CS626: Speech, NLP and the Web

Semantic Role Labelling (SRL)

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Week of 18th October, 2021

Motivation for Semantic Roles: Deeper than DP relations

Sentence	Shallow relation from Dependency Parsing	Deeper relation from Semantic Role Labeling
John broke the window	nsubj	Agent
The stone broke the window	nsubj	Instrument
The window broke	nsubj	Object
1947 saw the freedom of India	nsubj	Time
Delhi saw bloodshed when Nadir Shah attacked Delhi	nsubj	Place

Disambiguation is needed to convert shallow DP relations to semantic roles.

Question Answering

- Given “The stone broke the window”, if the question is “who broke the window?”

- The answer CANNOT be “the stone”

Motivation for Semantic Roles: Deeper than DP relations

Sentence	Shallow relation from Dependency Parsing	Deeper relation from Semantic Role Labeling
John broke the window	nsubj	*agt(break:02@past, John(icl>person))*
The stone broke the window	nsubj	*ins(break:02@past, stone:01@def)*
The window broke	nsubj	*obj(break:02@past, window:01@def)*
1947 saw the freedom of India	nsubj	*tim(see:01@past, 1947(icl>year))*
Delhi saw bloodshed when Nadir Shah attacked Delhi	nsubj	*plc(see:01@past, Delhi(icl>place)*

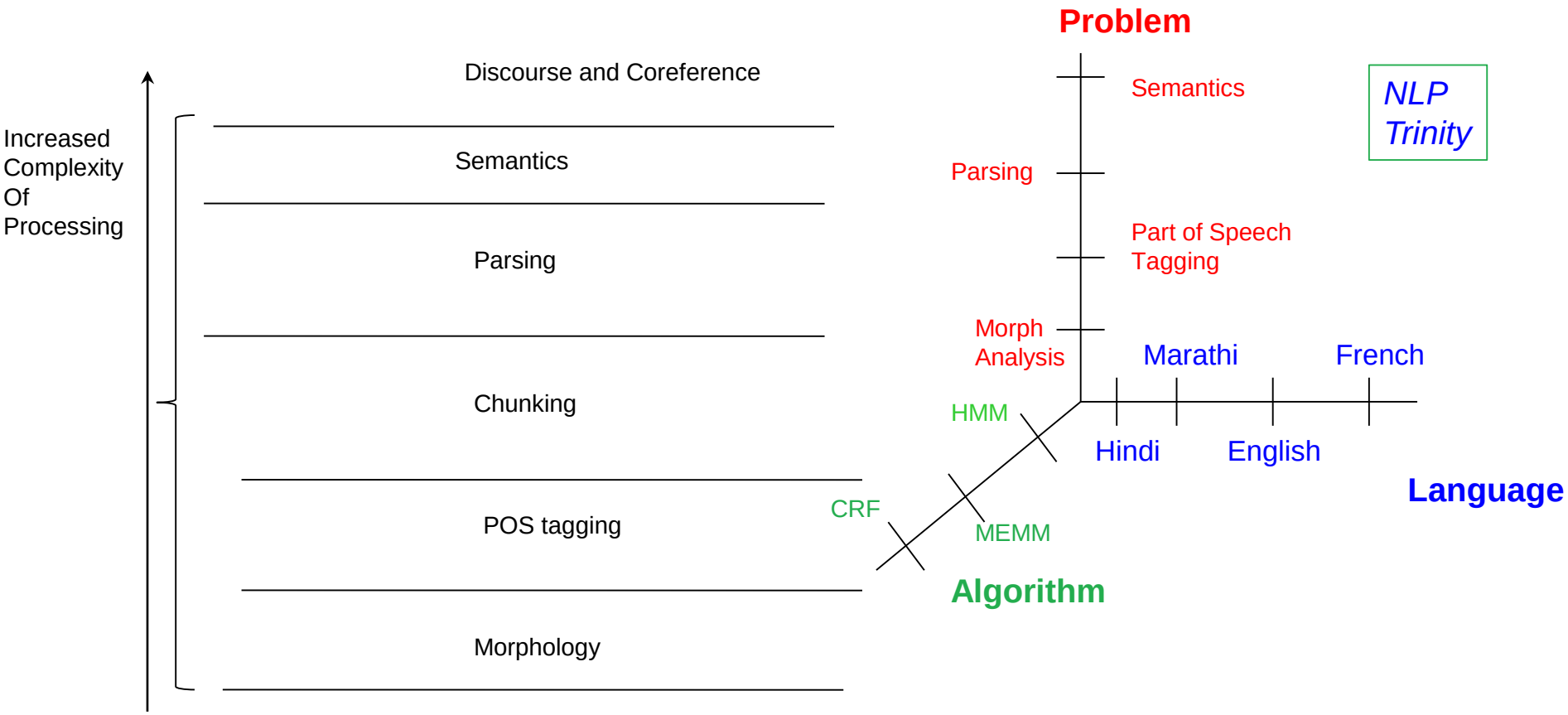
Disambiguation is also needed for word senses. Break:02 indicates the 2nd sense in wordnet

Question Answering supported by Semantic Graph

- Given “The stone broke the window”, if the question is “who broke the window?”

- The answer is “cannot tell”

NLP is layered Processing, Multidimensional too



A classic paper on SRL

Gildea and Jurafsky, Automatic Labeling of Semantic Roles, Computational Linguistics, 2002

Frames

Judgement Frame:

[*Judge* She] **blames** [*Evaluee* the Government] [*Reason* for failing to do enough to help]

Statement Frame:

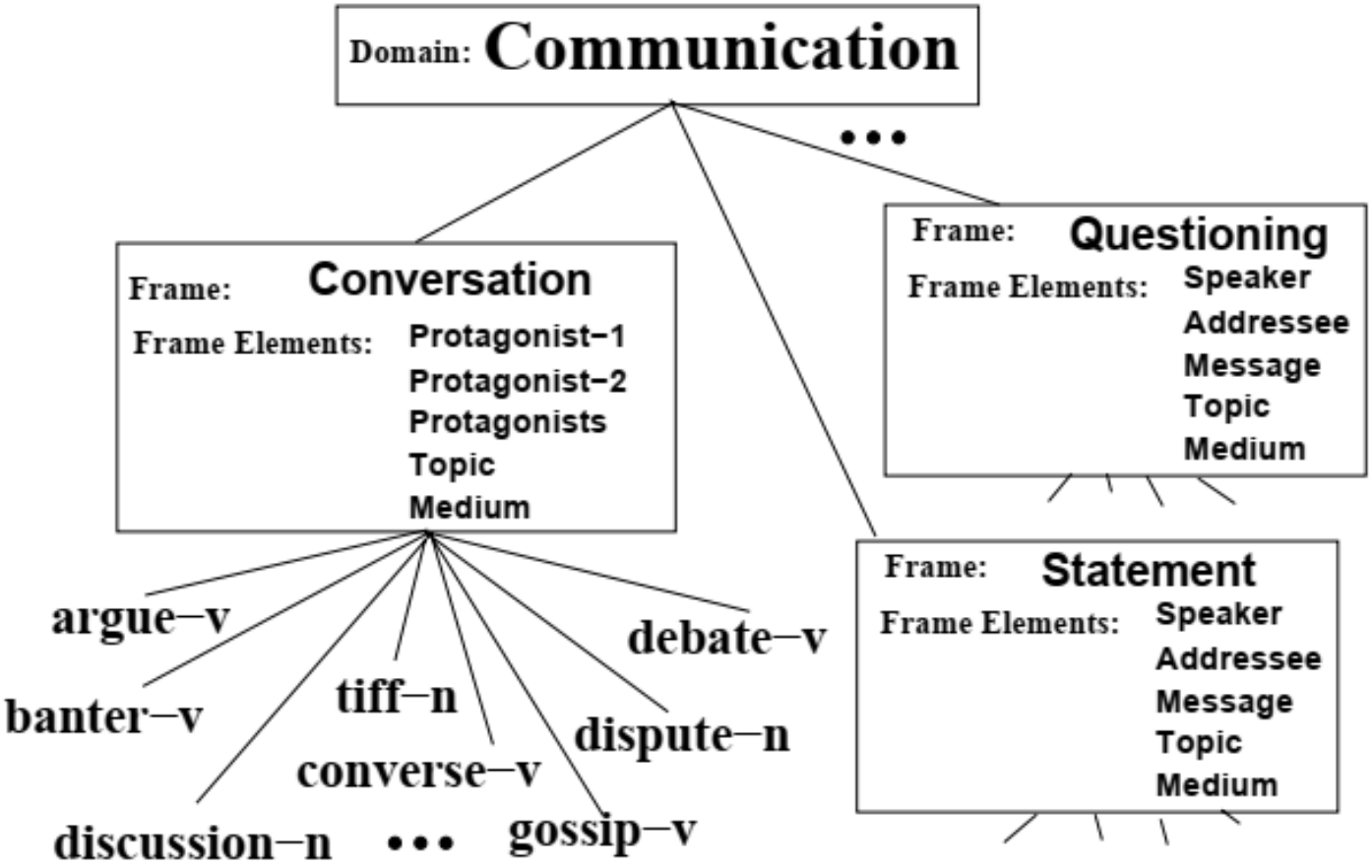
[*Message* “I’ll knock on your door at quarter to six”] [*Speaker* Susan] **said**

Frame Labelled Data

- Hand-labeled: the FrameNet database (Baker, Fillmore, and Lowe, 1998; Johnson et al., 2001)
- The FrameNet database defines a tagset of semantic roles called **frame elements**
- Roughly 50,000 sentences from the British National Corpus hand-labeled with these frame elements.

Example Frame: “Communication”

FrameNet project (Baker, Fillmore, and Lowe, 1998)



Makes use of huge amount of prior work on verbs

- Beth Levin's verb classes
- Indian ontology tradition: Amar Kosha
- FrameNets, Propbanks, VerbNets

Two disambiguations

- Identify correct frame
  - *Vaha khana khaa rahaa hai* (normal “eat” frame)
  - *Vaha meraa sar khaa rahaa hai* (metaphorical “to bore”, normal “eat” frame will not work)
- Identify correct role (slot in frame)
  - can face ambiguity: *visiting aunts can be interesting: aunts are “visited” (obj) or are visitors (agt)*
  - Similarly *aap ko mujhe mithai khilani padegi*

Stages in getting the SRL

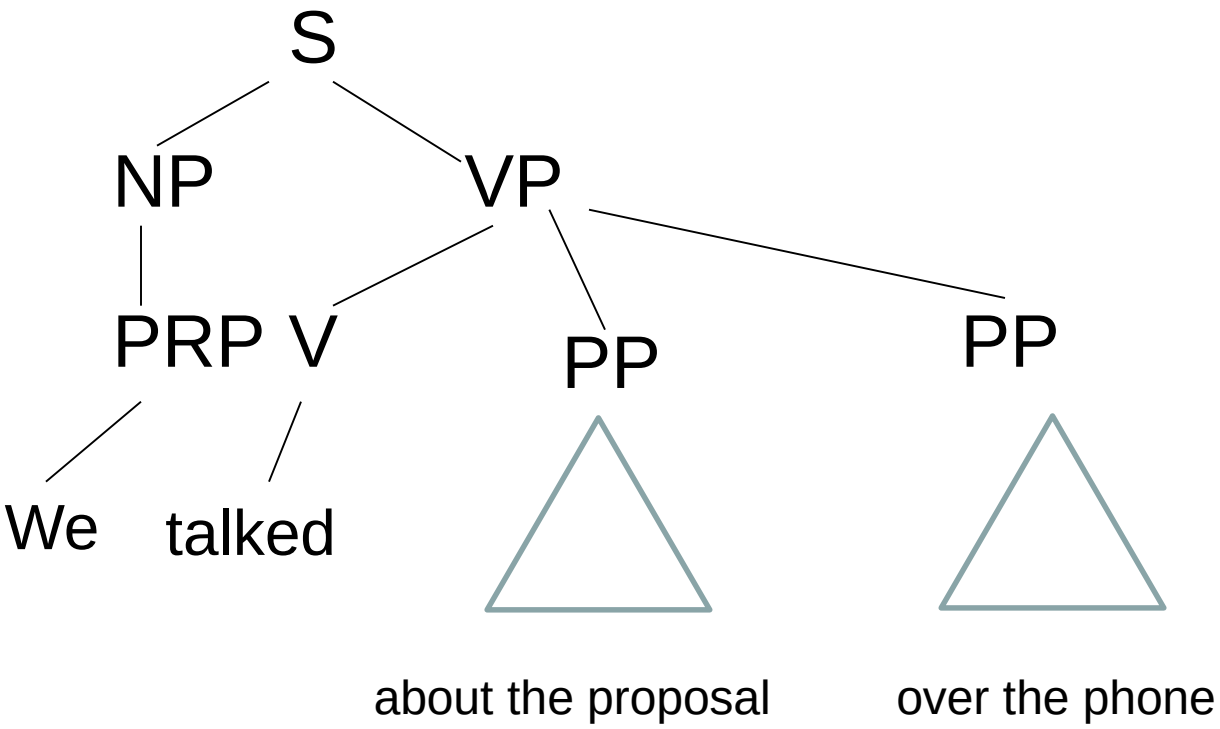
- Sentence- and constituent-level features
- Use these features to calculate probabilities for predicting frame element labels
- Labels roles using the human-annotated boundaries for the frame elements

Features

- Statistical classifier, trained by first using an automatic syntactic parser to analyze the 36,995 training sentences
- Match annotated frame elements to parse constituents
- Extract features from the string of words and the parse tree
- During testing, run the parser on the test sentences and the same features are extracted
- Probabilities for each possible semantic role *r* are then computed from the features

Feature-1: Phrase Type

[*Speaker* We] talked [*Topic* about the proposal] [*Medium* over the phone]

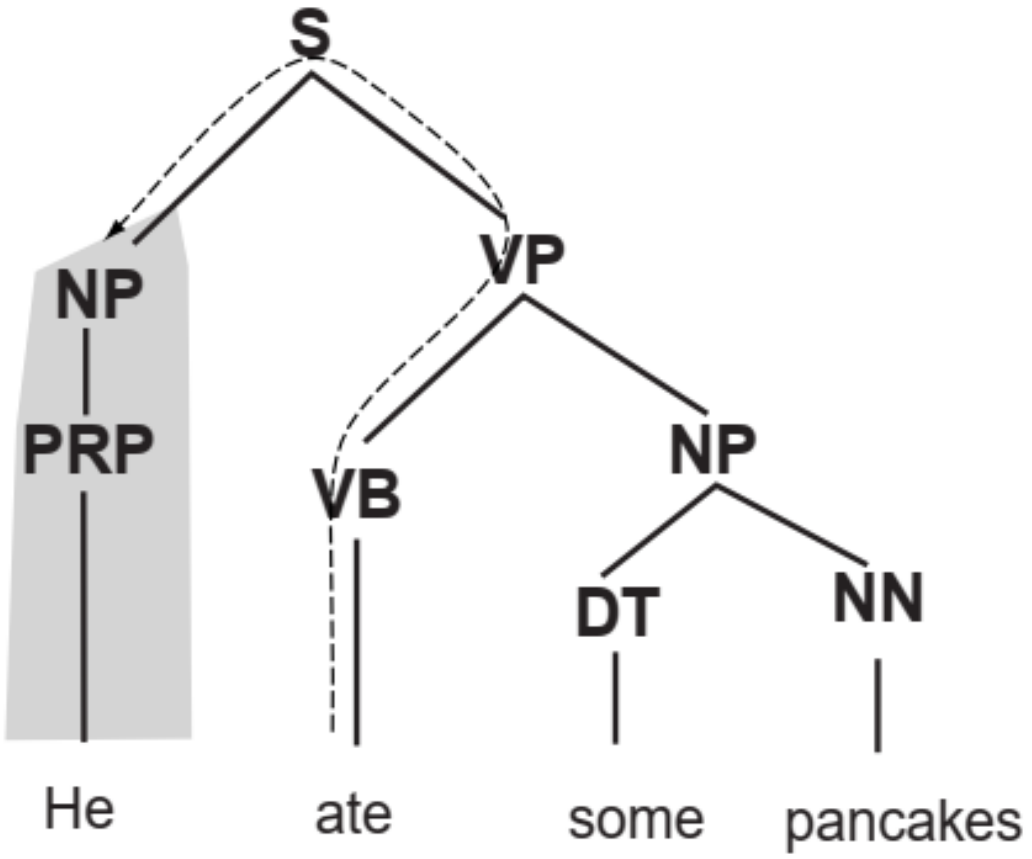


Phrase Type: Stats

- NP: 47% of frame elements in the training set
- PP: 22%
- ADVP: 4%
- PRT (e.g., “make something *up*): 2%
- SBAR (clauses): 2%
- S: 2%

Feature-2: Parse Tree Path

path(ate, he): $VB_{up}VP_{up}S_{up}NP_{dn}$



Path Features: Stats

Frequency	*Path*	*Description*
14.2%	VB↑VP↓PP	PP argument/adjunct
11.8	VB↑VP↑S↓NP	subject
10.1	VB↑VP↓NP	object
7.9	VB↑VP↑VP↑S↓NP	subject (embedded VP)
4.1	VB↑VP↓ADVP	adverbial adjunct
3.0	NN↑NP↑NP↓PP	prepositional complement of noun
1.7	VB↑VP↓PRT	adverbial particle
1.6	VB↑VP↑VP↑VP↑S↓NP	subject (embedded VP)
14.2		no matching parse constituent
31.4	Other	

Exercise: for each pattern, find an example phrase/sentence; e.g., *$VB_{up}VP_{dn}PP$* in *eats with spoon*

Feature-3: Position

- Whether the constituent to be labeled occurs before or after the predicate defining the semantic frame
- This feature is highly correlated with grammatical function, since subjects will generally appear before a verb, and objects after (for English)
- E.g.- *The boys ate the bananas*
  - *Predicate- “ate”, NP1- “The boys” BEFORE predicate, NP2-”the bananas” AFTER predicate*

Feature-4: Voice

- The distinction between active and passive verbs- Important for connection between semantic role and grammatical function
- Direct objects of active verbs often correspond in semantic role to subjects of passive verbs
  - *Ram saw Shyam; Shyam was seen by Ram*
  - *agt(see, Ram), obj(see, Shyam) in both cases*
- Classify verbs as active or passive by thru 10 passive identifying patterns- passive auxiliary (some form of *to be* or *to get*) and a past participle (*was seen)*

Feature-5: Head Word

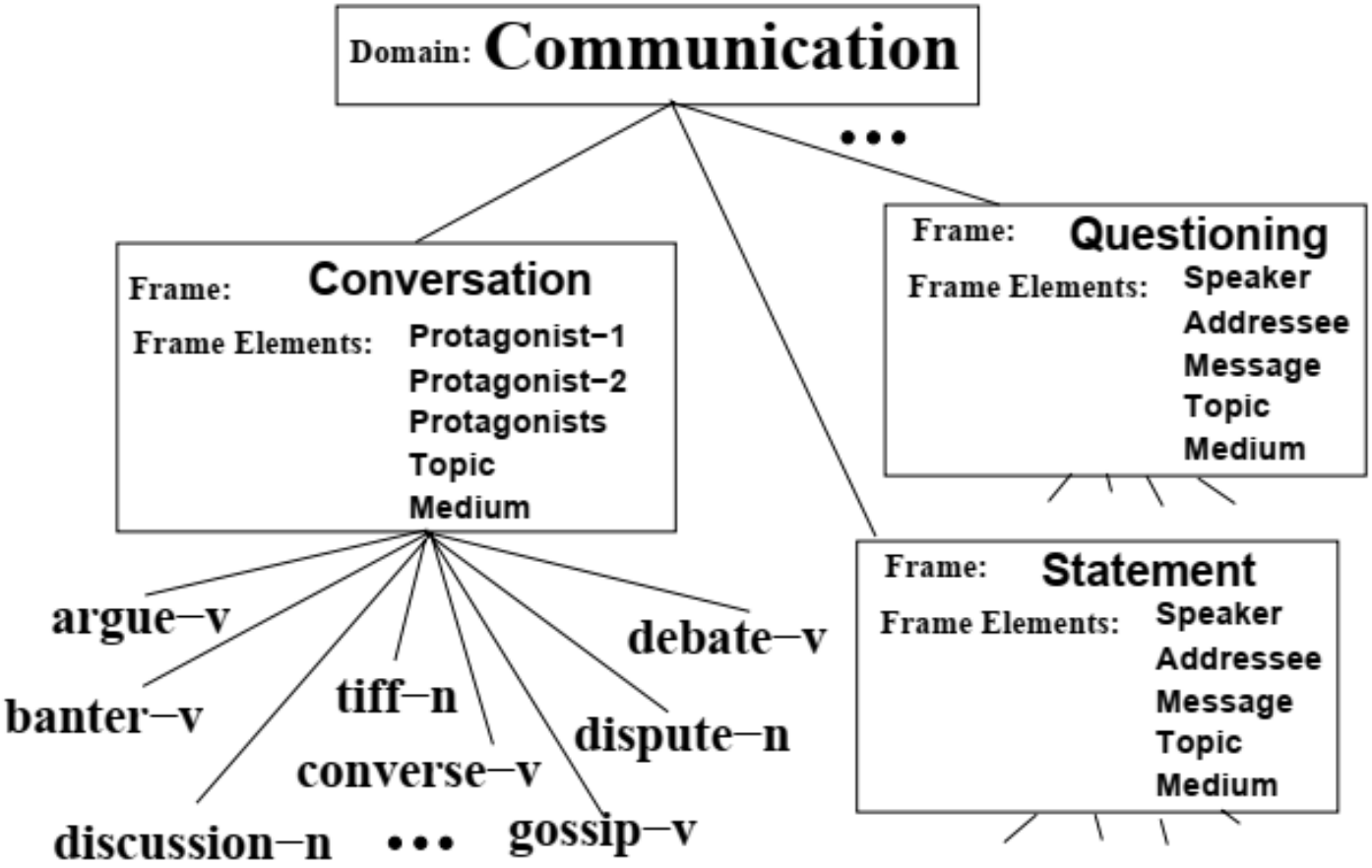
- lexical dependencies extremely important in labeling semantic roles
- Head words of noun phrases can be used to express selectional restrictions (called "akangksha" and "yogyata" in Indian linguistics) on the semantic types of role fillers
  - Selectional Restriction: suitability of being an argument, e.g., "eat"→ agt: animate, obj: edible
- E.g., in a communication frame, noun phrases headed by *Bill*, *brother*, or *he* are more likely to be the SPEAKER, while those headed by *proposal*, *story*, or *question* are more likely to

Preparing the data

- one-tenth of the annotated sentences for each target word were reserved as a test set, and another one-tenth were set aside as a tuning set
- Average number of sentences per target word is only 34, and the number of sentences per frame is 732

(e.g., target word: *speak*, candidate frame: conversation; sentence: *Ram spoke to Shyam)*

Remember: “Communication”



Modeling

$$P(r|h, pt, \mathbf{gov}, position, voice, t)$$

$$P(r|h, pt, gov, position, voice, t) = \frac{\#(r, h, pt, \mathbf{gov}, position, voice, t)}{\#(h, pt, \mathbf{gov}, position, voice, t)}$$

Semantic Role Labeling with Neural Network Factors

Nicholas FitzGerald, Oscar Täckström, Kuzman Ganchev, Dipanjan Das, EMNLP 2015

Problem Statement and the Challenge

- Semantic role labeling (SRL) is the task of identifying the semantic arguments of a predicate and
  labeling them with their semantic roles.
- A key challenge in this task is sparsity of labeled data
- A given predicate-role instance may only occur a handful of times in the training set.

Contribution

- “In this paper, we present a new model for SRL that embeds candidate arguments and semantic roles (in context of a predicate frame) in a shared vector space.”
- “A neural network is learned to capture correlations of the respective embedding dimensions to create argument and role representations.”

Some issues related to Embeddings

- AGT, OBJ, INS etc. can be represented as 1-hot vectors of dim N, where N is the #SRLs
- 1-hot cannot capture inherent similarities
  - agt, obj are in some sense similar (arguments of verbs)
- Word vectors can be pretrained embeddings
- BUT: how do we get vector for break:02 instead of break
- Have to use masked language model like BERT
- General embedding of “break” will be different from “break:02”

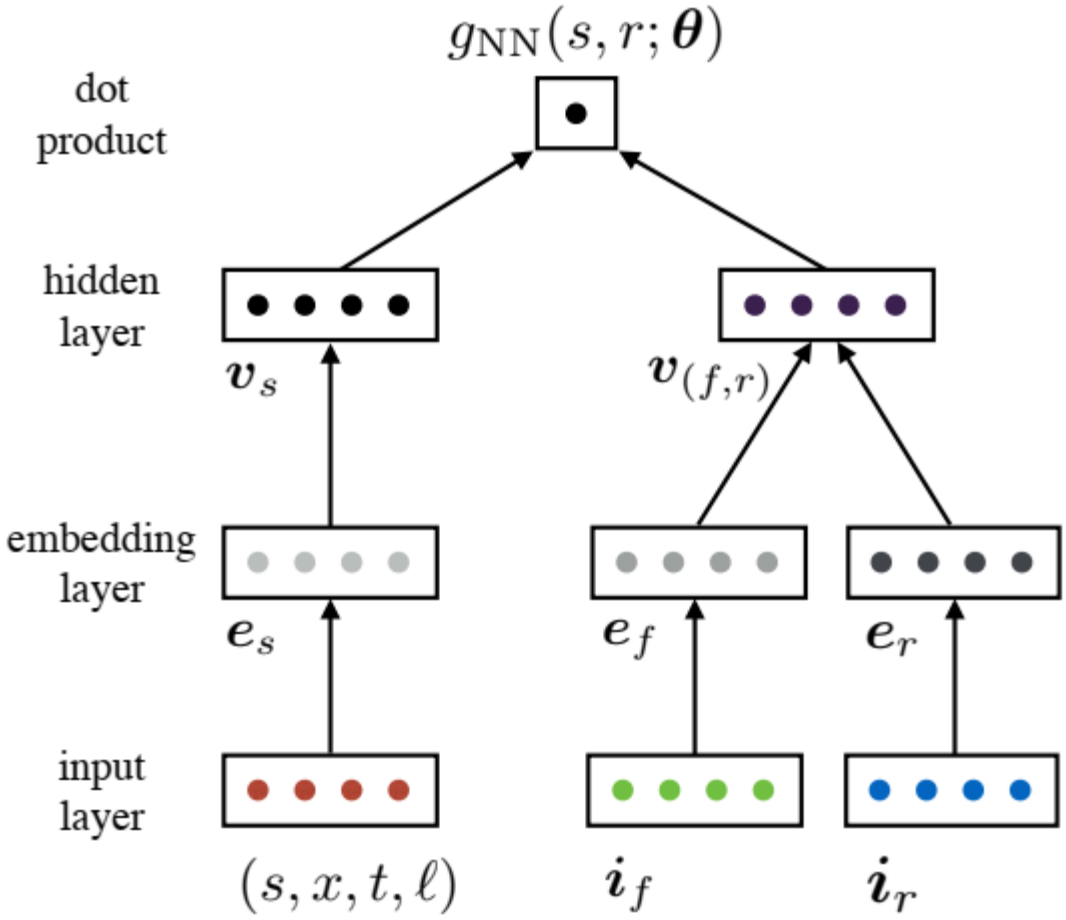
The crux of the work

- The similarity of these two representations, as measured by their dot product, is used to score possible roles for candidate arguments within a graphical model.
- This graphical model jointly models the assignment of semantic roles to all arguments of a predicate, subject to structural linguistic constraints.

SRL Data Sets

- CoNLL 2005 and 2012 data annotated with PropBank conventions
- FrameNet 1.5 data
- Dependency-based CoNLL 2009 shared tas

Neural Net Model



Features

- first word of s
- last word of s
- head word of s
- bag of words in s
- cluster of s's head
- t's children words
- tag of the first word of s
- tag of the last word of s
- tag of the head word of s
- bag of tags in s
- linear *distance* of s from t
- word cluster of s's head
- dependency path between s's head and t
- subcategorization frame of s
- *position* of s w.r.t. t (*before, after, overlap or same*)
- predicate use voice (*active*, *passive*, or *unknown*)
- whether the subject of t is missing (*missingsubj*)
- *position* of s w.r.t. t (*before, after, overlap or same*)
- word, tag, dependency label and cluster of the words immediately to the left and right of s

Governing Equations

Let $\mathcal{R}^{|\mathcal{S}|}$ denote the set of all possible assignments of semantic roles to argument spans (s_i, r_i) for $s_i \in \mathcal{S}$ that satisfy the constraints. Given a potential function $g(s, r) \triangleq g(s, r; \boldsymbol{\theta})$, the probability of a joint assignment $\boldsymbol{r} \in \mathcal{R}^{|\mathcal{S}|}$, subject to the constraints, is given by

$$p(\boldsymbol{r} \mid x, t, \ell, f) = \exp\left(\sum_{s_i \in \mathcal{S}} g(s_i, r_i) - A(\mathcal{S})\right), \tag{1}$$

where the log-partition function $A(\mathcal{S})$ sums over all satisfying joint role assignments:

$$A(\mathcal{S}) = \log \sum_{\boldsymbol{r}' \in \mathcal{R}^{|\mathcal{S}|}} \exp\left(\sum_{s_i \in \mathcal{S}} g(s_i, r_i')\right). \tag{2}$$

To be continued…

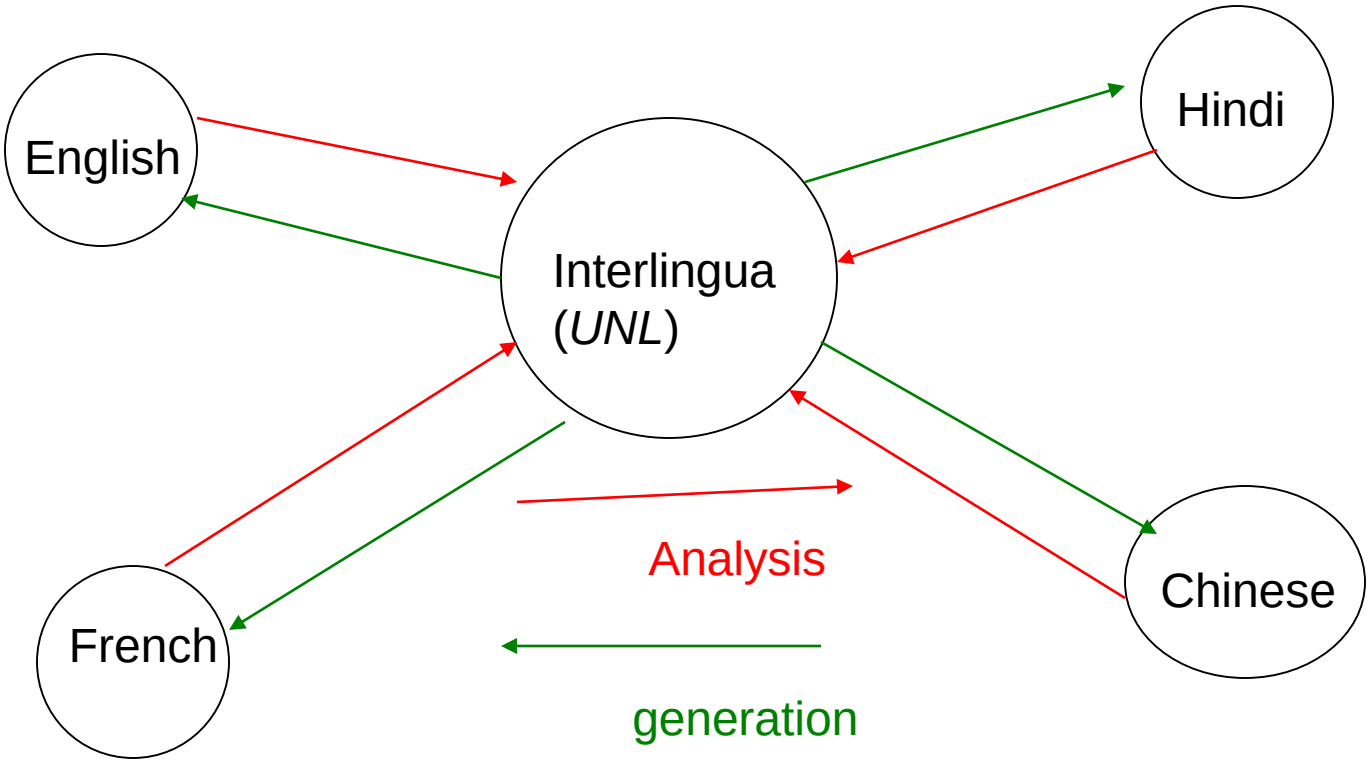
Universal Networking Language: Foundations and Applications

Introduction

Motivation

- Extraction of *semantics, i.e., deep meaning* is important for many applications.
  - Machine Translation, Meaning-based IR, CLIR
- Robust, scalable & efficient methods of knowledge extraction required
- Machine Translation and Cross Lingual IR: a need of the hour for crossing language barrier

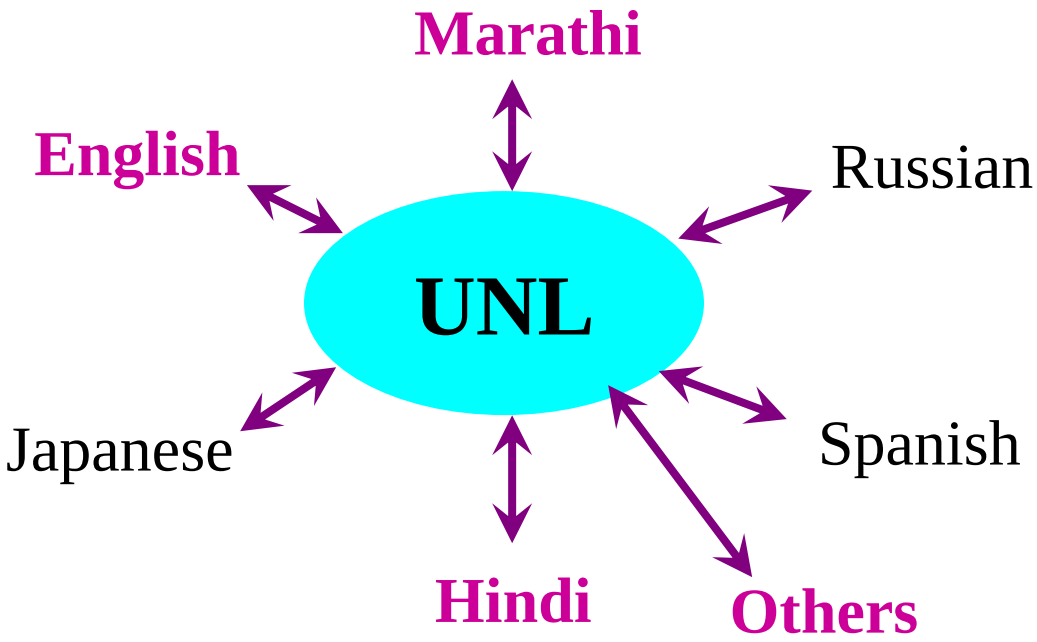
Interlingua: *a vehicle for machine translation*



UNL: a United Nations project

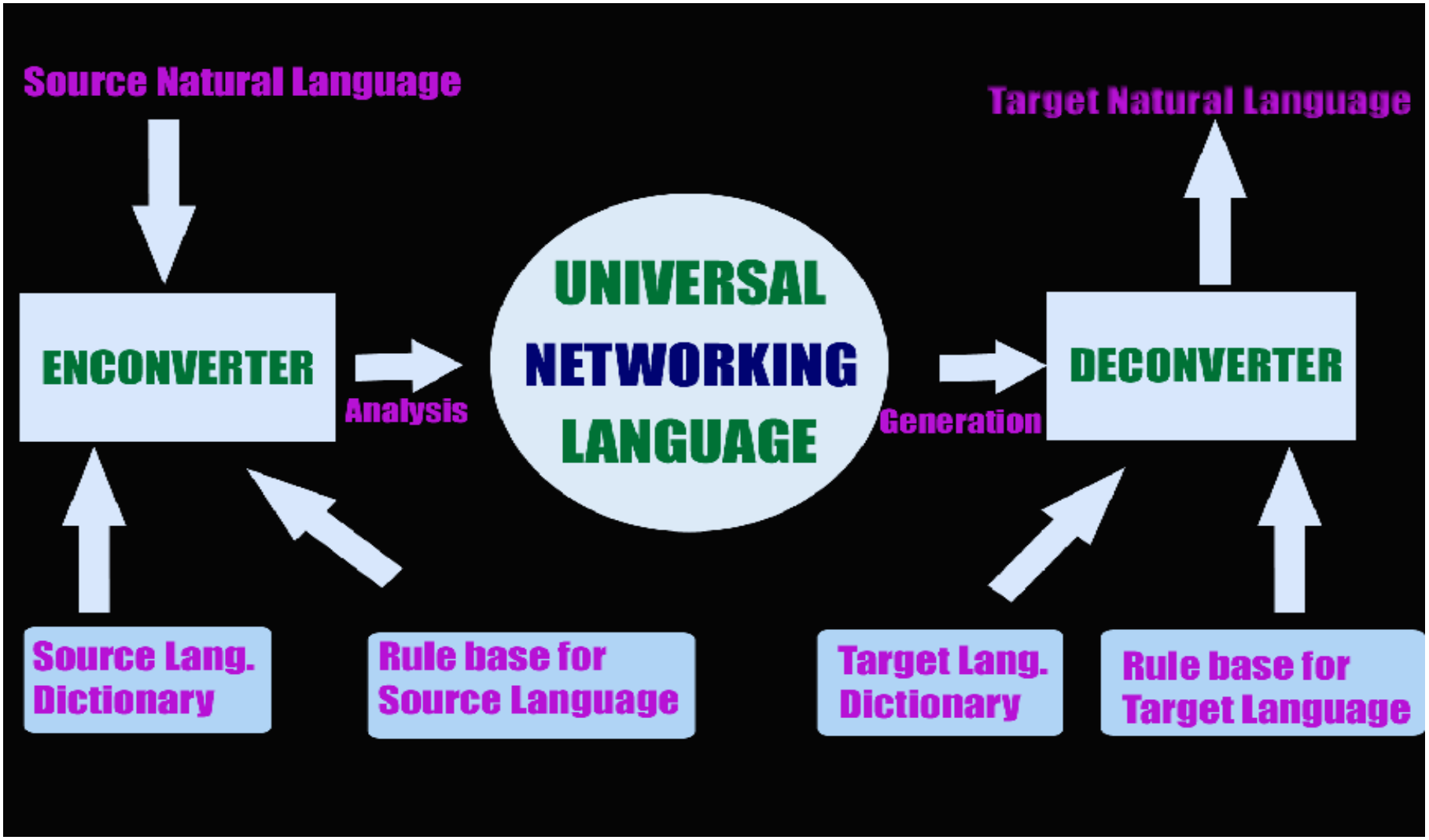
- Started in 1996
- 10 year program
- 15 research groups across continents
- First goal: generators
- Next goal: analysers (needs solving various ambiguity problems)
- Current active language groups
  - UNL_French (GETA-CLIPS, IMAG)
  - UNL_Hindi (IIT Bombay with additional work on UNL_English)
  - UNL_Italian (Univ. of Pisa)
  - UNL_Portugese (Univ of Sao Paolo, Brazil)
  - UNL_Russian (Institute of Linguistics, Moscow)
  - UNL_Spanish (UPM, Madrid)

World-wide
Universal Networking Language (UNL) Project

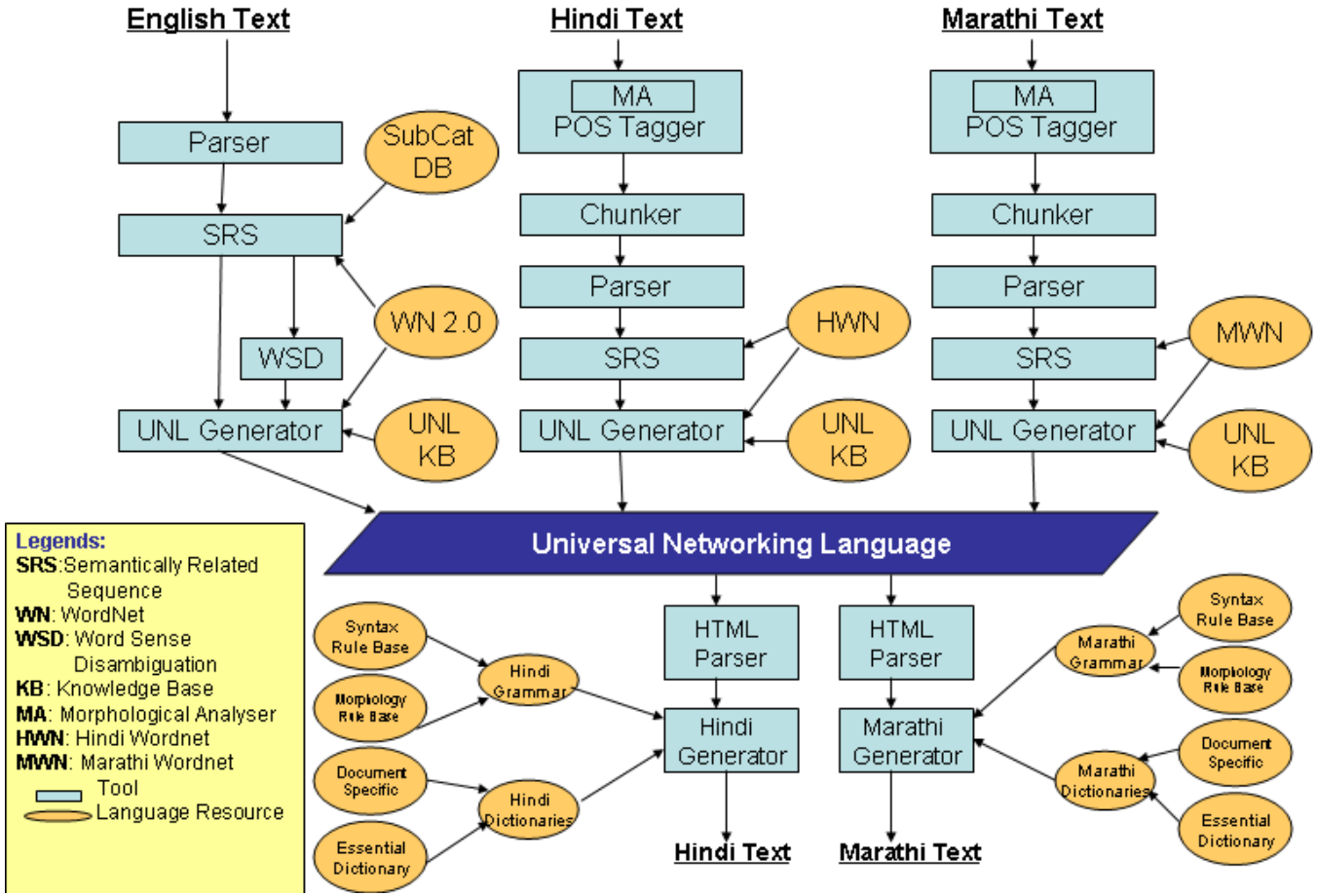


- Language independent meaning representation.

The UNL MT System: an Overview



NLP@IITB


English Text
Hindi Text
Marathi Text
Parser
SubCat DB
SRS
WN 2.0
WSD
UNL Generator
UNL KB
MA
POS Tagger
Chunker
Parser
HWN
SRS
UNL Generator
UNL KB
MA
POS Tagger
Chunker
Parser
MWN
SRS
UNL Generator
UNL KB
Universal Networking Language
Legends:
SRS:Semantically Related Sequence
WN: WordNet
WSD: Word Sense Disambiguation
KB: Knowledge Base
MA: Morphological Analyser
HWN: Hindi Wordnet
MWN: Marathi Wordnet
Tool
Language Resource
Syntax Rule Base
Morphology Rule Base
Document Specific
Essential Dictionary
Hindi Grammar
Hindi Dictionaries
HTML Parser
Hindi Generator
Hindi Text
HTML Parser
Marathi Generator
Marathi Text
Syntax Rule Base
Marathi Grammar
Morphology Rule Base
Document Specific
Marathi Dictionaries
Essential Dictionary

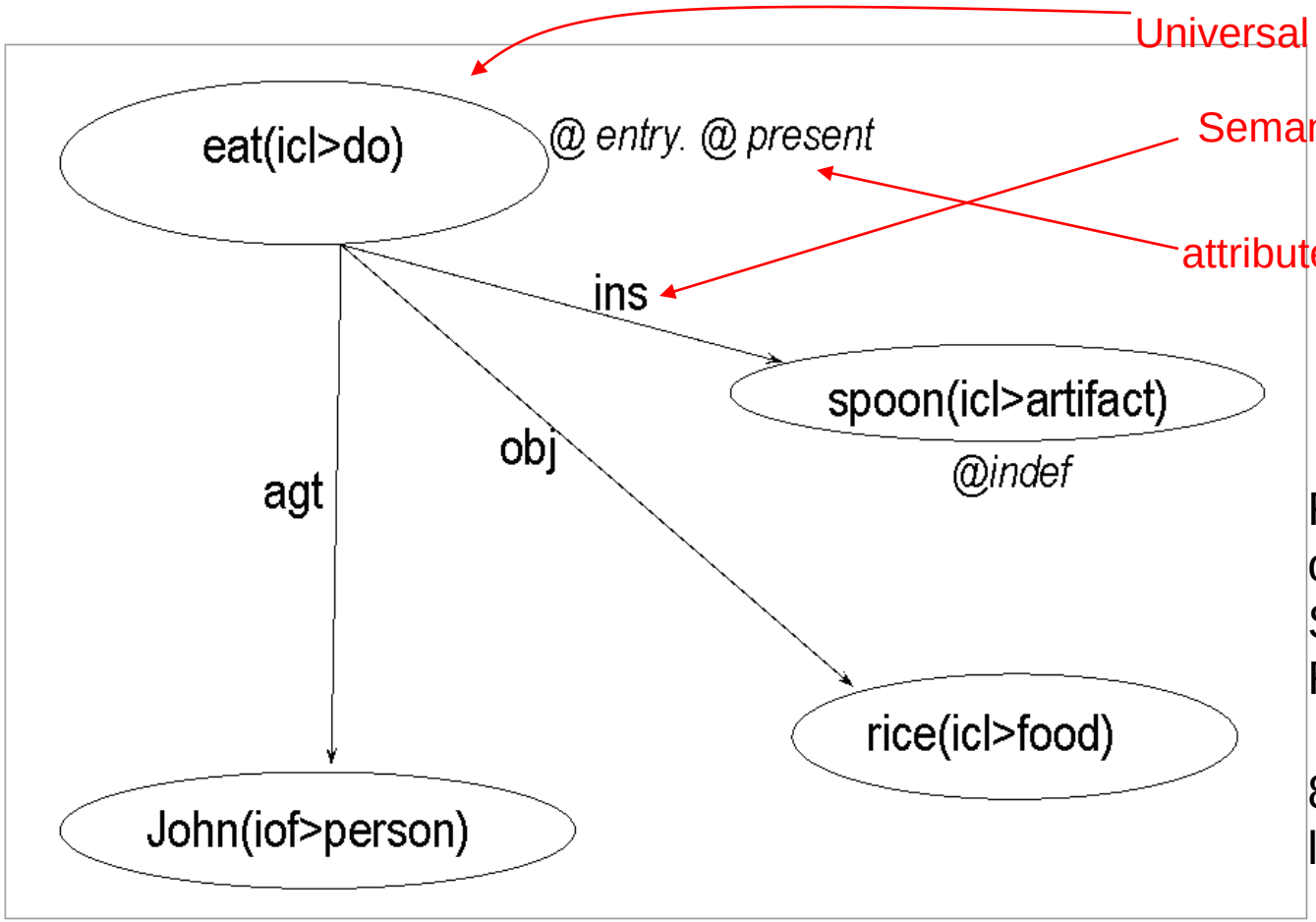
Roadmap and timemap

- UNL Foundations
  - Semantic Relations (0.5 hr)
  - Universal Words (0.5 hr)
  - Attributes (0.25 hr)
  - How to write UNL expressions (0.25 hr)
- UNL Applications
  - Machine Translation (1.5 hr)
  - Search (0.5 hr)
  - Text Entailment (0.5 hr)

The UNL System

UNL represents knowledge:

John eats rice with a spoon



Repository of 42 Semantic Relations and 84 attribute labels

Sentence embeddings

Deepa claimed that she had composed a poem.

[UNL]

agt(claim.@entry.@past, Deepa)

obj(claim.@entry.@past, :01)

agt:01(compose.@past.@entry.@complete, she)

obj:01(compose.@past.@entry.@complete,
poem.@indef)

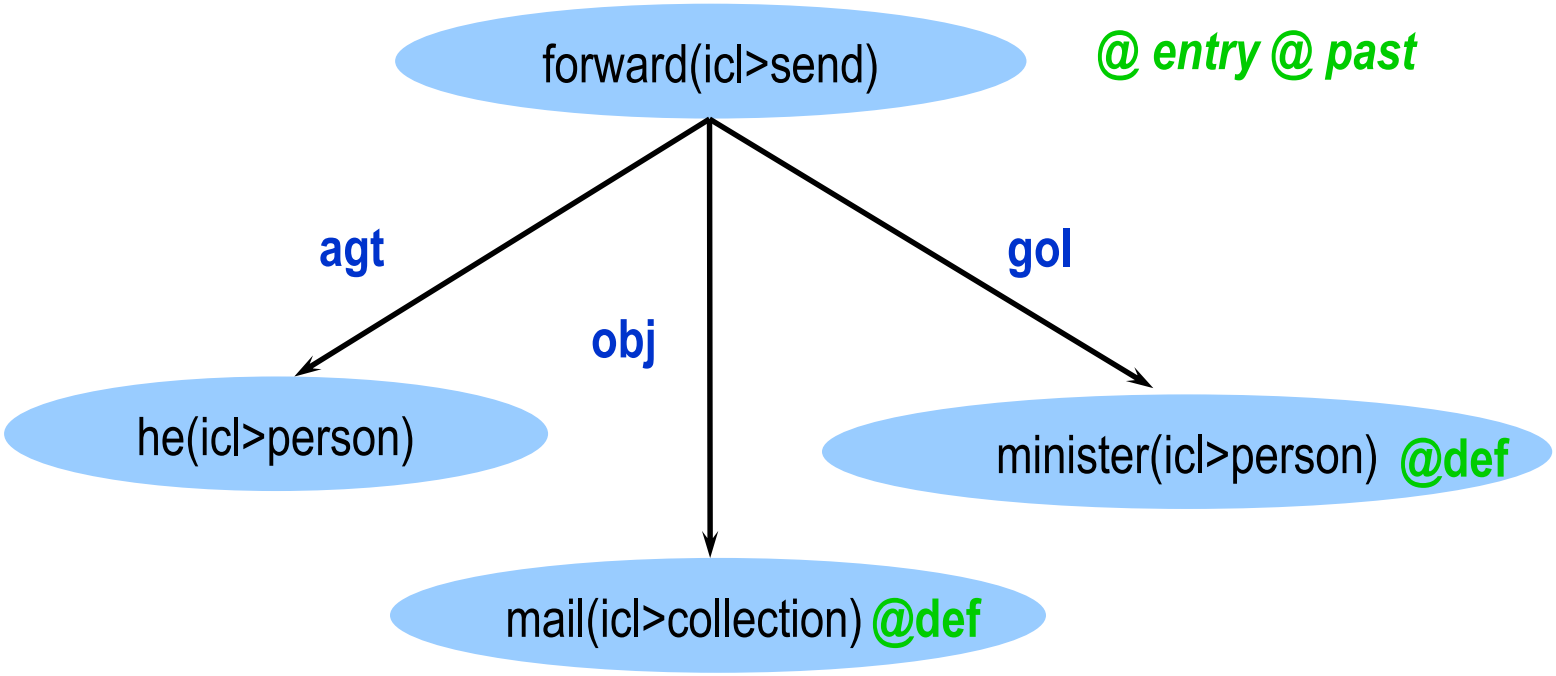
[\UNL]

Universal Networking Language

- Universal Words (UWs)
- Relations
- Attributes
- Knowledge Base

UNL Graph

He forwarded the mail to the minister.



UNL Expression

agt (forward(icl>send).@ ***entry*** @ ***past***, he(icl>person))

obj (forward(icl>send).@ ***entry*** @ ***past***, minister(icl>person))

gol (forward(icl>send).@ ***entry*** @ ***past***, mail(icl>collection). @**def**)

What is a Universal Word (UW)?

- Words of UNL
- Constitute the UNL vocabulary, the syntactic-semantic units to form UNL expressions
- A UW represents a concept
  - Basic UW (an English word/compound word/phrase with no restrictions or Constraint List)
  - Restricted UW (with a Constraint List)
- Examples:

  “crane(icl>device)”

  “crane(icl>bird)”

The Lexicon

Format of the dictionary entry

[headword] {} “Universal word“ (Attribute list);

e.g., [minister] {} “minister(icl>person)” (N,ANIMT,PHSCL,PRSN);

- Head word
- Universal word
- Attributes
  - Morphological - Pl(plural), V_ed(past tense form)
  - Syntactic - V(verb),VOA(verb of action)
  - Semantic - ANIMT(animate), PLACE, TIME

The Lexicon (cntd)

He forwarded the mail to the minister.

Content words:

[forward] {} “forward(icl>send)” (V,VOA) <E,0,0>;

[mail] {} “mail(icl>message)” (N,PHSCL,INANI) <E,0,0>;

[minister] {} “minister(icl>person)” (N,ANIMT,PHSCL,PRSN) <E,0,0>;

Headword Universal Word Attributes

The Lexicon (cntd)

He forwarded the mail to the minister.

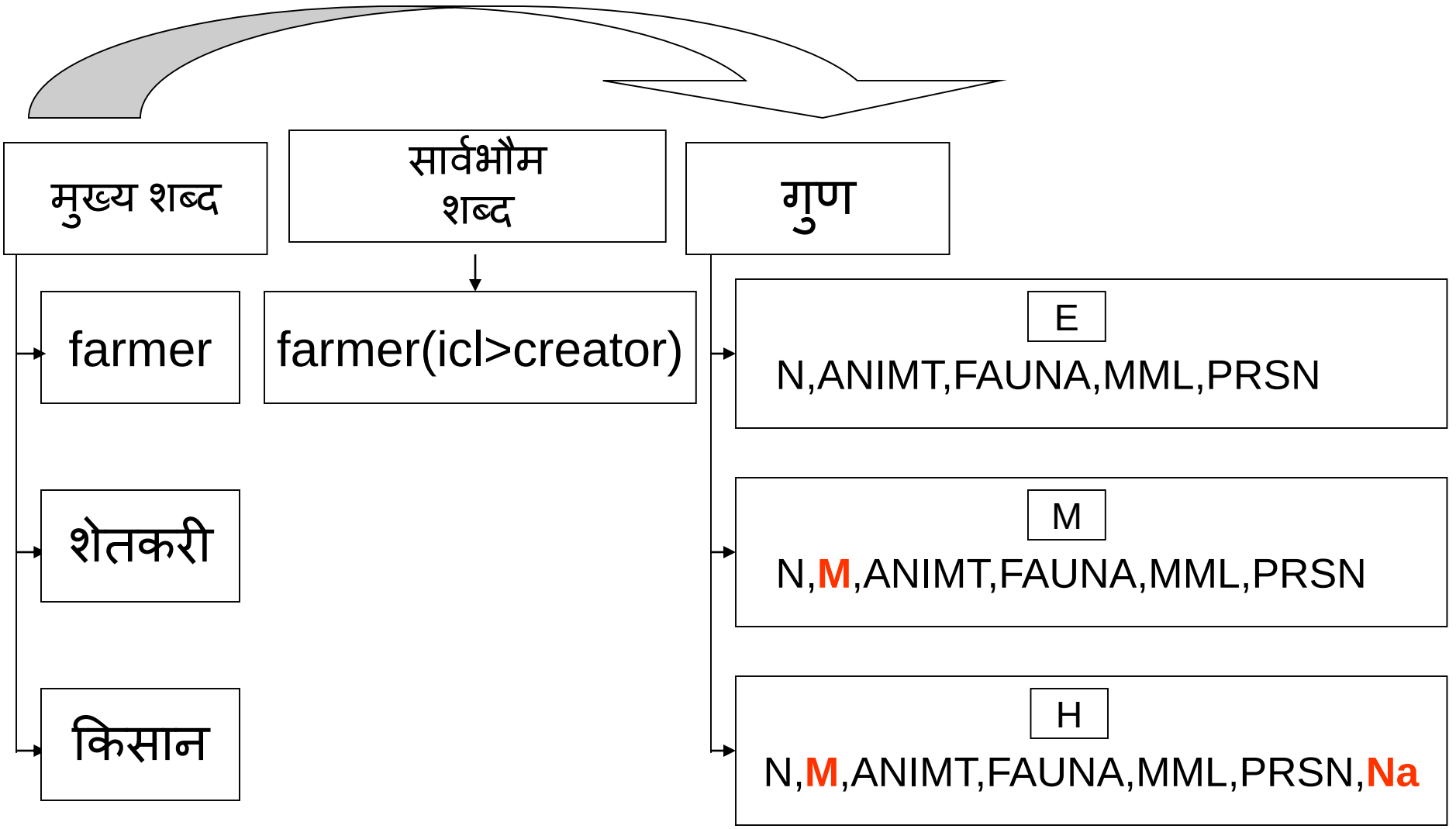
function words:

[he] {} “he” (PRON,SUB,SING,3RD) <E,0,0>;

[the] {} “the” (ART,THE) <E,0,0>;

[to] {} “to” (PRE,#TO) <E,0,0>;

Hindi example: संज्ञा का उदाहरण १/२



The Features of a UW

- Every concept existing in any language must correspond to a UW
- The constraint list should be as small as necessary to disambiguate the headword
- Every UW should be defined in the UNL Knowledge-Base

Restricted UWs

- Examples
  - *He will hold office until the spring of next year.*
  - *The spring was broken.*
- Restricted UWs, which are Headwords with a constraint list, for example:

  “spring(icl>season)”

  “spring(icl>device)”

  “spring(icl>jump)”

  “spring(icl>fountain)”

How to create UWs?

- Pick up a concept
  - the concept of "***crane***"
    as "***a device for lifting heavy loads"***
    or
    as "***a long-legged bird that wade in water in search of food***"
- Choose an English word for the concept.
  - In the case for "***crane***", since it is a word of English, the corresponding word should be '***crane***'
- Choose a constraint list for the word.
  - **[ ] 'crane(icl>device)'**
  - **[ ] 'crane(icl>bird)'**

UW construction procedure

Acknowledgement: Igor Boguslavsky, Russian LC,
U++ Meeting, July 2007

UW construction procedure

- Input: Local word in L in context C.
- Output: UW for L

- Resources:
  - UW dictionary
  - Local L-English dictionary (LED)
  - English dictionary (ED)

Major steps

- Headword selection
- Construction of ontological constraints
- Construction of semantic constraints
- Construction of argument constraints

I. Headword selection

1. Consult the UW dictionary and check whether it contains pairs of the types: [L]–UW

Yes: 2. Check if among UWs in these pairs there is one that suits context C.

Yes: Finish

No: Goto (3)

No: Goto (3)

Example: English-UW dictionary

[spring] “spring(icl>season)”

[spring] “spring(icl>device)”

[spring] “spring(icl>jump)”

[spring] “spring(icl>fountain)”

I. Headword selection (contd.)

3. Consult LED for L. Find the meaning of L that suits context C. Check how many translations has this meaning of L in the LED:

   One:

4. Make this translation headword for L (HW) and go to 2

   More than one:

5. Check if among these translations there is one more general and neutral than others

      Yes:

6. Take it and go to 4

      No:

7. Check if these translations are synonymous

      Yes: Take the first of them and go to 4

            No: Go to 4 with each translation

Example: Hindi word *ghar*

- *ghar- house*
  - *usne garmii me ghar kii marammat kii*
  - *he renovated the house in the summer*
- *ghar- home*
  - *office ke baad ghar louto*
  - *return home after office*
- *Ghar- family*
  - *bade ghar kii betii*
  - *girl from a renowned family*

Example: *ghar (cntd)*

- *ghar- own country*
  - *bahut saal bidesh me kaam karke ghar louta aayaa*
  - *returned home after working abroad for many years*
- *Ghar- astrological position*
  - *ashtam ghar par budh hai*
  - *Mercury in in the eighth house*

House in English Wordnet

- 1. (1029) house -- (a dwelling that serves as living quarters for one or more families; "he has a house on Cape Cod"; "she felt she had to get out of the house")
- 3. (51) house -- (a building in which something is sheltered or located; "they had a large carriage house")
- 4. (39) family, household, house, home, menage -- (a social unit living together; "he moved his family to Virginia"; "It was a good Christian household“;)

House in English Wordnet

- 7. (13) house -- (aristocratic family line; "the House of York")
- 11. sign of the zodiac, star sign, sign, mansion, house, planetary house -- ((astrology) one of 12 equal areas into which the zodiac is divided)

II. Construction of ontological constraints

1. Check the category of HW

Noun

2. Check whether HW is an instance of concept X

Yes:

3. Make constraint (iof>X).

No:

4. Take the closest hypernym X of HW and make constraint (icl>X>thing).

II. Construction of ontological constraints

Verb

5. Check to what semantic class HW belongs

Action

6. Make (icl>do).

Process

7. Make (icl>occur).

State

8. Make (icl>be).

II. Construction of ontological constraints

Adjective

9. Make (icl>adj)

Adverb, preposition, conjunction

10. Make (icl>how)

Semantic and Argument constraints: *need linguistic background*