

CS626 : NLP, Speech and the Web

Avishek Dan
CSE Dept.,
IIT Bombay

Lecture 16: Semantics; SRL; Universal Networking
Language

20th September, 2012

Overview

- Semantics
- A Perspective
- Signifier-denotata dichotomy
- UNL
- Components of UNL
 - Universal Words
 - Relations
 - Attributes
- UNL representation
- Issues to handle
- Summary

Semantics

According to Wikipedia

- **Semantics** (from [Greek](#) *sēmantiká*, neuter plural of *sēmantikós*) is the study of [meaning](#).
- It typically focuses on the relation between *signifiers*, such as [words](#), [phrases](#), [signs](#) and [symbols](#), and what they stand for, their [denotata](#).
- **Computational semantics** is the study of how to automate the process of constructing and reasoning with [meaning representations](#) of [natural language](#) expressions.

A Perspective

Discourse

Pragmatics

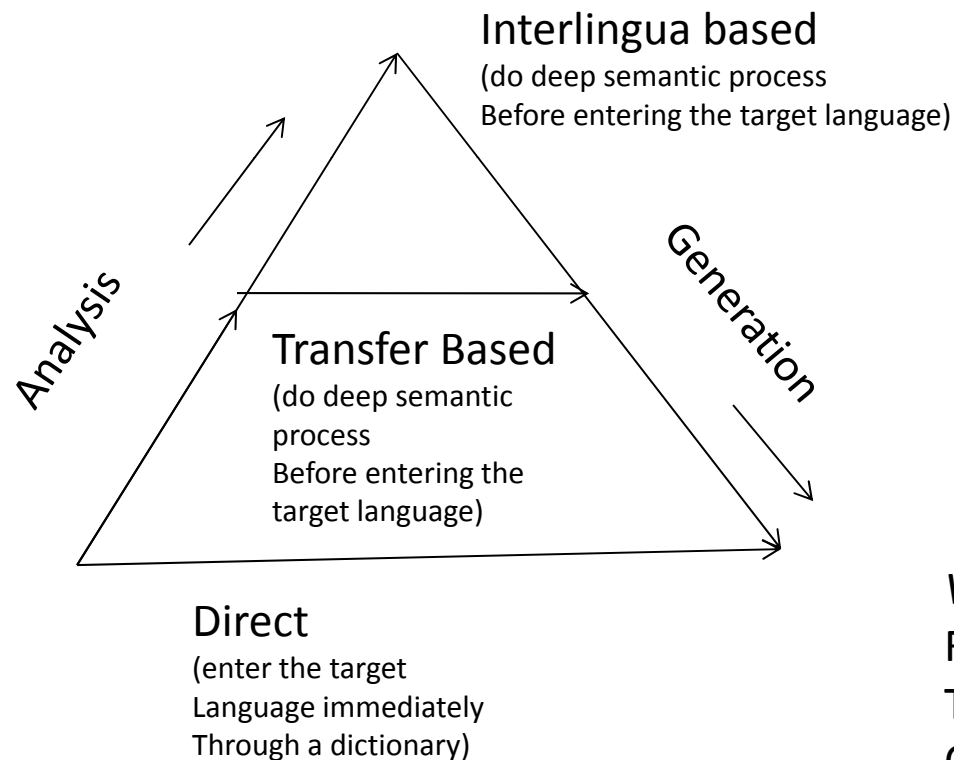
Semantics

Syntax

Lexicon

Morphology

Vaquious Triangle



Vaquious: an eminent
French Machine
Translation Researcher-
Originally a Physicist

Semantic Role Labeling

- For each clause, determine the semantic role played by each noun phrase that is an argument to the verb.

agent patient source destination instrument

– John drove Mary from Austin to Dallas in his Toyota Prius.

– The hammer broke the window.

- Also referred to as “shallow semantic parsing”

Semantic Roles

- A variety of semantic role labels have been proposed, common ones are:
 - Agent: Actor of an action
 - Patient: Entity affected by the action
 - Instrument: Tool used in performing action.
 - Beneficiary: Entity for whom action is performed
 - Source: Origin of the affected entity
 - Destination: Destination of the affected entity

Use of Semantic Roles

- Semantic roles are useful for various tasks.
- Question Answering
 - “Who” questions usually use Agents
 - “What” question usually use Patients
 - “How” and “with what” questions usually use Instruments
 - “Where” questions frequently use Sources and Destinations.
 - “For whom” questions usually use Beneficiaries
 - “To whom” questions usually use Destinations
- Machine Translation Generation
 - Semantic roles are usually expressed using particular, distinct syntactic constructions in different languages.

SRL and Syntactic Cues

- Frequently semantic role is indicated by a particular syntactic position (e.g. object of a particular preposition).
 - Agent: subject
 - Patient: direct object
 - Instrument: object of “with” PP
 - Beneficiary: object of “for” PP
 - Source: object of “from” PP
 - Destination: object of “to” PP
- However, these are preferences at best:
 - The hammer hit the window.
 - The book was given to Mary by John.
 - John went to the movie with Mary.
 - John bought the car for \$21K.
 - John went to work by bus.

Selectional Restrictions

- **Selectional restrictions** are constraints that certain verbs place on the filler of certain semantic roles.
 - Agents should be animate
 - Beneficiaries should be animate
 - Instruments should be tools
 - Patients of “eat” should be edible
 - Sources and Destinations of “go” should be places.
 - Sources and Destinations of “give” should be animate.
- Taxonomic abstraction hierarchies or ontologies (e.g. hypernym links in WordNet) can be used to determine if such constraints are met.
 - “John” is a “Human” which is a “Mammal” which is a “Vertebrate” which is an “Animate”

Use of Sectional Restrictions

- Selectional restrictions can help rule in or out certain semantic role assignments.
 - “John bought the car for \$21K”
 - Beneficiaries should be Animate
 - Instrument of a “buy” should be Money
 - “John went to the movie with Mary”
 - Instrument should be Inanimate
 - “John drove Mary to school in the van”
“John drove the van to work with Mary.”
 - Instrument of a “drive” should be a Vehicle

Selectional Restrictions and Syntactic Ambiguity

- Many syntactic ambiguities like PP attachment can be resolved using selectional restrictions.
 - “John ate the spaghetti with meatballs.”
“John ate the spaghetti with chopsticks.”
 - Instruments should be tools
 - Patients of “eat” must be edible
 - “John hit the man with a dog.”
“John hit the man with a hammer.”
 - Instruments should be tools

Selectional Restrictions and Syntactic Ambiguity

- Many syntactic ambiguities like PP attachment can be resolved using selectional restrictions.
 - “John ate the spaghetti with meatballs.”
“John ate the spaghetti with chopsticks.”
 - Instruments should be tools
 - Patients of “eat” must be edible
 - “John hit the man with a dog.”
“John hit the man with a hammer.”
 - Instruments should be tools

Signifier-denotata dichotomy

- Divide between a word and what it stands for
- How many meanings does “present” have in English?

Signifier-denotata dichotomy

- Divide between a word and what it stand for
- How many meanings does “present” have in English?
- According to WordNet: 15

Signifier-denotata dichotomy

- Divide between a word and what it stand for
- How many meanings does “present” have in English?
- According to WordNet : 15
- Some are:
 - (noun) nowadays : वर्तमान
 - (noun) gift : उपहार
 - (adj) in attendance : मौजूद
 - (verb) “what I am doing (presentation)” : प्रस्तुत करना etc.

Signifier-denotata dichotomy

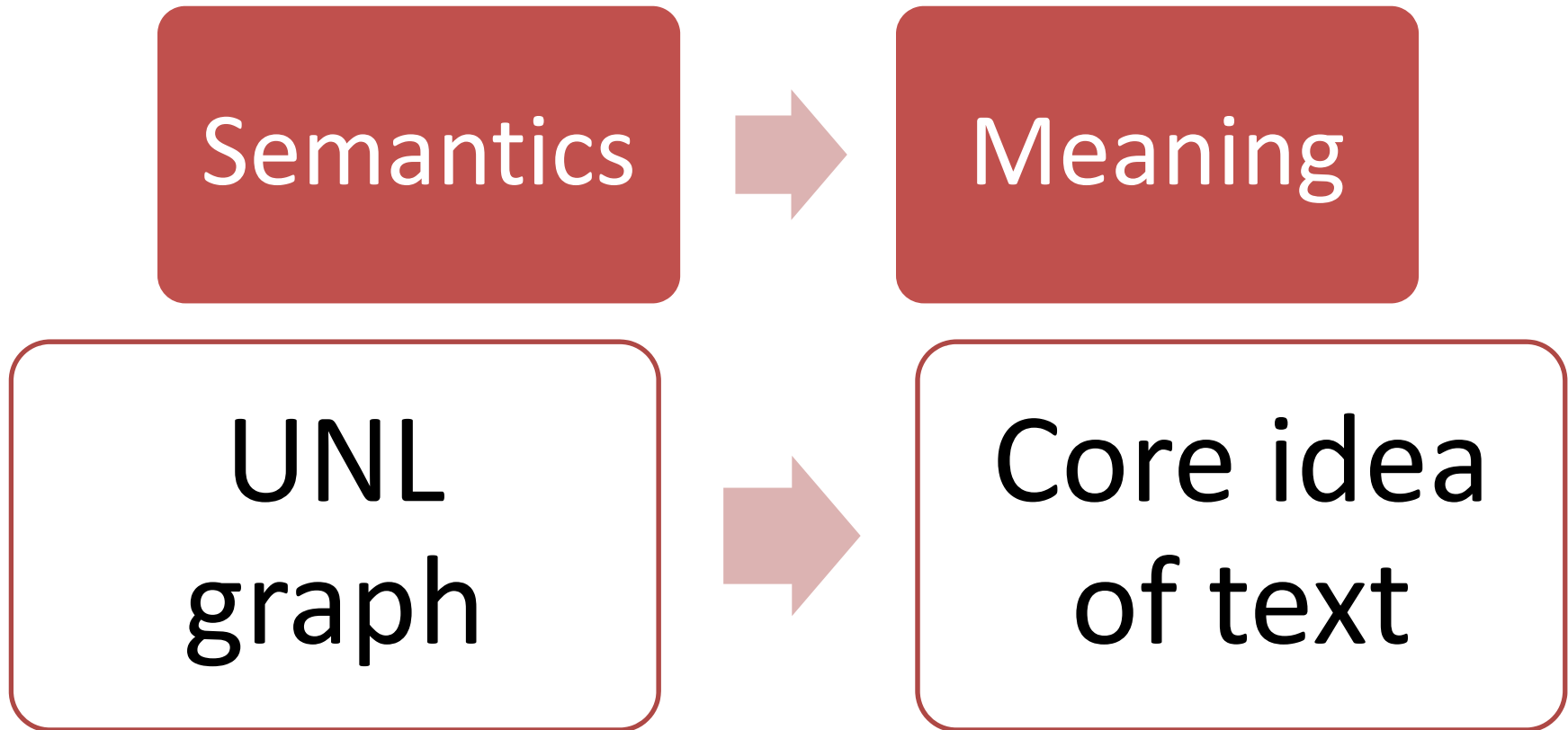
- Interestingly Hindi has different words for all of these
- These 15 meanings represent different ***concepts***
- Using the tool of English the concepts have the same representation (“present”)
- With change in tool (Hindi) the representation is distinct for each of them

Universal Networking Language

- The **Universal Networking Language** (UNL) is an artificial language created to process information across language barriers.
- Features of UNL
 - represent "what was meant" and not "what was said", i.e. provides an **interpretation** rather than a translation of a given utterance
 - represent knowledge in a machine-tractable format
 - UNL is self-sufficient, i.e., should be as semantically complete and saturated as possible
 - UNL is independent from any particular natural language
 - UNL representation is unambiguous

UNL graph

- UNL is a semantic graph



Example

- “John eats rice with a spoon”

- Main verb:

eats



eat (present)

- Who performed this action:

John



A person

- What is the object of the action:

rice



a kind of food

- What is the instrument of the action?

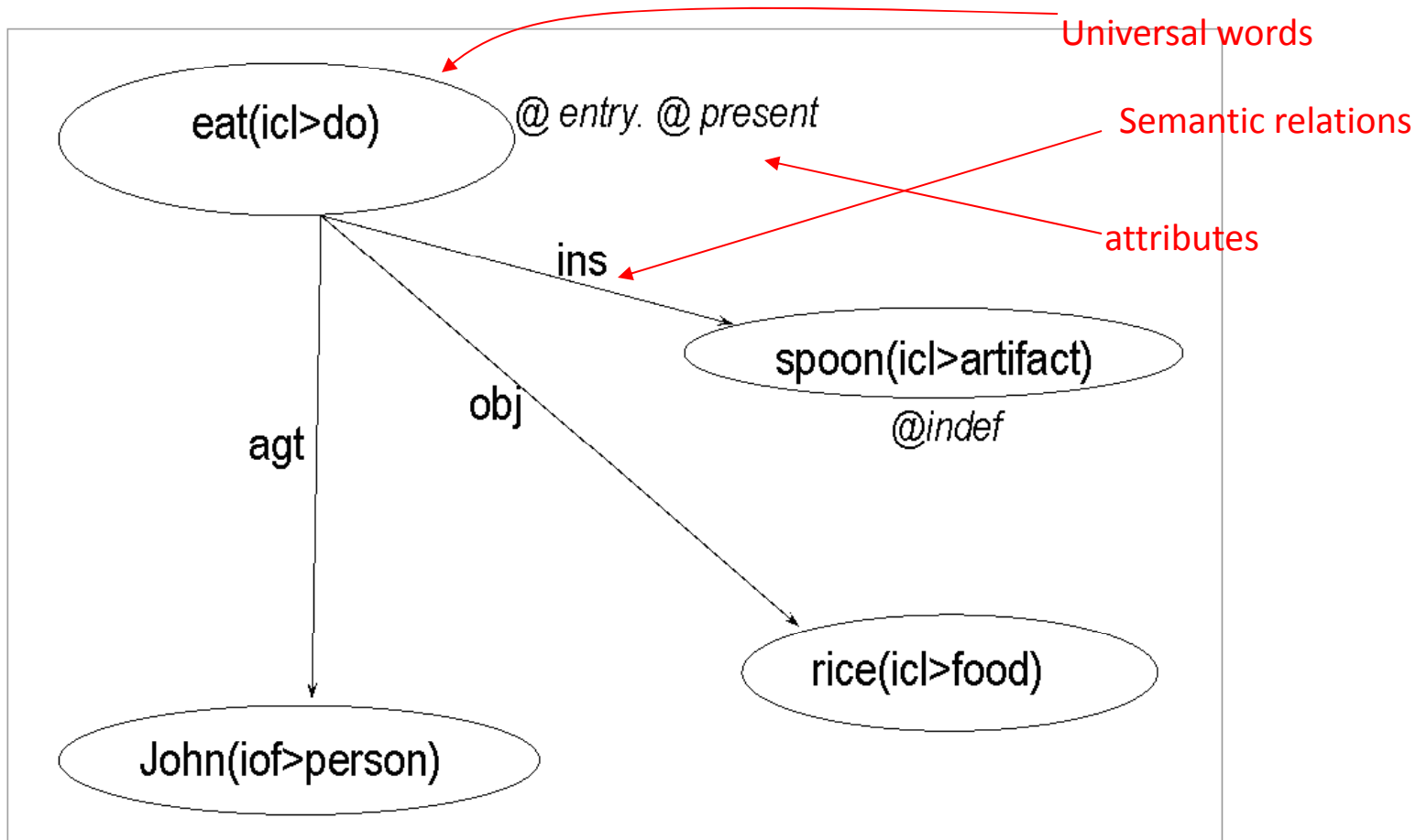
spoon



An artifact

UNL represents **knowledge**:

John eats rice with a spoon



Components of UNL

- **Universal Words**
- Relations
- Attributes

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)
 - Temporary UW
- Examples:
 - “crane(icl>device)”
 - “crane(icl>bird)”

Restricted Universal Words

- Restricted UWs have :

- Headword

- Constraint list

E.g. eat (icl>do)

John (iof>person) [*Named Entity*]

New Delhi (iof>city) [*Named Entity*]

week-long (icl>period) [*Multiword*]

look for (icl>do, obj>thing) [*English equivalent of Spanish word “Buscar”*]

Restricted Universal Words

- Constraint list serve the following purpose:
 - **Ontological** : Give ontological relations (icl(class of), iof(instance of)) with other UWs
 - **Semantic**: Disambiguate the headword
e.g. bank(icl>financial institution)
bank(icl>slope, icl<riverbank)
 - **Argument Frame** : Give argument frame requirements of the UW
e.g. look for (icl>do, **obj>thing**) (look for **something**)

Temporary Universal Words

- Strings of symbols that make part of the text but do not belong to the vocabulary of UNL
- E.g. Numbers like 1234

Components of UNL

- Universal Words
- **Relations**
- Attributes

Relations

- Labelled arcs connecting a node to another node in a UNL graph
- Correspond to two-place semantic predicates holding between two Universal Words
- Used to represent semantic cases or thematic roles (such as agent, object, instrument, etc.) between UWs
- Represented as strings of 3 characters or less
- Relations are organized in a hierarchy where lower nodes subsume upper nodes
 - topmost level is the relation "rel"
 - Next level contains participant (ptp), attribute (aoj), specifier (mod) and adjunct (adj)
- A set of 46 relations specified in UNL (e.g., *agt*, *aoj*, *ben*, *gol*, *obj*, *plc*, *src*, *tim*,...)

Relations

- Some important relations are:

Relation	Description
agt	Agent of the action
obj	Object of the action
plc	Place of the action
tim	Time of the action
pur	Purpose
rsn	Reason
aoj	Attribute of the object
plf	Source place
plt	Destination place

Hierarchy of Relations

rel

[-] ptp (participant)

- [-] agt (agent, cause or natural force)
- [-] cag (co-agent)
- [-] obj (patient)
- [-] cob (co-object)
- [-] ptn (partner)
- [-] ins (instrument)
- [-] ben (beneficiary)
- [-] gol (recipient, addressee)
- [-] aoj (experiencer)

[-] aoj (attribute)

- [-] ant (antonymy)
- [-] cnt (content, theme)
- [-] equ (synonymy)
- [-] fld (field)
- [-] icl (hyponymy, a kind of)
- [-] iof (instance of)
- [-] pof (meronymy, part of)
- [-] nam (name)

[-] mod (specifier)

- [-] pos (possessor)
- [-] qua (quantity)
- [-] frm (origin)
- [-] to (destination)

[-] adj (adjunct)

- [-] and (conjunction)
- [-] or (disjunction)

[-] plc (location)

- [-] plf (initial place)
- [-] plt (final place, direction)
- [-] via (intermediate place, path)
- [-] scn (logical place, scene)

[-] tim (time)

- [-] tmf (initial time)
- [-] tmt (final time)
- [-] dur (duration)
- [-] coo (co-occurrence)
- [-] seq (sequence)

[-] man (manner)

- [-] bas (basis for comparison)
- [-] con (condition)
- [-] gol (final state)
- [-] mat (material)
- [-] met (method)
- [-] per (proportion, rate, distribution)
- [-] pur (purpose)
- [-] rsn (reason)
- [-] src (initial state)

AGT / AOJ / OBJ

- AGT (Agent)

Definition: *Agt defines a thing which initiates an action*

- AOJ (Thing with attribute)

Definition: *Aoj defines a thing which is in a state or has an attribute*

- OBJ (Affected thing)

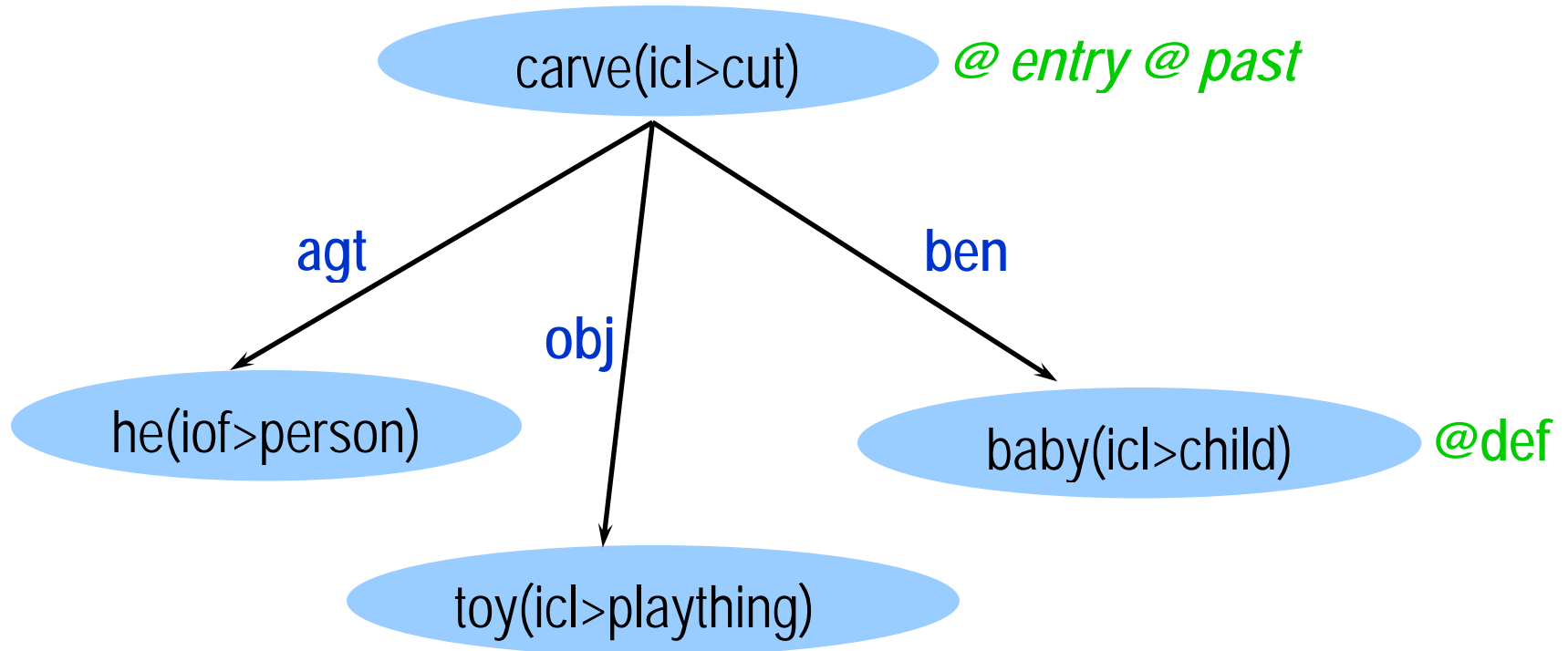
Definition: *Obj defines a thing in focus which is directly affected by an event or state*

Examples

- John broke the window.
agt (break.@entry.@past, John)
- This flower is beautiful.
aoj (beautiful.@entry, flower)
- He blamed John for the accident.
obj (blame.@entry.@past, John)

Example: UNL Graph with agt, obj, ben

He carved a toy for the baby.



GOL / SRC

- GOL (Goal : final state)

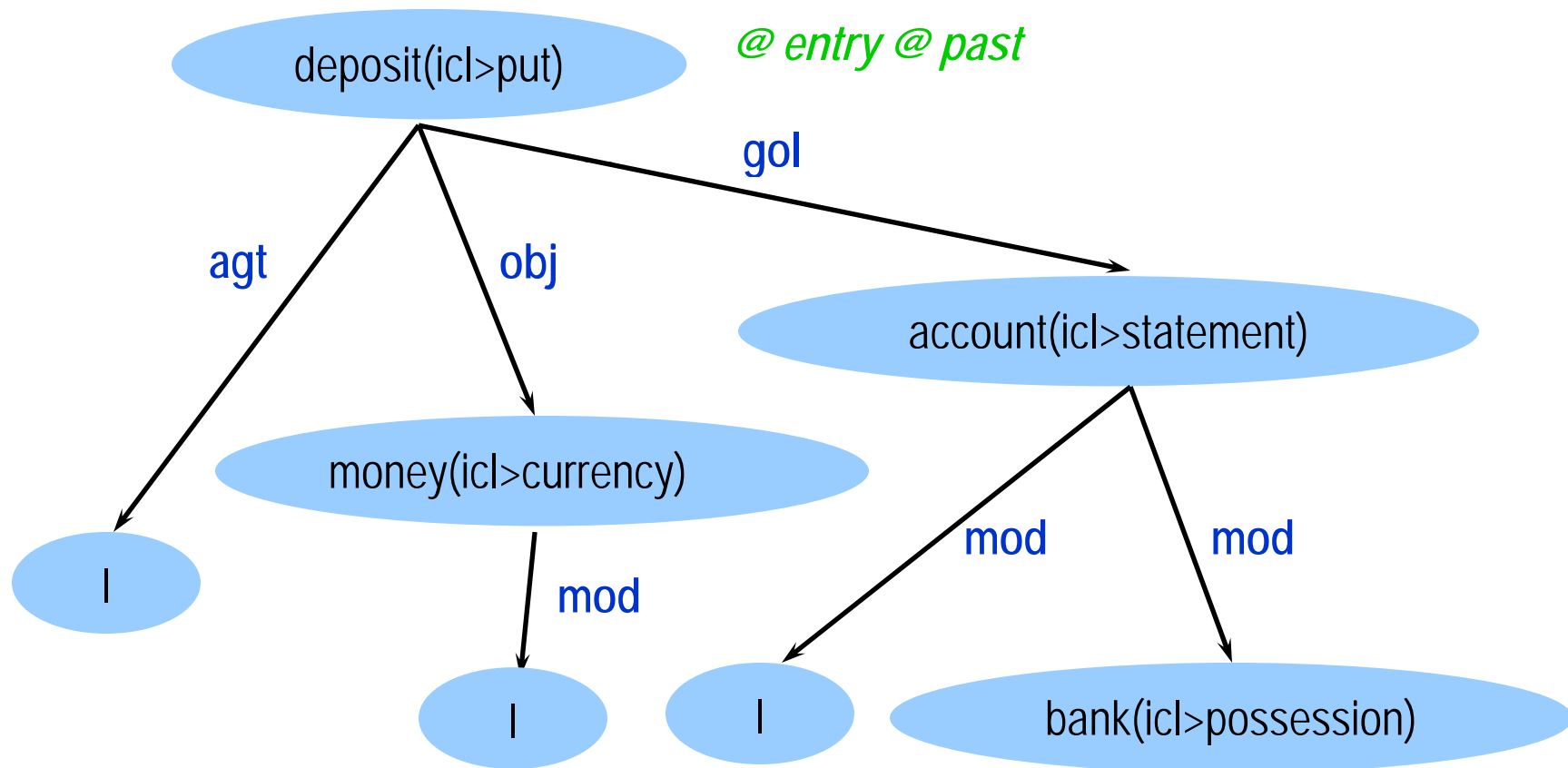
Definition: *Gol defines the final state of an object or the thing finally associated with an object of an event*

- SRC (Source : initial state)

Definition: *Src defines the initial state of object or the thing initially associated with object of an event*

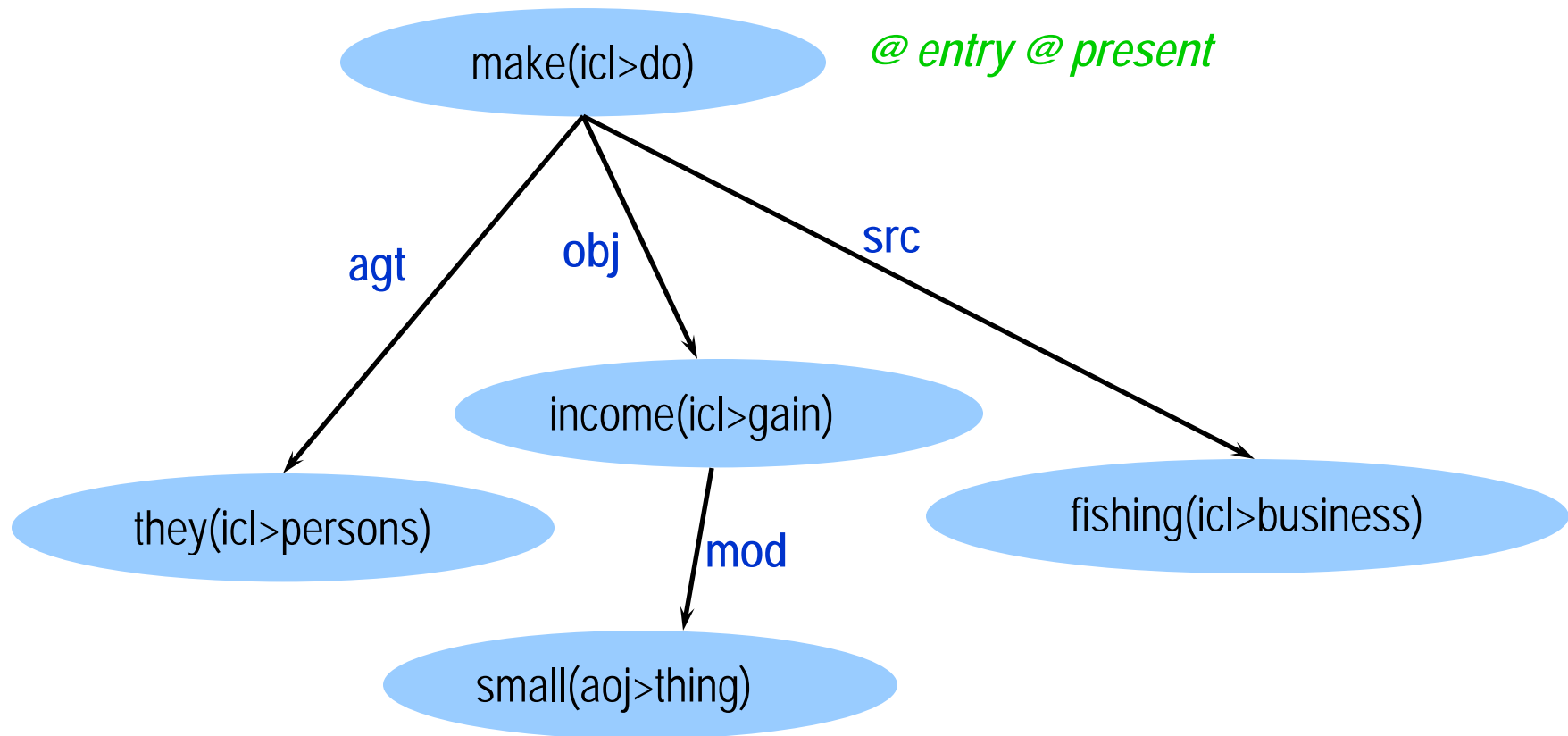
GOL

- I deposited my money in my bank account.



SRC

- They make a small income from fishing.



PUR

- PUR (Purpose or objective)

Definition: *Pur defines the purpose or objectives of the agent of an event or the purpose of a thing exist*

- *This budget is for food.*

pur (food.@entry, budget)
mod (budget, this)

RSN

- RSN (Reason)

Definition: *Rsn defines a reason why an event or a state happens*

- *They selected him for his honesty.*

agt(select(icl>choose).@entry, they)

obj(select(icl>choose) .@entry, he)

rsn (select(icl>choose).@entry, honesty)

TIM

- TIM (Time)

Definition: Tim defines the time an event occurs or a state is true

- *I wake up at noon.*

agt (wake up.@entry, I)

tim (wake up.@entry, noon(icl>time))

PLC

- PLC (Place)

Definition: *Plc defines the place an event occurs or a state is true or a thing exists*

- *Temples are very famous in India.*

aoj (famous.@entry, temple@pl)

man (famous.@entry, very)

plc (famous.@entry, India)

INS

- INS (Instrument)

Definition: *Ins defines the instrument to carry out an event*

- I solved it with computer

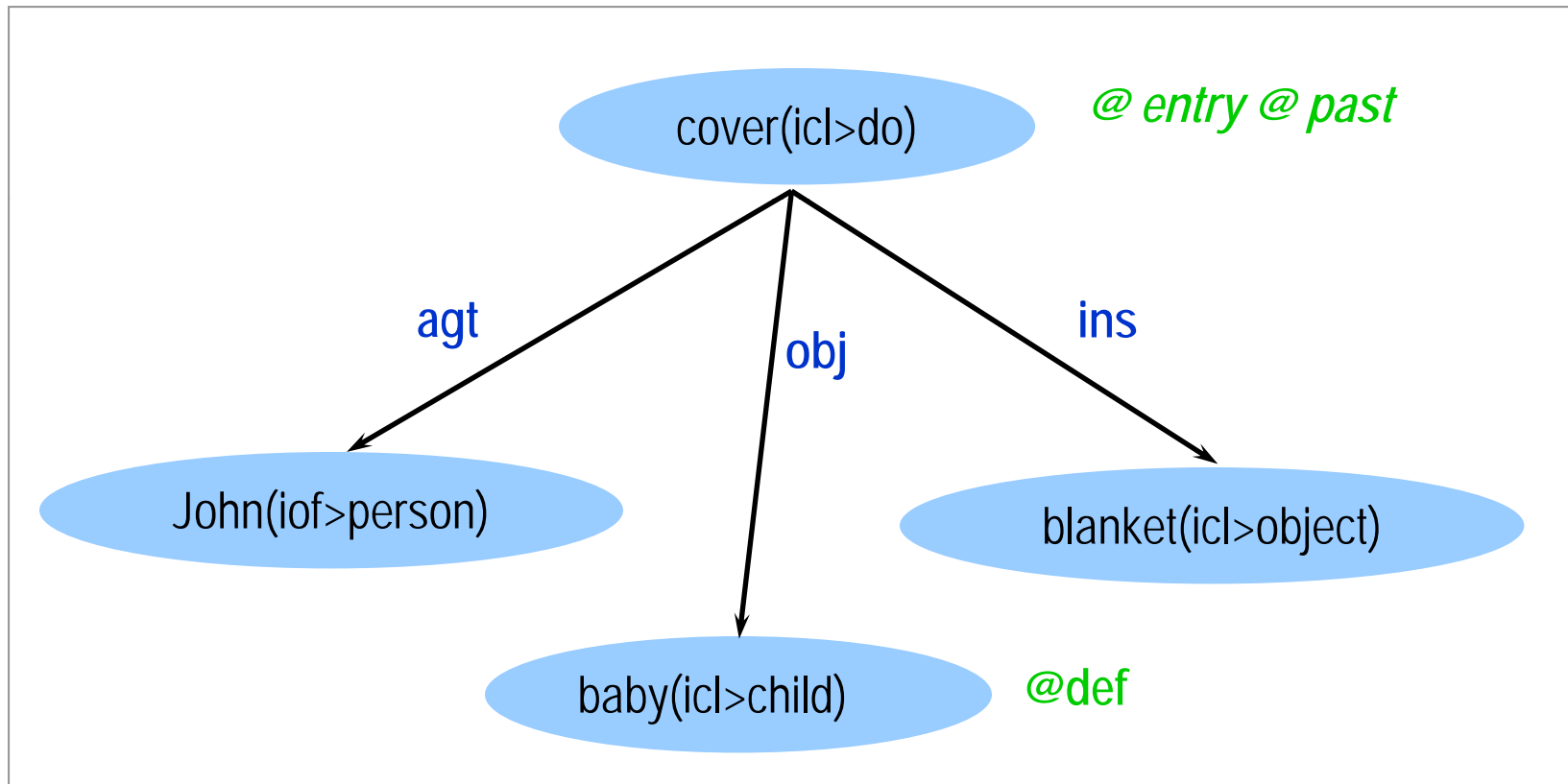
agt (solve.@entry.@past, I)

ins (solve.@entry.@past, computer)

obj (solve.@entry.@past, it)

INS

John covered the baby with a blanket.



Components of UNL

- Universal Words
- Relations
- **Attributes**

Attributes

- Arcs linking a node to itself
- Correspond to one-place predicates
- Attributes may convey three different kinds of information
 - Information on the role of the node in the UNL graph
 - Information conveyed by bound morphemes and closed classes
 - affixes (gender, number, tense, aspect, mood, voice, etc)
 - determiners (articles and demonstratives)
 - adpositions (prepositions, postpositions and circumpositions)
 - Conjunctions
 - auxiliary and quasi-auxiliary verbs (auxiliaries, modals, coverbs, preverbs)
 - degree adverbs (specifiers)
 - information on the (external) context of the utterance, i.e., non-verbal elements of communication, such as politeness, schemes, social deixis and speech acts

Categories of Attributes

- **Speech acts:**

@wish, @admire, @regret, @surprise

- **Propositional attitudes:**

@certain, @possible, @possible, @interrogativ

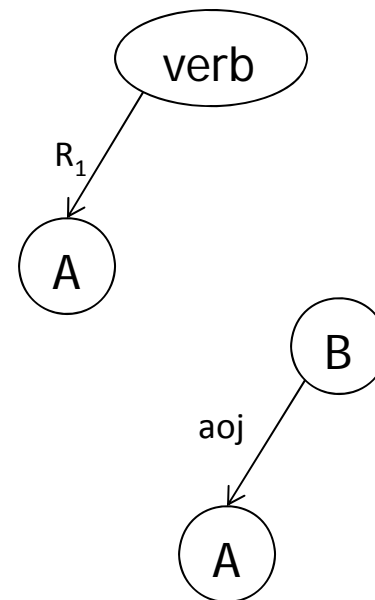
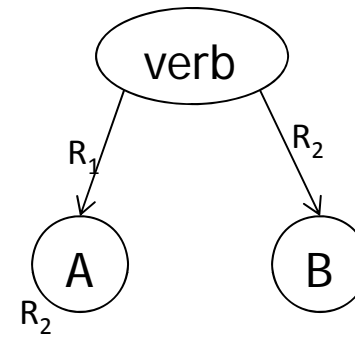
- **Focus:** @emphasis, @topic

- **Time with respect to speaker:**

@past, @present, @future etc.

English sentences: basic structure

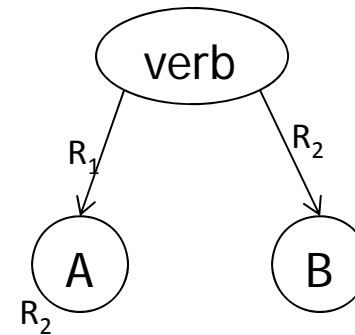
- $A <verb> B$
 - *John eats bread*
 - *agt(eat.@entry, John)*
 - *obj(eat.@entry, bread)*
- $A <verb>$
 - *John sleeps*
 - *aoj(sleep.@entry, John)*
- $A <be> B$
 - *John is good*
 - *aoj(good.@entry, John)*



Hindi sentences: basic structure

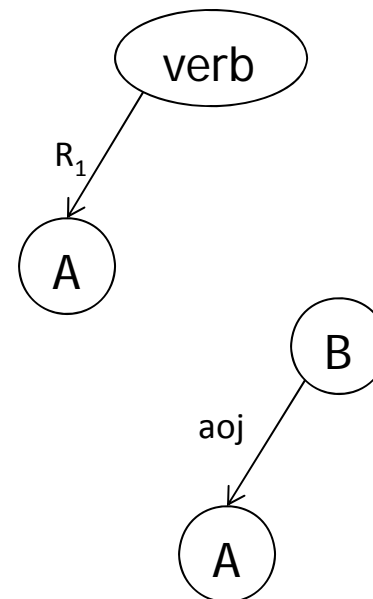
- **A B <verb>**

- *John roti khaata hai*
- *agt(eat.@entry, John)*
- *obj(eat.@entry, bread)*



- **A <verb>**

- *John sota hai*
- *aoj(sleep.@entry, John)*

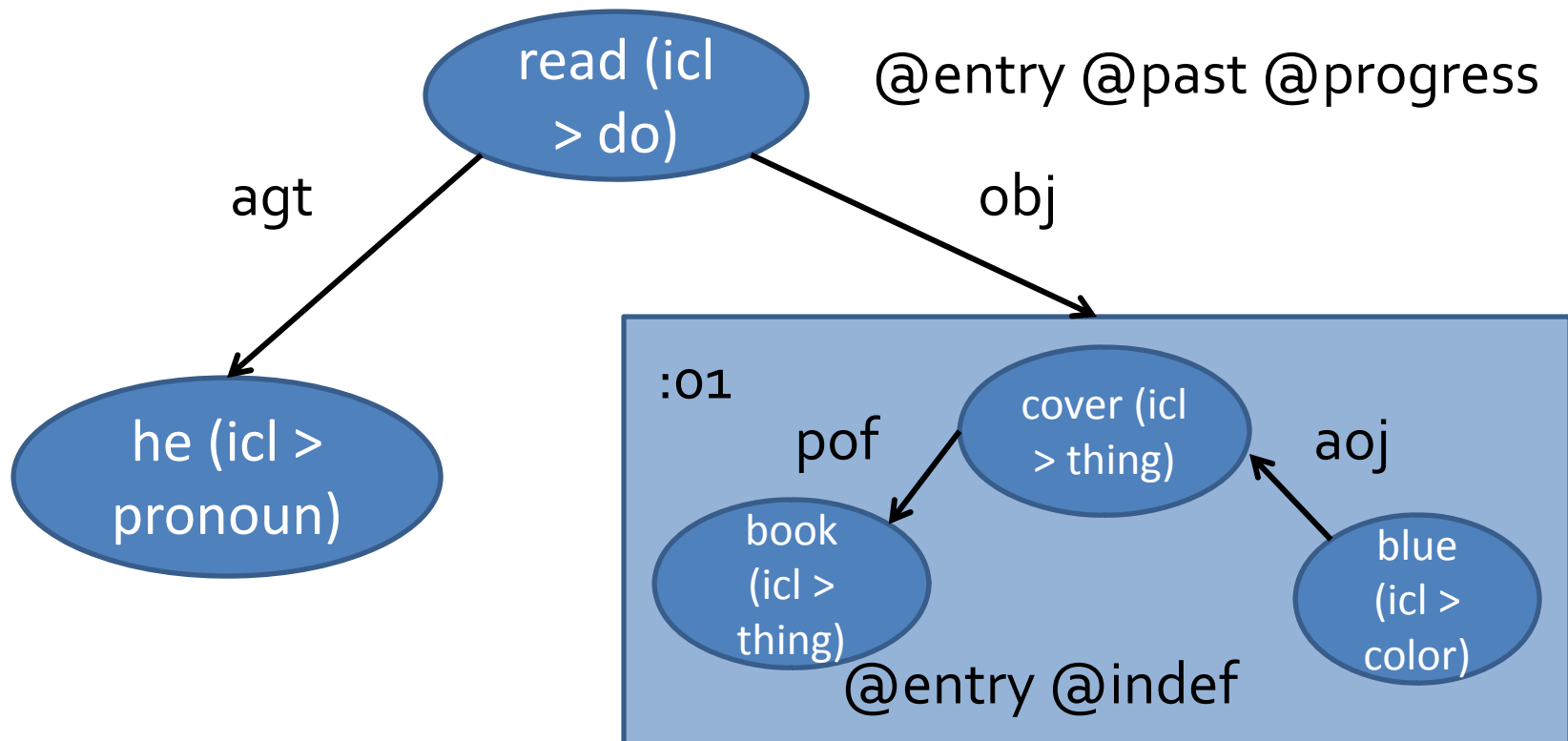


- **A <be> B**

- *John accha hai*
- *aoj(good.@entry, John)*

Scope example

- “He was reading a book with a blue cover.”



Non-graphical representation

- Like a markup language:

```
{unl}
```

```
agt (read (icl>do).@entry.@past.@progress, he (icl>pronoun));
```

```
obj (read (icl >do).@entry.@past.@progress, :01 );
```

```
pof:01 (cover (icl > thing) , book (icl > thing));
```

```
aoj:01(blue (icl > color , cover (icl > thing)))
```

```
{\unl}
```

UNL representation

{unl}

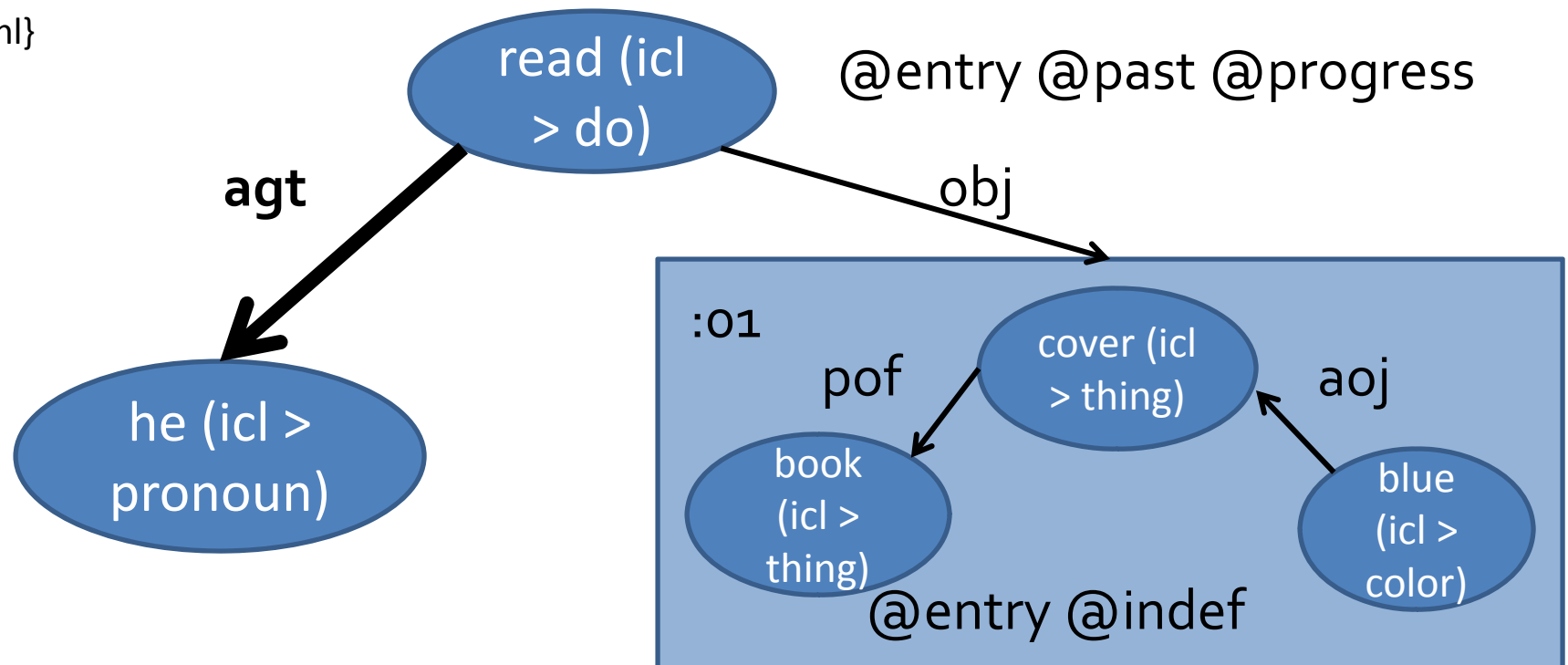
agt (read (icl>do).@entry.@past.@progress, he (icl>pronoun));

obj (read (icl >do).@entry.@past.@progress, :01);

pof:01 (cover (icl > thing) , book (icl > thing));

aoj:01(blue (icl > color , cover (icl > thing)))

{\unl}



UNL representation

{unl}

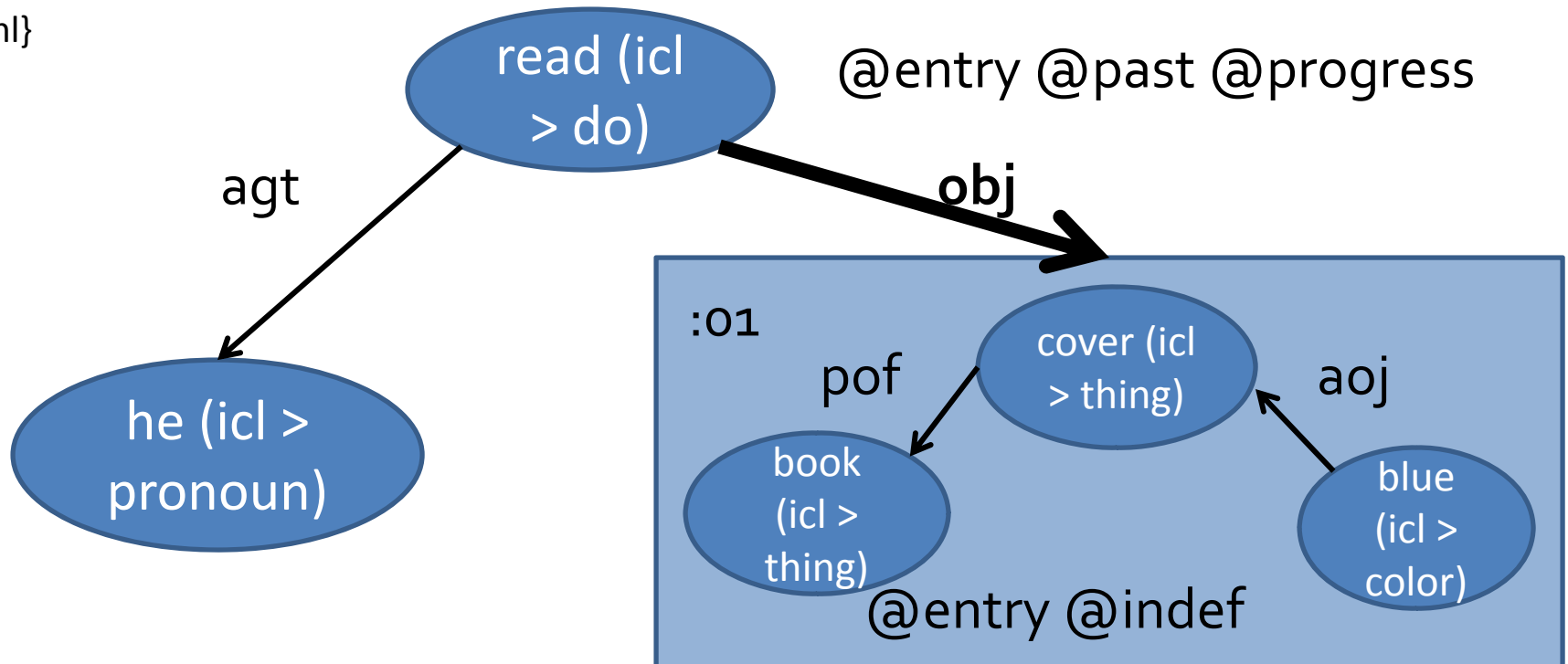
agt (read (icl>do).@entry.@past.@progress, he (icl>pronoun));

obj (read (icl >do).@entry.@past.@progress, :01);

pof:01 (cover (icl > thing) , book (icl > thing));

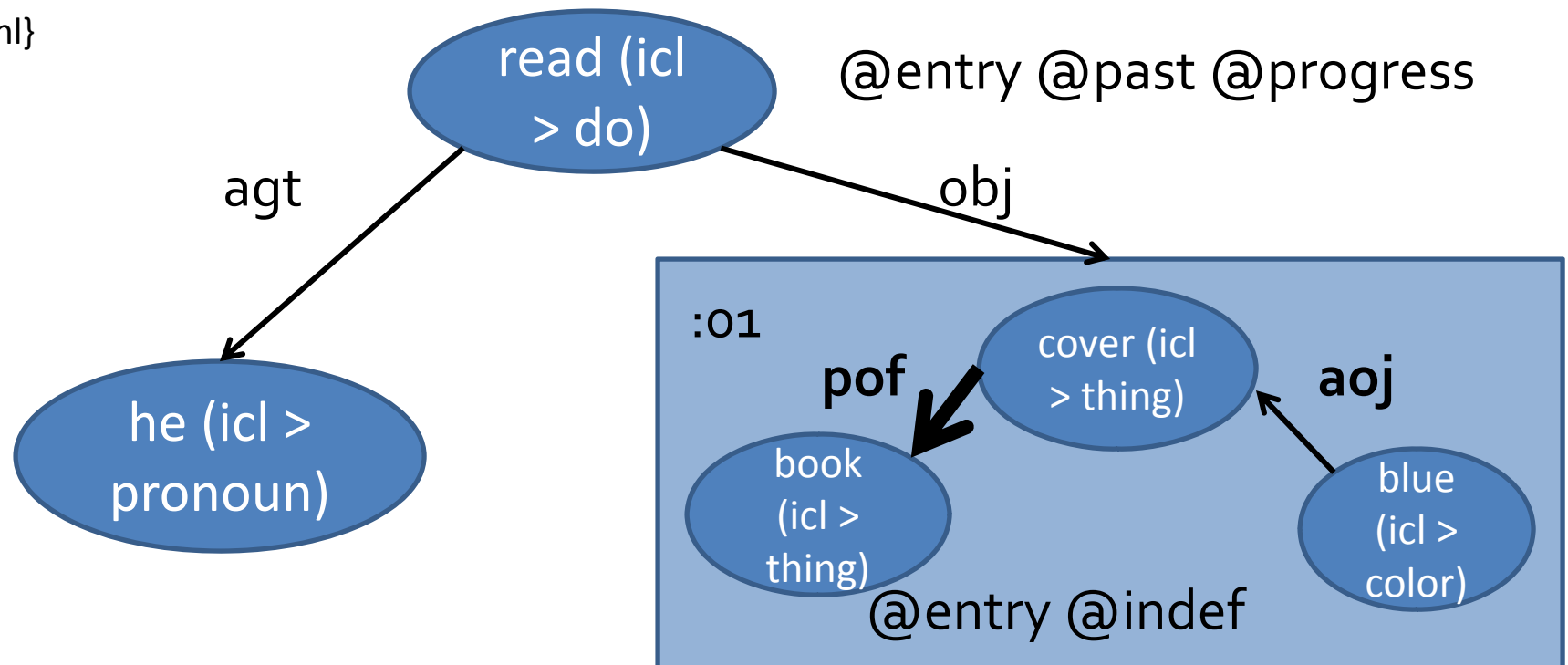
aoj:01(blue (icl > color , cover (icl > thing)))

{\unl}



UNL representation

```
{unl}  
agt (read (icl>do).@entry.@past.@progress, he (icl>pronoun));  
obj (read (icl >do).@entry.@past.@progress, :01 );  
pof:01 (cover (icl > thing) , book (icl > thing));  
aoj:01(blue (icl > color , cover (icl > thing)))  
{\unl}
```



UNL representation

{unl}

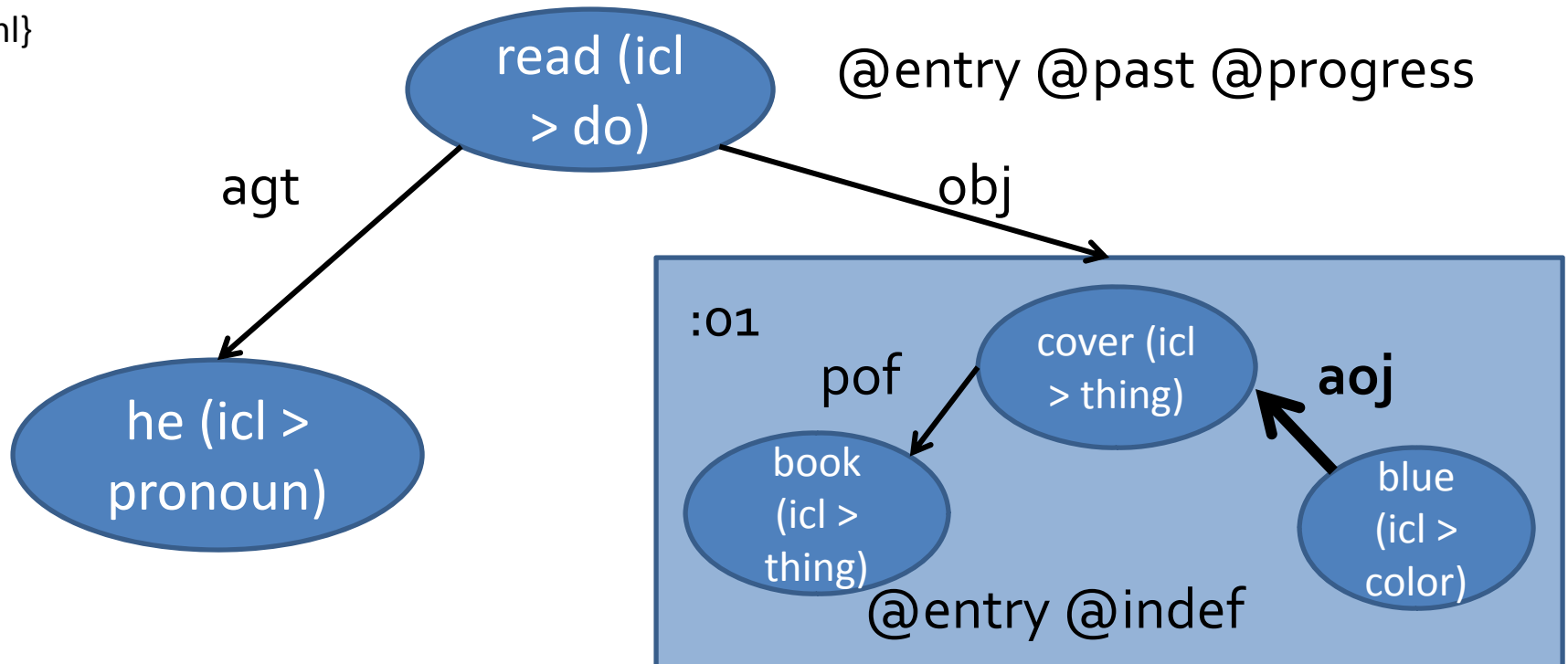
agt (read (icl>do).@entry.@past.@progress, he (icl>pronoun));

obj (read (icl >do).@entry.@past.@progress, :01);

pof:01 (cover (icl > thing) , book (icl > thing));

aoj:01(blue (icl > color , cover (icl > thing)))

{\unl}



UNL Enconverter at IIT Bombay

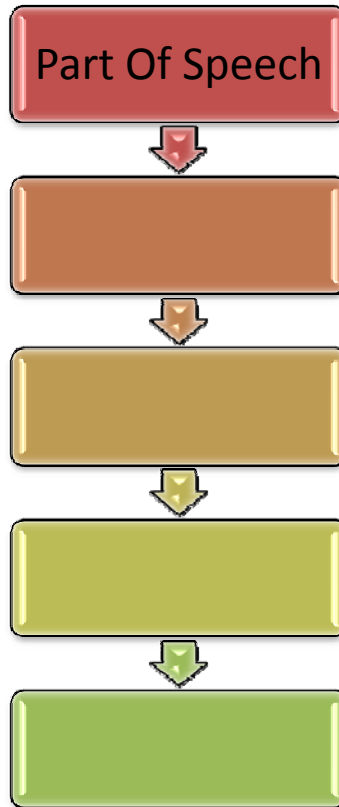
Parts of Speech Tagging

Sentence: *I went with my friend, John, to the bank to withdraw some money but was disappointed to find it closed.*

ISSUES

Part Of Speech

Which *bank* is it?
Noun or Verb



Named Entity Recognition

Sentence: *I went with my friend, John, to the bank to withdraw some money but was disappointed to find it closed.*

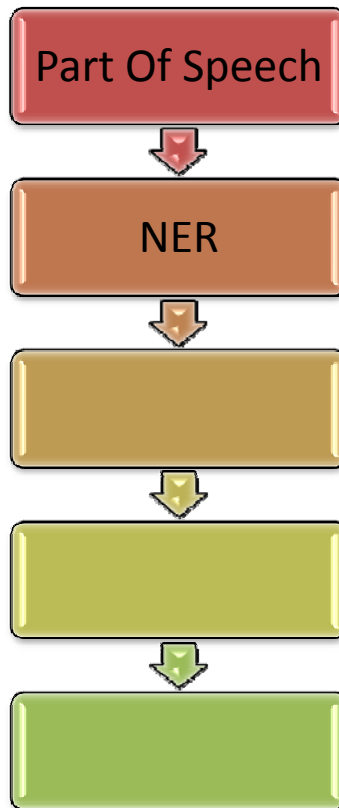
ISSUES

Part Of Speech

John is the name
of a PERSON



NER



Word Sense Disambiguation

Sentence: *I went with my friend, John, to the bank to withdraw some money but was disappointed to find it closed.*

ISSUES

Part Of Speech



NER

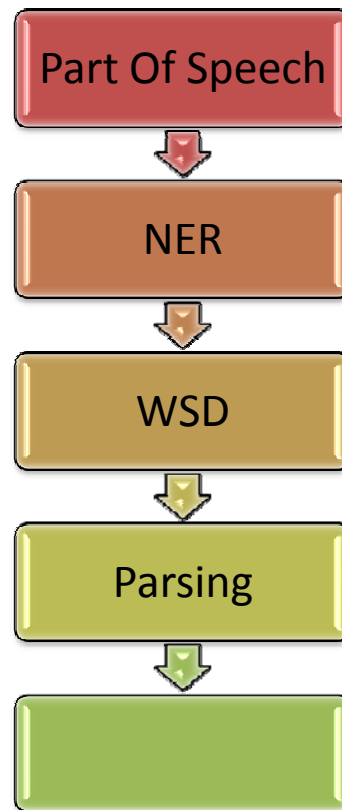


WSD



Which *bank* is it?
Financial bank or
River bank

Parsing



Issues: Co-reference Resolution

Sentence: *I went with my friend, John, to the bank to withdraw some money but was disappointed to find it closed.*

ISSUES

Part Of Speech



NER



WSD



Co-reference



"it" is referring to "bank".

Issues: Subject Drop

Sentence: *I went with my friend, John, to the bank to withdraw some money but was disappointed to find it closed.*

ISSUES

Part Of Speech



NER



WSD

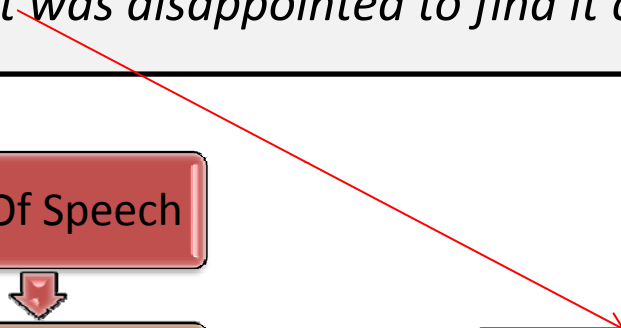


Co-reference



Subject Drop

Subject / is
dropped



System Demo

Summary

- UNL is an artificial language processable by machine
- Unambiguous representation unlike natural language
- Information divided in three components:
 - Universal words: disambiguated words each having a single sense
 - Relations: Binary semantic relations among universal words explaining their roles
 - Attributes: Giving subjective information with respect to the speaker
- For generation of UNL many different NLP tools are required together