bas	Basis	a thing used as the basis(standard) for expressing degree
Ben	Beneficiary	a not directly related beneficiary or victim of an event or state
Cag	co-agent	a thing not in focus which initiates an implicit event which is done in parallel
Cao	co-thing with attribute	a thing not in focus is in a state in parallel
Cnt	Content	an equivalent concept
Cob	affected co-thing	a thing which is directly affected by an implicit event done in parallel or an implicit state in parallel
Con	Condition	a non-focused event or state which conditioned a focused event or state
Coo	co-occurrence	a co-occurred event or state for a focused event or state
Dur	Duration	a period of time during an event occurs or a state exists
Fmt	Range	a range between two things
Frm	Origin	an origin of a thing
Gol	goal/final state	the final state of object or the thing finally associated with object of an event
Ins	Instrument	the instrument to carry out an event
man	Manner	the way to carry out event or characteristics of a state
Met	Method	a means to carry out an event
mod	Modification	a thing which restrict a focused thing
Nam	Name	a name of a thing
Obj	affected thing	a thing in focus which is directly affected by an event or state
Opl	affected place	a place in focus where an event affects
Or	Disjunction	disjunctive relation between two concepts
Per	proportion, rate or distribution	a basis or unit of proportion, rate or distribution
plc	Place	the place an event occurs or a state is true or a thing exists
Plf	initial place	the place an event begins or a state becomes true
plt	final place	the place an event ends or a state becomes false
Pof	part-of	a concept of which a focused thing is a part
Pos	possessor	the possessor of a thing
Ptn	Partner	an indispensable non-focused initiator of an action
Pur	purpose or objective	the purpose or an objective of an agent of an event or a purpose of a thing which exist
Qua	Quantity	a quantity of a thing or unit
Rsn	Reason	a reason that an event or a state happens
Scn	Scene	a virtual world where an event occurs or state is true or a thing exists
Seq	Sequence	a prior event or state of a focused event or state
Src	source/initial state	the initial state of an object or thing initially associated with the object of an event
Tim	Time	the time an event occurs or a state is true
Tmf	initial time	the time an event starts or a state becomes true
Tmt	final time	the time an event ends or a state becomes false
To	Destination	a destination of a thing
Via	intermediate place or state	an intermediate place or state of an event

3. Attributes

Attributes of UWs are used to describe subjectivity of sentences. They show what is said from the speaker's point of view: how the speaker views what is said. This includes phenomena technically called "speech acts", "propositional attitudes", "truth values", etc. Conceptual relations and UWs are used to describe objectivity of sentences. Attributed of UWs enrich this description with more information about how the speaker views these states-of-affairs and his attitudes toward them.

Time with respect to the speaker

Where does the speaker situate his description in time, taking his moment of speaking as a point of reference? A time before he spoke? After? At approximately the same time? This is the information that defines "narrative time" as past, present or future. These Attributes are attached to the main predicate.

Although in many languages this information is signaled by tense markings on verbs, the concept is not tense, but "time with respect to the speaker". The clearest example is the simple present tense in English, which is not interpreted as present time, but as "independently of specific times".

Consider the example: The earth is round.

This sentence is true in the past, in the present and in the future, independently of speaker time, so although the tense is "present" it is not interpreted as present time.

@past happened in the past

ex) He went there yesterday.

ex) It was snowing yesterday

@present happening at present

ex) It's raining hard.

@future will happen in future

ex) He will arrive tomorrow

Speaker's view of Aspect:

A speaker can emphasize or focus on a part of an event or treat it as a whole unit. This is closely linked to how the speaker places the event in time. These Attributes are attached to the main predicate.

He can focus on the beginning of the event, looking forward to it (@begin), or backward to it (@begin).

He can also focus on the end of the event, looking forward to it (@end) or backward to it from nearby (@end) or from farther away (@complete).

Degree of forwardness or backwardness (@soon, @just).

He can focus on the middle of the event (@progress).

The speaker can choose to focus on the lasting effects or final state of the event (@state) or on the event as a repeating unit (@repeat).

The feeling of incompleteness or not yet happen of an event with respect to the speaker (@yet).

@begin	beginning of an event or a state	
	ex) It began to work again.	work.@begin.@past
@complete	finishing/completion of a (whole) event.	
	ex) I've looked through the script	look.@entry.@complete
@continue	continuation of an event ex) He went on talking.	talk.@continue.@past

Speaker's view of Reference

Whether an expression refers to a single individual, a small group or a whole set is often not clear. The expression "the lion" is not sufficiently explicit for us to know whether the speaker means "one particular lion" or "all lions". Consider the following examples:

The lion is a feline mammal.
The lion is eating an anti-lope.

In the first example, it seems reasonable to suppose that the speaker understood "the lion" as "all lions", whereas in the second example as "one particular lion". The following Attributes are used to make explicit what the speaker's view of reference seems to be.

@generic	generic concept
@def	already referred
@indef	non-specific class
@not	complement set
@ordinal	ordinal number

These Attributes are usually attached to UWs that denote things.

Speaker's Focus

The speaker can choose to focus or emphasize the parts of a sentence to show how important he thinks they are in the situation described. This is often related to sentence structure.

@emphasis ***Emphasis***

ex) “I do like it”

@entry	Entry point or main UW of whole UNL expressions or in a hyper (scope) node
@title	Title
@topic	The topic UW of a sentence

One UW marked with "@entry" is essential to each UNL expression or in a Compound UW.

Speaker’s attitudes:

The speaker can also express, directly or indirectly, what his attitudes or emotions are toward what is being said or who it is being said to. This includes respect and politeness toward the listener and surprise toward what is being said.

@imperative **Imperative**

ex) “Get up!” ex) “You will please leave the room.”

@interrogative **Interrogation**

ex) “Who is it?”

The variety of possibilities reflects degrees of belief, emphasis, and the extent to which what is said should be interpreted as a suggestion or order, as well as many other social factors such as the relative status of the speakers.

Speaker's view point

The following labels are used to clarify the speaker’s viewpoint information

@ability ability, capability of doing things

ex) He can speak English but he can't write it very well.

@will will to do

ex) I shall not be long.

@obligation to oblige someone

ex) The vendor shall maintain the equipment in good repair.

Convention:

Typical UNL structures can be expressed by attribute, to avoid the complexity of enconverting and deconverting. These attributes do not express speaker’s information.

@pl Plural

@single quotation ‘ ’ is used

X BAR THEORY

We can represent the individual vocabulary items of a language as small pieces of syntactic structure, or elementary trees.

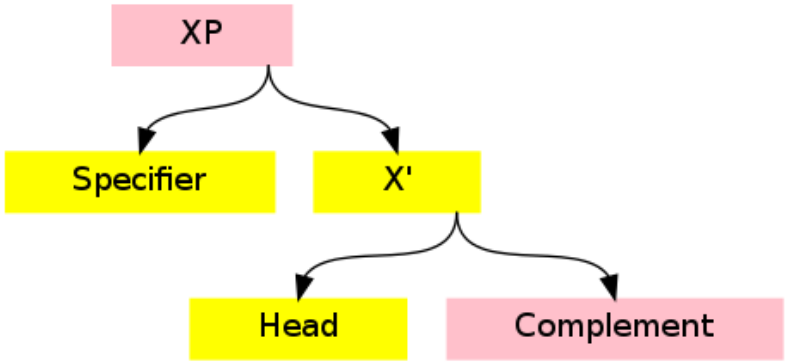
In this view, vocabulary items are comparable to the atoms of physical matter. Atoms do not combine into molecules just because they happen to be next to each other; rather, their combinatorial possibilities are governed by their internal structure (for instance, the number of electrons on an atom's outermost shell and the relative number of protons and electrons).

Accordingly, in the first part, we consider the internal structure of elementary trees. Later, we begin by focusing on how verbs combine with their arguments to form larger phrases. Movement and adjunction are ways for modification of trees

The X-bar schema:

(a) This is a model of the syntactic component of mental grammar

(b) We will first learn how the model works, and then begin to test it on/apply it to language data



Here

XP is the maximal projection

X' is the intermediate projection

Head: The word that is the “core” of the phrase. It determines the type of phrase (X is the head of XP, for any X)

Complement: The phrase that is the sister of X (the head). A complement is a phrase that the head requires inside its own phrase

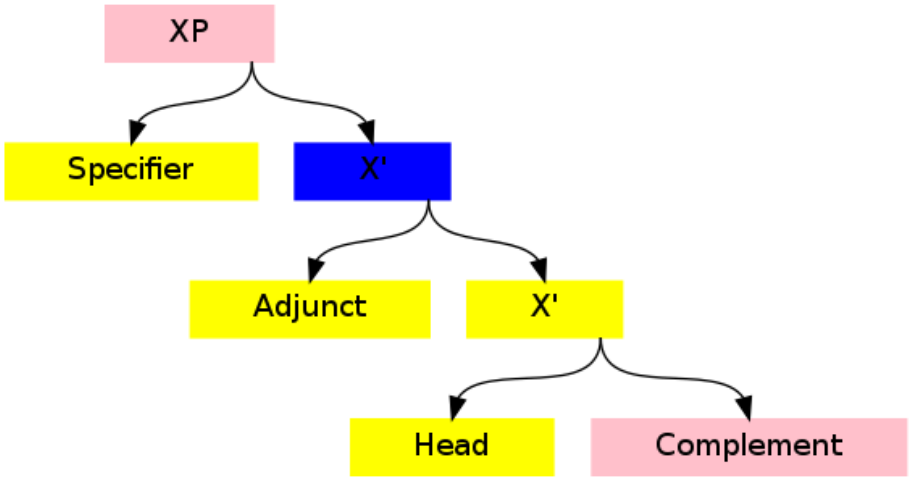
Examples: Direct object (complement of V); object of preposition (complement of P)

Specifier: The phrase that is the daughter of XP. A specifier is a “subject-like” phrase that occurs with a head

Examples: Possessor (spec. of N); certain adverbs (spec. of V, A, P)

The linear order of elements (left-to-right) is language-specific. Nodes are generally binary-branching. The nodes of the UNL hyper graph are lexicons. Each lexicon has a head which is the main word.

Adjuncts are ways to modify trees or connect two different trees. These cause recursive X' nodes to appear. There could be left or right adjunct based on if the adjunct is the left daughter or the right daughter of the X' node

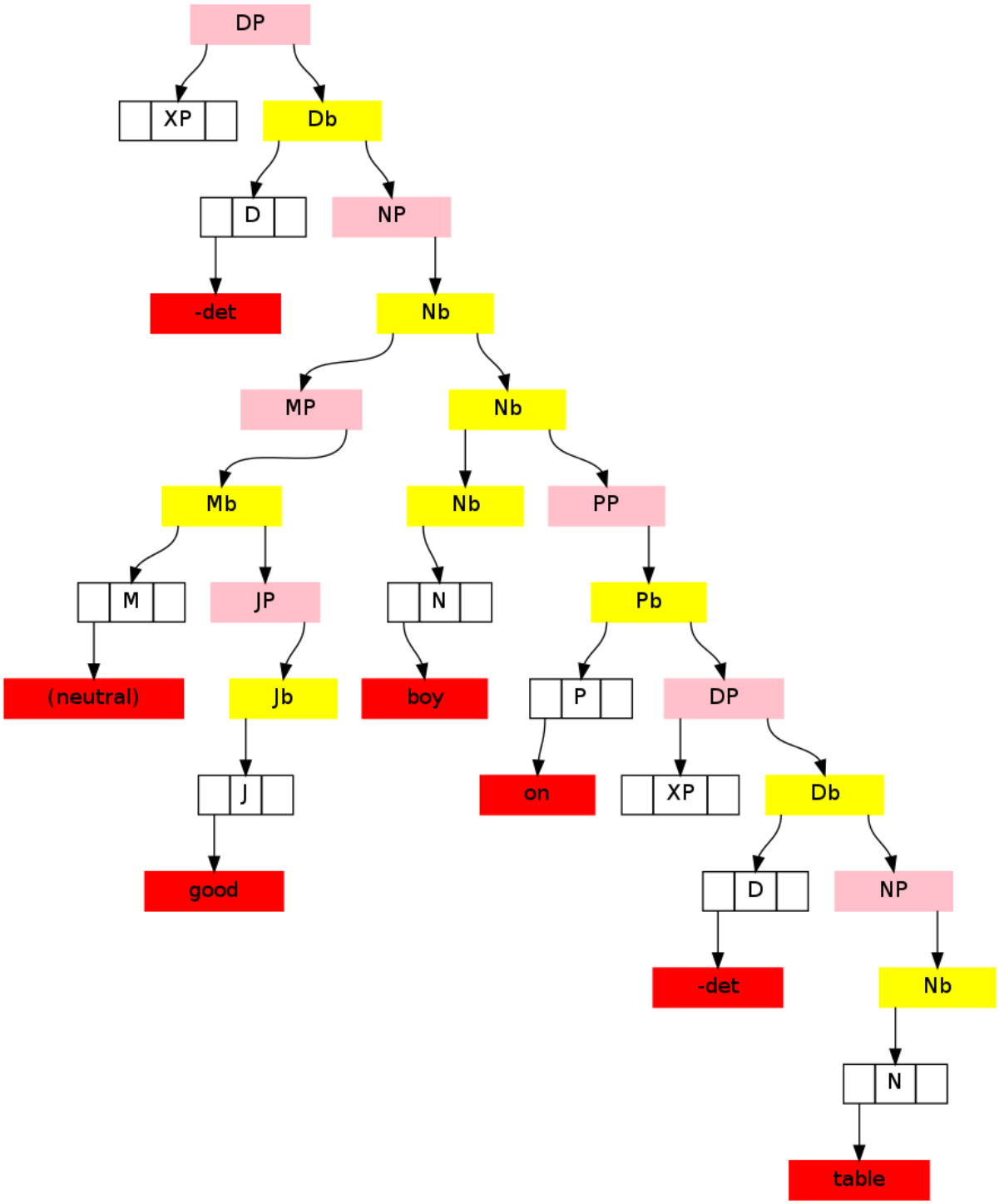


Heads can be divided into lexical categories (N, V, A, P). These words are heads of phrases (the phrases may also contain other elements, which generally provide more information about the head). Whenever you see a head of one of these categories, it must project a phrase.

Noun Phrases:

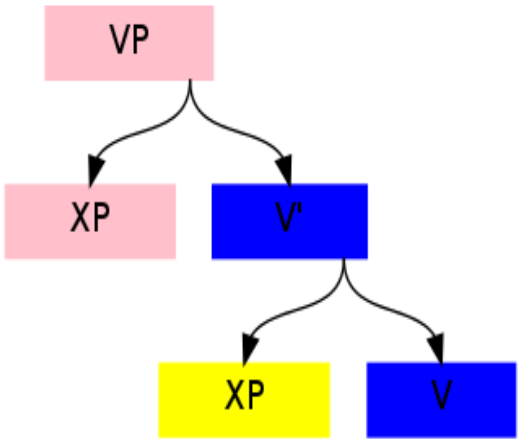
A striking fact about nouns is that they cannot in general function as arguments on their own, but must be accompanied by a determiner. So, NP and DP are considered to be two different categories. In English, prepositions are followed by nouns. So, a PP structure would capture these prepositions like in the example “on the table”, “across the road”, etc. Pronouns are also similarly handled like in the case of nouns using the DP structure. Adjectives and Adverbs are handled as Adjuncts to the NP or VP. Thus, structures are assigned to all types of lexicons which are nouns based on their role played in the sentence.

Examples:

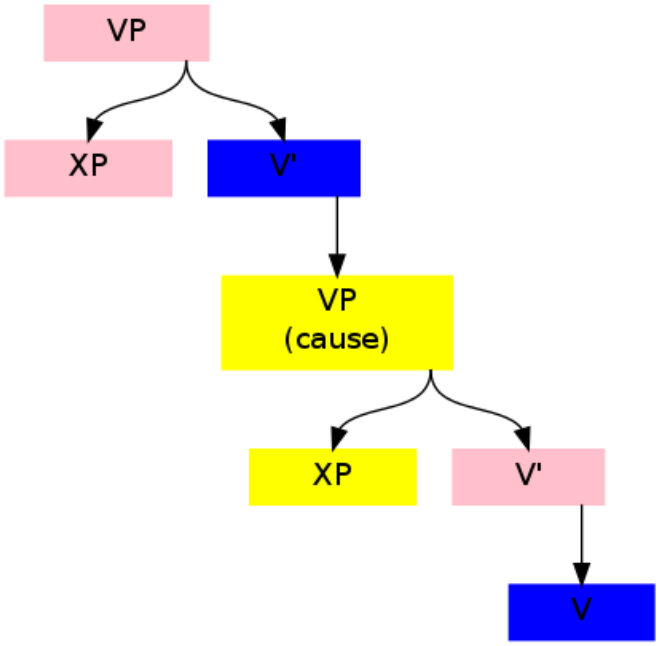


Verb Phrases:

Intransitive verbs are handled using the VP structure and Ditransitive verbs are handled using the VP shell. Structure for all verbs remains among these two structures but Attributes may vary depending on the verb. So, structures for verbs are assigned in the following ways based on the verb taking an animate form or an inanimate form of agent or object.



The



Sentence is considered to be an IP projection of the VP.

(a) The head of a sentence is I, the “inflectional head”; morphemes (possibly abstract/invisible ones) involving grammatical features such as verb tense and modality (possibility, necessity, etc.) typically go in this position

(b) The complement of an IP is the predicate of the sentence

(c) The specifier of an IP is the subject of the sentence

Complementizer phrases (CPs):

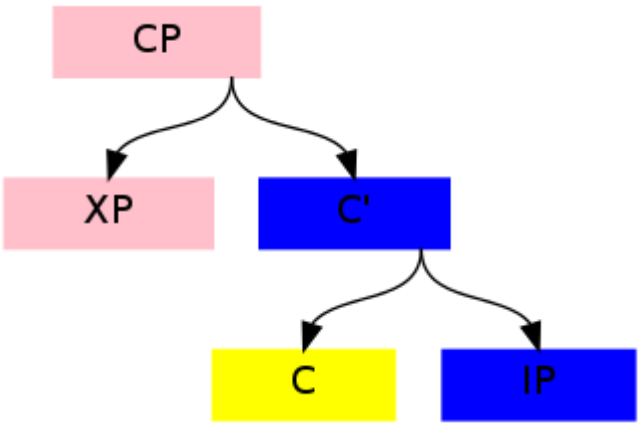
(a) A Complementizer (C) is a head (i.e., word) that turns a sentence (IP) into something that can be a complement

Example: Embedded clauses

(b) The complement of C is IP

(c) The specifier of C is a very useful position to move things into in some languages

(d) A main-clause (matrix) IP is probably also contained inside a CP, but we can’t necessarily see that until we start looking at the syntax of questions

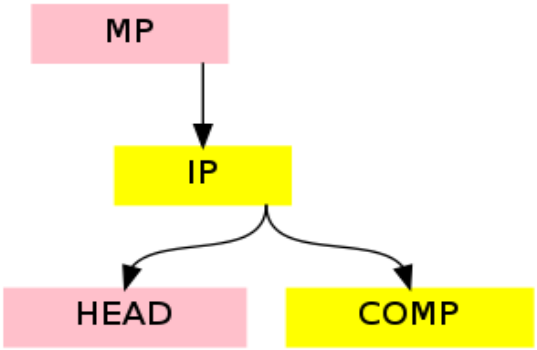


CHANGES IN THE ENGLISH DECONVERTER TO HINDI DECONVERTER

Internal Structure of Elementary trees:

English is a Head start language. Elementary trees of all syntactic categories follow the same structural schema. So, sentences are built by reading the leaves of the tree in the order

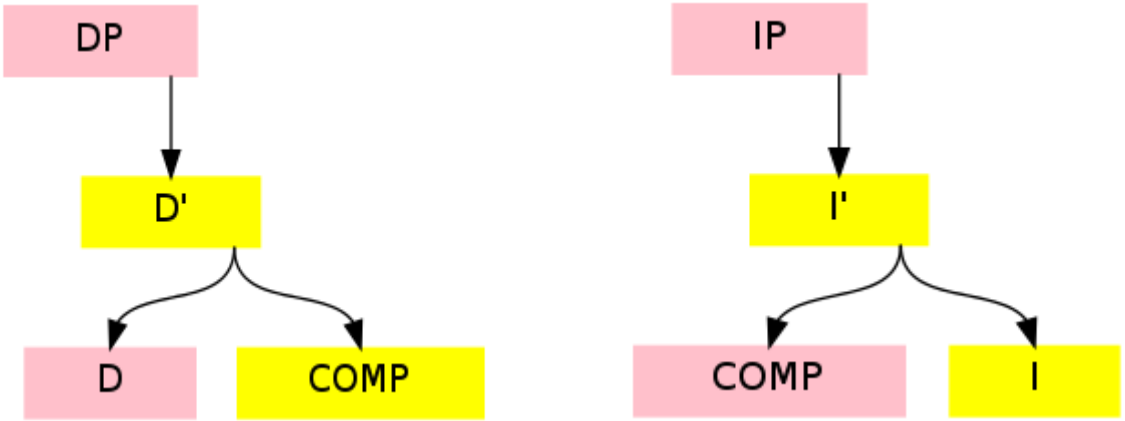
Specifier -> Head -> Complement



Hindi does not follow the same structural schema for all the syntactic category. 'DP' and 'CP' follow the head start structure whereas 'IP' and 'VP' follow the complement start structure. So, Sentences are built by reading the leaves differently for different category.

For 'DP', 'CP' : Specifier -> Head -> Complement

For 'VP', 'IP' : Specifier -> Complement -> Head



Handling of Sentences where there is UNL relation between two verbs:

In English, relations like 'obj', 'cnt', 'pur', 'rsn' could be there between two verbs. Two different cases arise depending on whether there is an 'agt' relation to the relation arrow pointing verb

Example:

1. I like to work
2. I like that i work

So, if there is no agent to the relation arrow pointing verb, then a 'infinitive' attribute is given to relation arrow pointing verb from which the word 'to' can be projected.

Also, there are sentences which may have infinitive attribute but do not project a 'to'

Example:

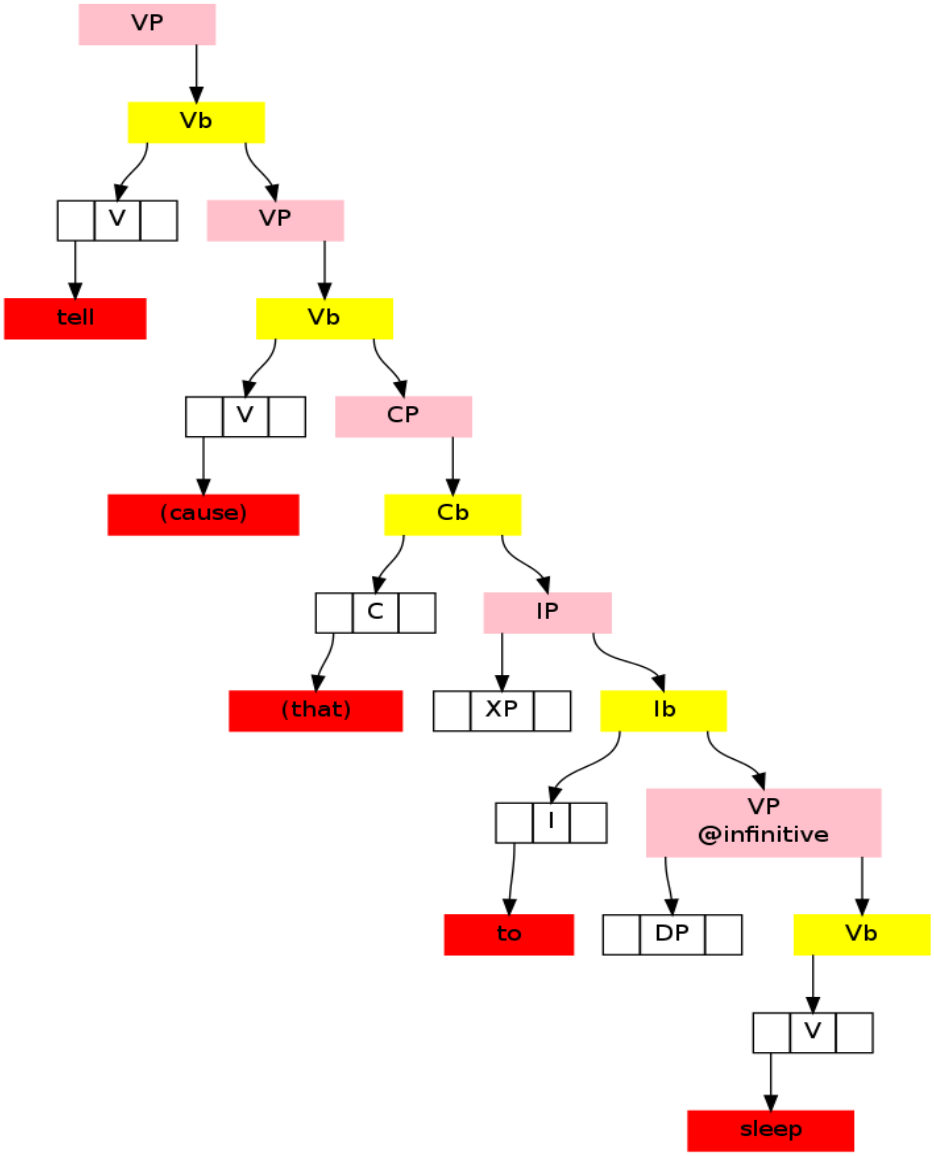
1. I let go
2. I like to work
3. I start running

Verbs like 'let' and 'start' are marked in dictionary with .INONE or .IGER which projected a masked 'to'.

In case of Hindi, the relation 'cnt' behaves in a very similar way as discussed above but relations like 'obj', 'pur', 'rsn' will have an infinitive attribute irrespective of whether the arrow pointing verb has a agent or not

Example:

1. pur mai tujhko bolanE kE liyE tujhko bulAtA hu
2. pur mai tujhko bolanE kE liyE mai tujhko bulAtA hu
3. rsn mai tujhko bolanE kI vajaha sE tujhko bulAtA hu
4. rsn mai tujhko bolanE kI vajaha sE mai tujhko bulAtA hu
5. cnt mai tujhko kehtha hu ki mai sotA hu
6. cnt mai tujhko sonE kE liyE kehthA hu



In English, infinitive attribute always projects a 'to'. In Hindi, infinitive attribute can project 'ko', 'ki', 'sE'. This is a property of the relation arrow starting verb and is included in the dictionary using the attributes .INko , .INsE, .Inki

In Hindi, the arrow starting verb decides the number of the arrow pointing verb. This again is also a property of the arrow starting verb.

Example:

1. singular mai sonE ko kahatA thA
2. plural mai sonA chAhatA hu

Copying attributes while projecting IP to a VP:

In English, attributes are assigned only to the projected IP and not the VP or the IP to which the new IP is projected

In Hindi, attributes of the newly projected IP is copied to the VP or IP to which it is projected

Example:

1. mai chaltA rahA hu
2. hum chaltE rahE hai

Copying attributes due to semantic relations:

In English, only number and person attributes are copied in case of agt relation and also in relation is rel(x,y) , attributes are copied from y to x.

Example:

1. I am going
2. He is going
3. We are going

In case of Hindi, Number, person and gender attributes are copied and they are copied in many relations like agt, aoj, pos, obj, mod, man, qua. Also, in these relations { rel(x,y) }, some relations require copying of attributes from x to y and others from y to x

Copying attributes:

1. Copying attributes from x to y : pos, man, mod, qua
2. Copying attributes from y to x : agt, obj, aoj

Example:

1. ladakA dekhtA thA
2. ladakI dekhtI thI
3. ladakA acchA thA
4. ladakI acchI thI

Case Agreement:

In English, there are few instances of Case assignment. It happens only in agt relation and case is assigned to the verb

Example:

1. I sleep.
2. He sleeps.

In Hindi, Case is assigned by the verb, nouns to nouns connecting through relations like agt, obj, gol, pos. In case of agt, case is assigned only when the verb is in perfect form. So, the case decided by tense, modality of verb and person, number and gender of the noun.

Example:

1. mai ladakE ko marathA hu
2. ladakE ki mAtA
3. ladakE kA pitA
4. mainE marA thA

Wh Movement:

There is no wh movement in Hindi as in English. whr and wh have different morphology unlike in English.

Example:

1. Who hit me?
2. You who hit me ran.

1. kaun mujhko mArathA hai ?
2. tu jo mujhko mArathA hai bhAgathA hai.

Morphological change due to modality of verb:

In Hindi, there are modality like obligation which has morphological changes to verb unlike in English

Example:

1. Vah jAtA hai.
2. usE jAnA chahiyE.

Framing rules of the lexicons remain the same for both the languages but the semantic rules and morphological rules changes

RESULTS

Few examples where the x bar tree shown when UNL sentence is passed through a deconverter .

Example:

UNL Sentence:

agt (201405044, 00.@1)

obj (201405044, 110285313)

Hindi Output: मै लडका को मारता हुं.

English Output: I hit the boy.

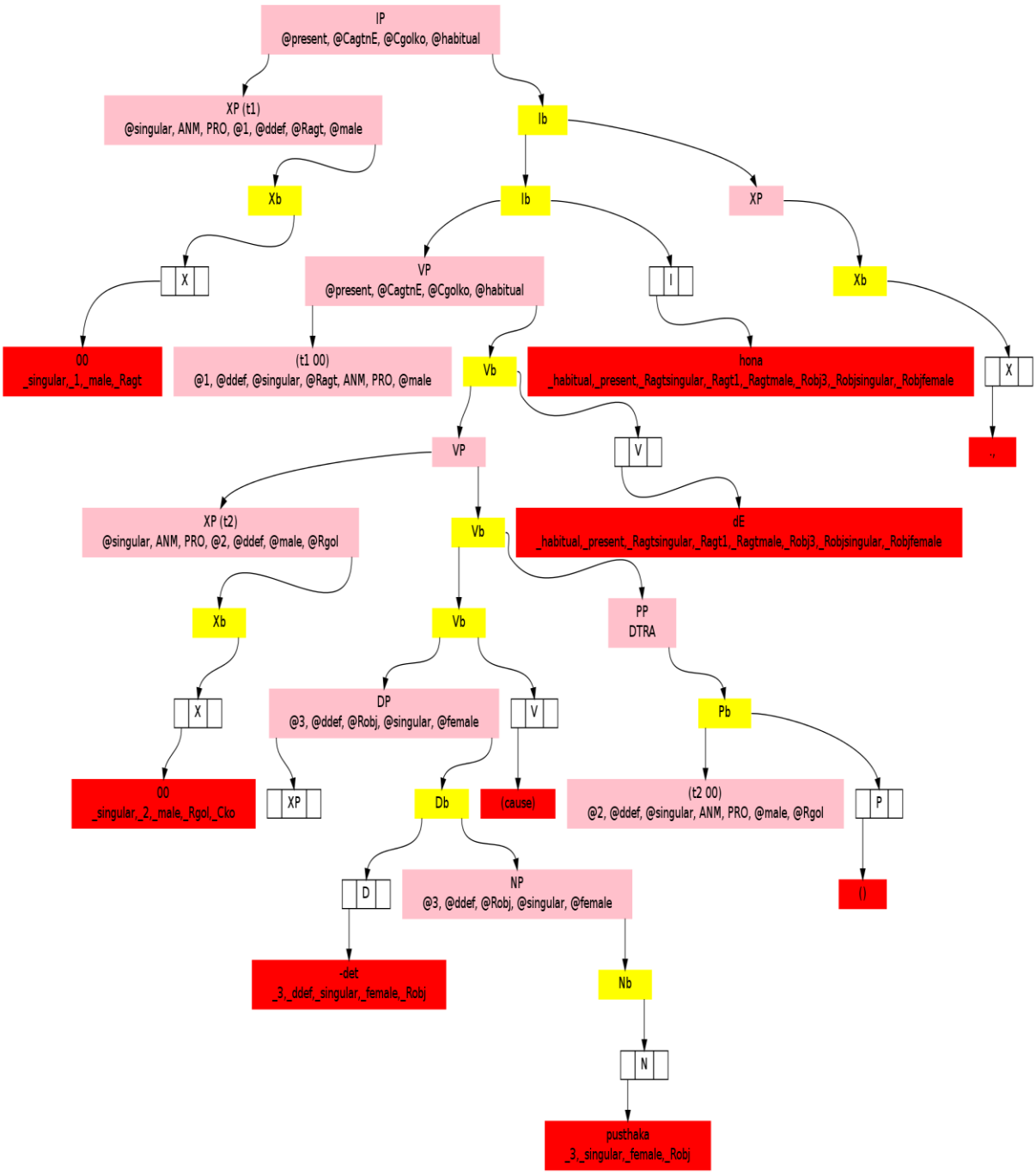
Example 2:

UNL sentence:

1. agt (202316868, 00.@1)
2. obj(202316868, 106410904)
3. gol(202316868, 00.@2)

Hindi Output : मै तुज्को पुस्तक देता हुं.

English Output : I gave you a book.



Example 3:

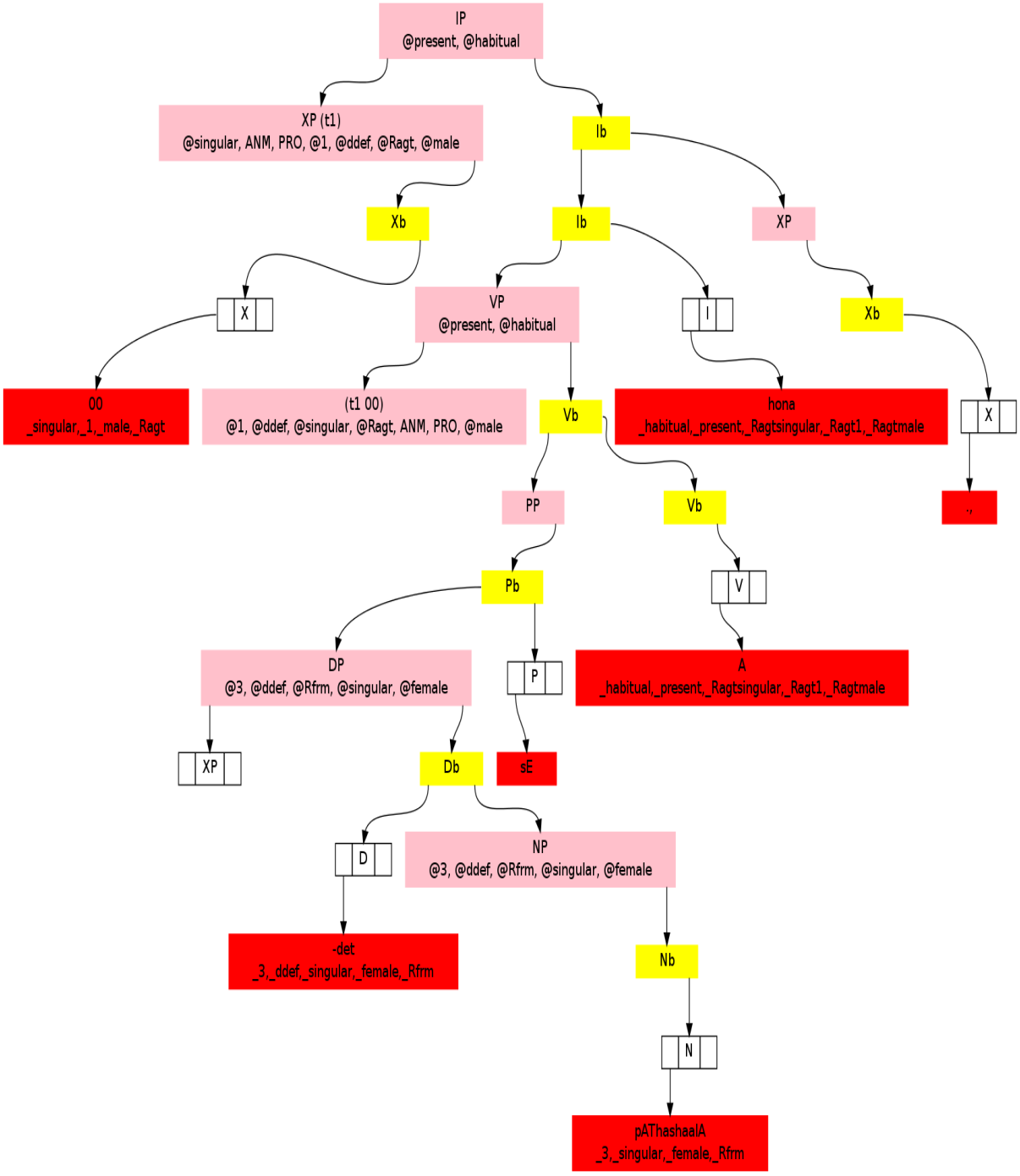
UNL sentence:

agt (201904930, 00.@1)

to(201904930, 108559508)

Hindi Output: मै घर चलता हुं.

English Output: I walk home.



Example 4:

UNL sentence:

agt(201170052,00:2.@2)

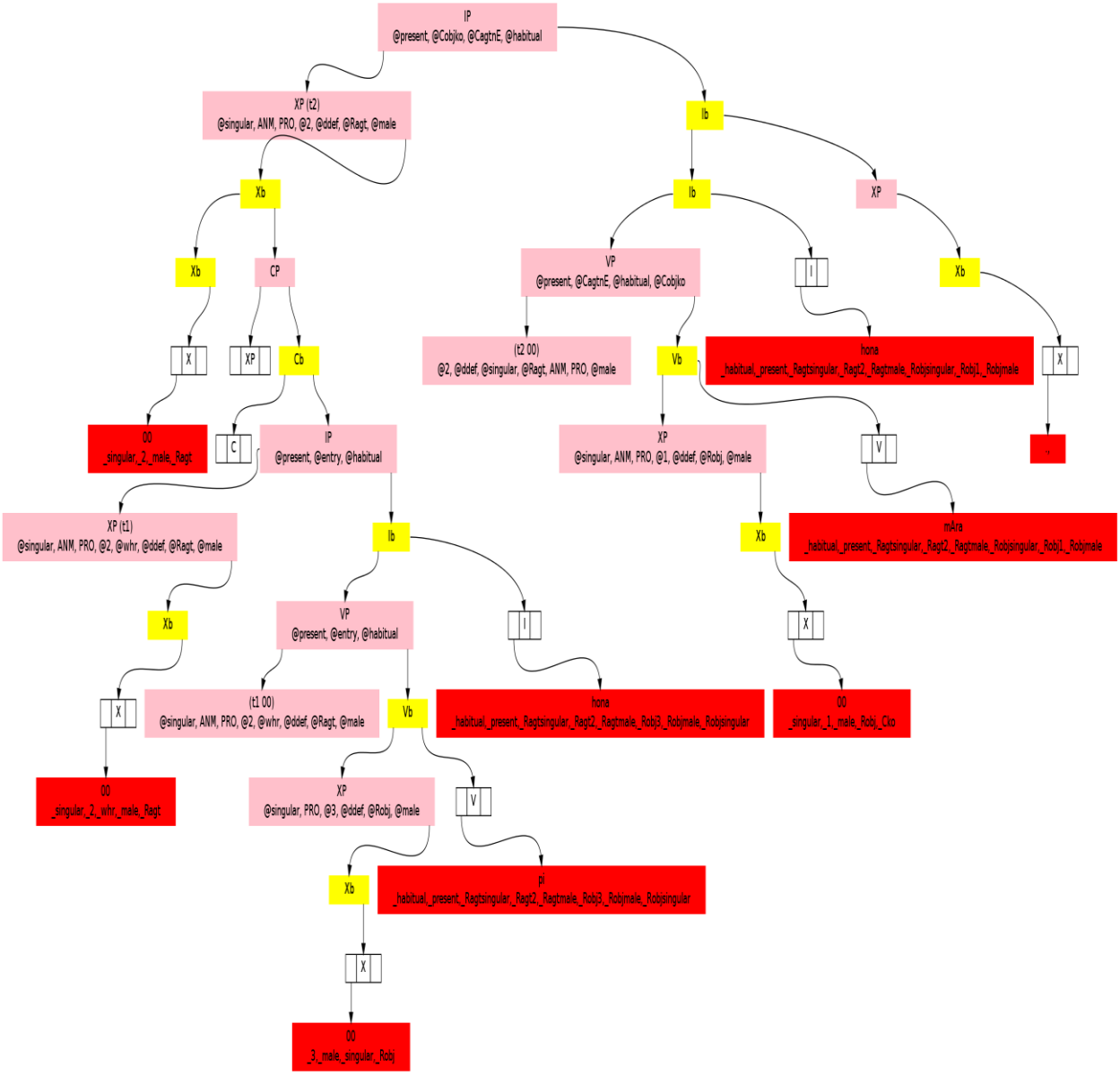
obj(201170052,00)

agt(201405044,00:2.@2)

obj(201405044,00.@1)

Hindi Output: तु जो वह सुनता है मुझ्को मारता है.

English Output: you who heard it hit me.



REFERENCES

1. Free Speech application code by Inventions Labs
2. http://www.ling.upenn.edu/~beatrice/syntax-textbook/
3. http://www.unlweb.net/unlweb/
4. http://www.undlfoundation.org/undlfoundation/
5. http://en.wikipedia.org/wiki/Hindi_grammar