



Universal Networking Language: A Semantic Network

Rajat Kumar Mohanty
rkm@cse.iitb.ac.in

Centre for Indian Language Technology
Department of Computer Science and Engineering
Indian Institute of Technology Bombay
Mumbai, India
<http://www.cfilt.iitb.ac.in>

**2nd Asian Regional Training on Local Language Computing 2005
Siem Reap, Cambodia**

Introduction

Stages of Language Processing

- Phonetics and Phonology: process *sound*
- Morphology: process *word formation*
- Lexical Analysis: process *words*
- Syntactic Analysis: process *structure*
- Semantic Analysis: process *meaning*
- Pragmatics: process *user intention*
- Discourse: process *connected text*

Motivation for Semantic Analysis

- Words can have several distinct meanings for any syntactic class. (e.g., *crane, bank, etc*)
- Semantics: what words mean and how these meanings combine in sentences to form sentence meanings.
- Same verb but different meaning
 - *John teaches NLP to Mary.*
 - *Mary teaches NLP to John.*
- Different verb but same meaning
 - *John bought a house from Mary.*
 - *Mary sold a house to John.*

Roadmap

- Linguistic Background
 - Predicate-Argument Structure
 - Subcategorization
 - Selectional Restrictions
 - Lexical Conceptual Structure
 - Thematic Roles
 - Some Linguistically Relevant Concepts
- Resources for Semantic Processing
- Universal Networking Language (UNL)
 - UNL: The Framework
 - The Process of Enconversion
- Conclusion

Linguistic Background

Predicate-Argument Structure

- Natural Languages have a form of predicate-argument arrangement at the core of their semantic structure.
- Examples
 - (a) I want an NLP book.
 - (b) I want to spend less than \$5 for this book.
 - (c) I want John to read this book.
- Syntactic Frames:
 - *[NP-want-NP]*
 - *[NP-want-inf_VP]*
 - *[NP-want-NP-inf_VP]*

Subcategorization

- Ungrammaticality
 - The boy relied on the girl.
 - * The boy relied the girl.
 - *The boy relied.
- Specify the categorial class of the lexical item
- Specify the environment
- Examples
 - cry: [V; _]
 - frighten: [V; _NP]
 - rely: [V; _PP]
 - forward: [V; _ NP PP]

Subcategorization Frame

forward
V
— NP PP

e.g., We will be **forwarding** our new catalogue to you

invitation
N
— PP

e.g., An **invitation** to the party

accessible
A
— PP

e.g., A program making science is more **accessible** to young people

Selectional Restrictions

- Grammatically Sound but Semantically Odd
 - *The boy frightens sincerity.
 - *Sincerity kicked the boy.
- Inherent Properties of Nouns:
[+/- ABSTRACT], [+/- ANIMATE]
- *E.g.*,
Sincerity [+ ABSTRACT]
Boy [+ANIMATE]
- $V \longrightarrow \textit{frighten} / \left\{ \begin{array}{l} [+/-\text{ABSTRACT}] \text{ ---} \\ \text{---} [+ANIMATE] \end{array} \right\}$

Arguments and Adjuncts

- Arguments: The participants are referred to as arguments of the verb.
 - We will be **forwarding** [our new catalogue] [to you]
- Obligatory vs Optional Arguments
- Adjuncts: Spatial and temporal expressions, manner adverbs, and other expressions that can go with almost any sentences.
 - We will be **forwarding** our new catalogue to you [on Monday].

Lexical Conceptual Structure

- Example:
 - *Bill ran into the room*
- Syntactic Structure:
 - [S [NP Bill] [VP ran [PP into [NP the room]]]]
- Conceptual Structure:

GO ([Thing Bill], [Path TO [Place IN [Thing the room]]])
Event

Lexical Conceptual Structure

- Primitives
 - GO, BE, STAY, CAUSE...
 - TO, FROM, AWAY...
- Events can be elaborated as either of the two Event-functions GO or STAY, each of which takes two arguments

Conceptual Constituents

- The semantic structure of a sentence is built up from a hierarchical arrangement of *conceptual constituents*.
- Conceptual constituents are realized syntactically by means of major phrasal constituents (such as, *NP, S, PP, AP, AdvP*)
- Each of the conceptual constituents belongs to a major ontological category or *semantic part of speech*.
- Semantic Parts-of-Speech: *Thing, Place, Path, Event, State, Manner, and Property*

Semantic Parts-of-Speech

- The correspondence of semantic to syntactic categories is governed by markedness convention.
- The unmarked realization of
 - *Thing* is *NP*
 - *Place* and *Path* is *PP*
 - *Property* is *AP*
 - *Manner* is *AdvP*
 - *Event* and *State* is *S*
- However, the marked realizations also occur
 - The NP *a bumper* expresses *Property*
 - The NP *earthquake* expresses an *Event*

Place- and Path-function

PLACE → $\left\{ \begin{array}{l} \text{PLACE-FUNCTION ([THING])} \\ \text{Place} \end{array} \right\}$

(e.g., in the room)

PATH → $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{TO} \\ \text{FROM} \\ \text{TOWARD} \\ \text{VIA} \end{array} \right\} \\ \text{Path} \end{array} \right\} \text{ ([THING])}$

(e.g., to the station)

Motion Verbs

- Motion verbs map two arguments, a **Thing** and a **Path**, into an **Event** consisting of the *Thing* traversing the *Path*. (e.g., *John ran into the room*)

- Lexical entry:

run

+V, -N

[_ PP_{*j*}]

[_{Event} GO ([_{Thing} *i*], [_{Path} *j*])]

- Motion verbs are represented generically by *GO* and differentiated from one another by various markers of manner.

Mapping a *Thing* into a *Path*

- The preposition *into* is a function that maps a thing –the *reference object* – into a *Path*.
- Lexical entry:
$$\left(\begin{array}{l} \text{into} \\ \text{P} \\ [_ \text{NP}_j] \\ [\text{Path TO } [\text{Place IN } [\text{Thing } j]]] \end{array} \right)$$
- To satisfy the well-formedness conditions on the use of *into*, its sister phrase must be an NP (the syntactic condition) and must express a concept of a category *Thing* (the semantic condition).

Example

- *John entered the room*

- Lexical entry: $\left(\begin{array}{l} \text{enter} \\ +V, -N \\ [_ NP_j] \\ [_{\text{Event}} \text{GO} ([_{\text{Thing}} i], [_{\text{Path}} \text{TO} [_{\text{Place}} \text{IN} [_{\text{Thing}} j]])] \end{array} \right)$

- The position and category of the variable in *enter* result in a syntactic realization as transitive verbs plus NP instead of verb plus PP.

Thematic Roles: Motivation

- The case of *open* (*Are these sentences underlying related?*)
 - *John opened the door with a key.*
 - *The door was opened by John with a key.*
 - *The key opened the door.*
- *Thematic Roles* are part of the level of conceptual structure, not part of syntax.

Thematic Roles

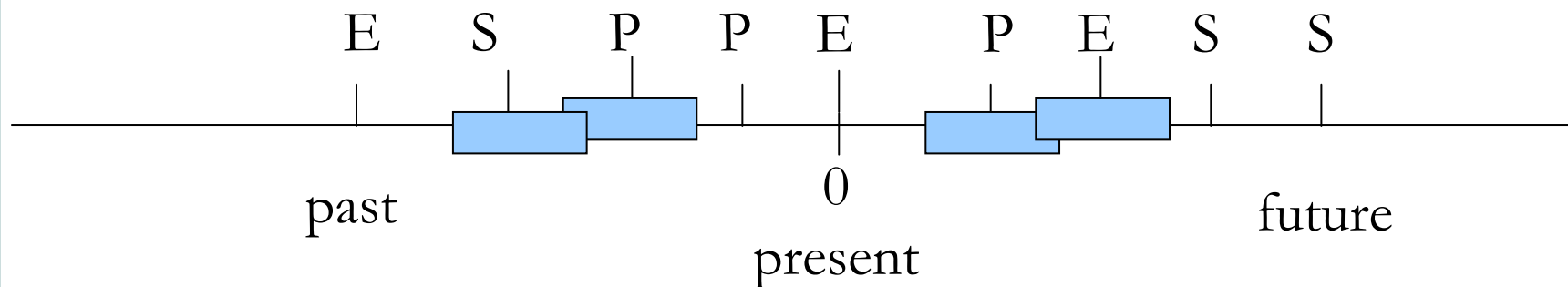
- Agent: The instigator of an event
- Patient: A patient is directly affected by an action
- Theme: the object in motion or being located
- Source: the object from which motion proceeds
- Goal: the object to which motion proceeds

Some Linguistically Relevant Concepts

- Time and Tense
 - Grammaticalized Location in Time
 - The Domain of Tense (Syntactic and Semantic)
- Aspect
 - Viewpoint type
 - Situation type
- Modality
 - Propositional Modality
 - Event Modality

Time Line

- *Time* can be represented as a straight line, with the past represented conventionally to the left and the future to the right
- Whether the *time line* is bounded at either end : linguistically irrelevant, but philosophically important



Representation of situations in the time line

Time and Language

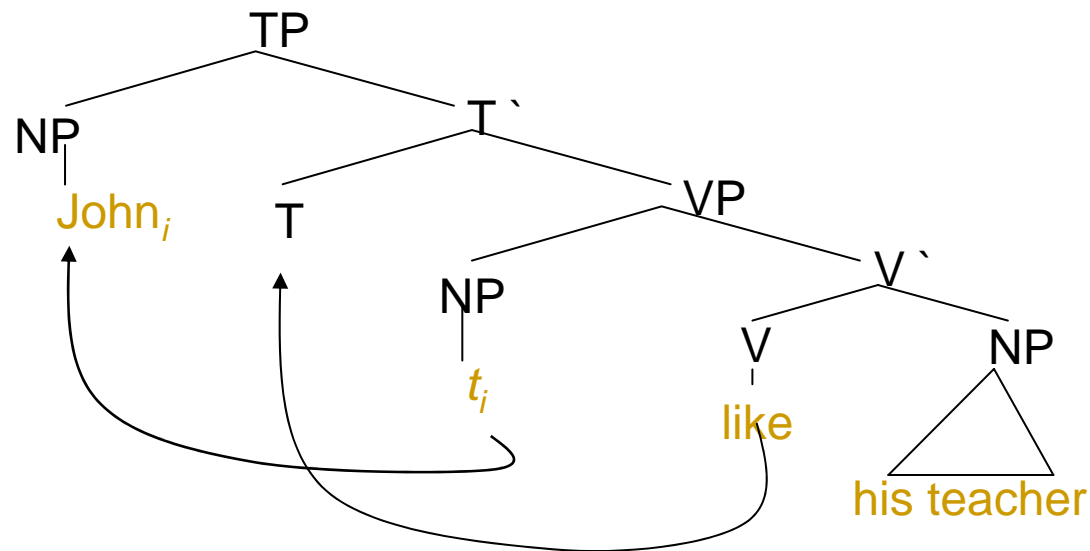
- All languages allow speakers and listeners to locate situations (*i.e.*, *event*, *process*, *state*) in time
- The idea of locating *situations* in *time* is purely a conceptual notion
- The ways of locating *situation* in *time* differ from one another on two parameters
 - the degree of accuracy of temporal location that is achievable in different languages
 - the relative weight assigned to the lexicon and to the grammar in establishing location in time

Grammaticalized Location in Time

- Expressions for locating in time can be divided into three classes
 - Lexically composite expressions (e.g., *five minutes after John left, last year*)
 - Lexical items (e.g., *now, today, yesterday, last*)
 - Grammatical categories (e.g., *tense, aspect, modals*)
- Languages express location in *time* by means of grammatical categories, e.g.,
 - (1) John sang (2) John sings
 - (3) John has read this story (4) John is reading this story
- The difference in (1) and (2) in English is *tense*
- The difference in (3) and (4) is *aspect*

Tense in the Syntactic Domain

- The syntactic domain of tense is the clause
- A tense morpheme is generated in each clause
- In Principles and Parameters Theory (now *The Minimalist Program*) tense heads a functional category, *i.e.*, *TP*



Tense in the Semantic Domain

- Semantic interpretation of tense requires information from the contexts
- Properties of Time in the semantic domain
 - There is the *time* at which the act of speaking itself takes place
 - There is a large set of *times* at which various events take place
 - There are all the *times* other than the time of speaking, which can serve as alternative points of reference for the speaker
- Each tense
 - involves three times: *speech time*, *reference time* and *event time*
 - conveys information about two relations between them
 - A relation between Speech Time and Reference Time
 - And, a relation between Reference Time and Event Time

Tense in the Semantic Domain

- Examples

(1) Mary has arrived. (RT is the same as SpT)

(2) Mary arrived. (RT is the same as ET)

(3) On Sunday, Mary had already arrived.

- (3) is a past perfect that needs three different times for semantic interpretation: SpT, RT (the Sunday before), and ET

- The schematic meanings for the tenses of (1-3):

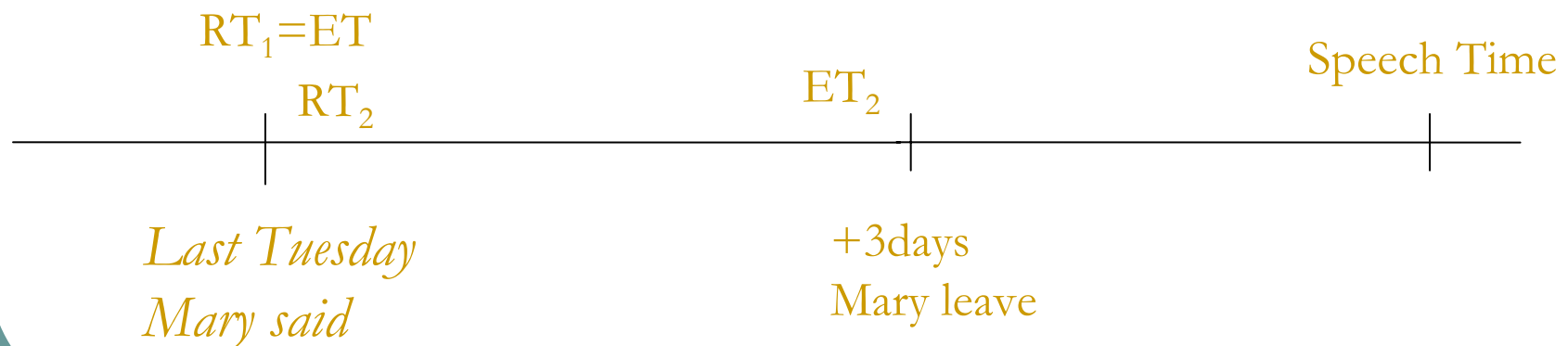
- Present: RT = SpT, ET = RT

- Past: RT < SpT, ET = RT

- Past Perfect: RT < SpT, ET < RT

Tense in the Semantic Domain

- Mary said last Tuesday that she was leaving in 3 days
- Temporal Interpretation
 - Main clause: $RT_1 < SpT$; $RT = \text{last Tuesday}$; $RT = ET_1$
 - Embedded Clause: $RT_2 = RT_1$; $ET_2 > RT_2$

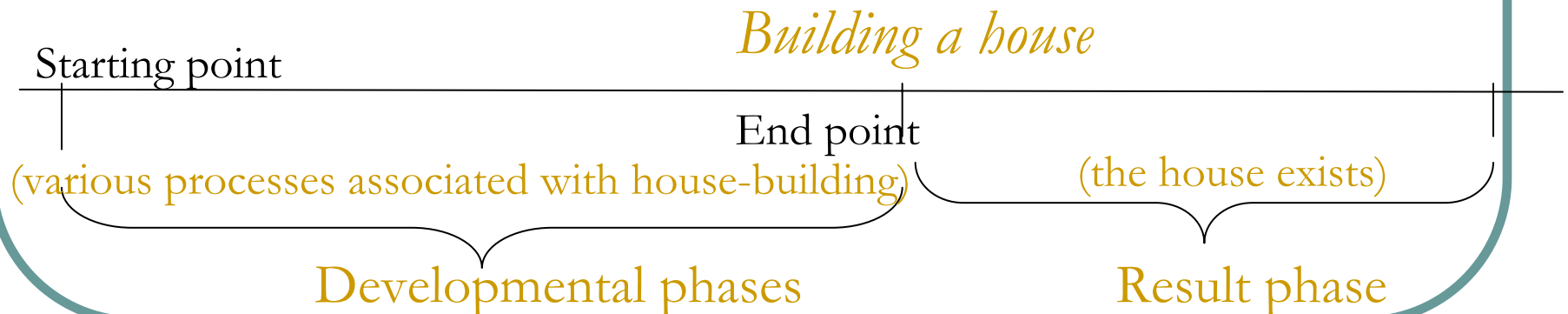


Tense in the Semantic Domain

- The function of tense is to locate the position of the speaker's reference time, by relating it to a position of the time of speaking
- Three primary possibilities for this relationship
 - Past $RT < SpT$
 - Present $RT = SpT$
 - Future $SpT < RT$

Aspect

- An event is said to evolve through a series of temporal *phases*
- One of these temporal phases is the actual time of the event itself, inclusive of its end-point
 - *E.g., an event of building a house*
- *Verb aspect* involves reference to one of the temporally distinct phases in the evolution of *building a house* through time



Aspectual Information

- Tense interpretation interacts with aspectual information
- Examples
 - (1) Mary walked to school.
 - (2) Mary was walking to school.
 - (3) Mary walked in the park.
- (1) presents a complete event that has an endpoint, and the information that the goal was reached.
- (2) presents a part of the same type of event, but does not convey whether the goal was reached.
- (3) presents a complete event that does not involve a goal, and the information that the event was terminated.

The Two Components in Sentences

- The aspectual information is given by the linguistic forms that appear in a sentence.
- Aspectual systems have two components:
viewpoint and *situation* type
- Aspectual situation type indirectly classifies a sentence as expressing an eventuality, a state or an event. The information is conveyed by the verb constellation.
- Aspectual viewpoint, conveyed morphologically, focuses all or part of the eventuality.

Aspectual Viewpoint Types

- Perfective

- Perfective viewpoints focus a situation in its entirety, including both initial and final endpoints.
- *e.g.*, John has built a house.

- Imperfective

- Imperfective viewpoints focus part of a situation, including neither initial nor final endpoints.
- *E.g.*, Mary was walking in the park

The English Verb Group

- On the original timetable for this project, by this time the report *would have been being printed*.

WOULD	HAVE	BEEN	BEING	PRINTED
Modal	perfective	progressive	passive	Main verb
<i>Will +</i>	<i>have</i>	<i>be + -en</i>	<i>be + -ing</i>	<i>print + -en</i>
Past tense	(+ -en →)	(+ -ing →)	(+ -en →)	
↑	↑	↑	↑	
Modal aux	Aspectual aux	Aspectual aux	Copular aux	

Temporal Properties

- Dynamism: [static/dynamic]
 - *States* are static (undifferentiated period of states).
Events are dynamic (successive stages of events).
- Telicity: [telic/atelic]
 - Events may be telic or atelic. A telic event has a natural final endpoint. An atelic event has no outcome. It has arbitrary final endpoint. (e.g., *walked to school, walked in the park, etc*)
- Duration: [durative/instantaneous]
 - Situations are durative or instantaneous. (e.g., *strolled in the park, win the race, etc*)

Telic Events

- Telic events may be classified by the type of result they bring about.
- Major type of results
 - Affected object: *bend an iron bar, break a pot, wrinkle a dress*
 - Constructed object: *build a house, write a letter*
 - Consumed object: *drink a glass of wine, destroy a house*
 - Affected experiencer: *amuse Mary*
 - Path-Goal: *walk to the lake, work from 2 to 3*

Temporal Features of the Situation Types

situations	static	durative	telic	examples
stative	[+]	[+]	[-]	<i>Know the answer</i>
activity	[-]	[+]	[-]	<i>Laugh, stroll in the park</i>
accomplishment	[-]	[+]	[+]	<i>Walked to school</i>
semelfactive	[-]	[-]	[-]	<i>Tap, knock</i>
achievement	[-]	[-]	[+]	<i>Reach the top, win the race</i>

The Activity Situation Type

- Temporal features: [dynamic], [atelic], [durative]
- Examples
 - He is sleeping.
 - They are widening the road.
- Activities terminate or stop; the notion of completion is irrelevant to a process event.
- Temporal schema of Activities:
 - I.....F_{Arb}

The Accomplishment Situation Type

- Temporal features: [dynamic], [telic], [durative]
- Examples
 - Build a bridge, walk to school, drink a glass of wine
- Accomplishments have successive stages in which the process advances to its natural final endpoint. They result in a new state.
- The event is completed and cannot continue.
- Temporal schema of Activities:
 - I.....F_{Nat R}

The Semelfactive Situation Type

- Semelfactives are single-stage event with no result or outcome.
- They occur very quickly.
- Temporal features: dynamic, atelic, instantaneous
- Examples
 - Knock at the door
 - The light flicker
 - Blink, cough
- Temporal schema of Activities: E

The Achievement Situation Type

- Achievements are instantaneous events that result in a change of state
- Temporal features: dynamic, telic, instantaneous
- Examples
 - Leave the house, recognize the person, reach the top, break a glass
- Temporal schema of Activities:
 - E_R

Major types of Results of Achievement

- Affected object: *break a cup, tear a paper*
- Constructed object: *imagine a city, define a parameter*
- Consumed object: *explode a bomb*
- Affected experiencer: *see a comet*
- Path-Goal: *reach the top, arrive in Mumbai*

Stative Situation Type

- States are stable situations which hold for a moment or an interval.
- Temporal features: static, durative
- Examples
 - Own the farm, be tall, believe in ghosts, hope that..., fears that..., know that...
- Temporal schema of Activities:
 - (I).....(F)

Illustrations

- John knew the truth (stative)
- Suddenly Bill knew the truth (achievement)
- Mary coughed (semelfactive)
- Mary coughed for an hour (activity)
- John played cricket yesterday (accomplishment)
- Mary swam in the pond. (Activity)
- John wrote a letter. (Accomplishment)
- Mary knocked at the door (Semelfactive)
- Mary reached the guest house (achievement)

Modality

- Modality differs from tense and aspect in that it does not refer directly to any characteristic of the events, but simply to the status of the proposition.
- *Propositional Modality* (concerned with the speaker's judgment of the proposition)
 - *John may be at home now. (it is possible that...)*
 - *John must be at home now. (it is necessary that...)*
- *Event Modality* (concerned with the speaker's attitude towards a potential future event)
 - *John may come in now. (it is possible for John to come in now...)*
 - *John must come in now. (it is necessary for John to come in now...)*

Propositional Modality

- Epistemic
 - The speaker expresses his judgment about the factual status.
 - *John must be in the office* (The speaker makes a firm judgment, on the basis of evidence, e.g., *that the office lights are on; that he is not at home, etc.*)
- Evidential
 - The speaker indicates the evidence they have for its factual status.
 - He is said to be extremely rich. (reported)
 - He claims to have shot down a mosquito. (reported)

Event Modality

- Deontic

- The conditioning factors are external. It relates to obligation or permission emanating from an external source.

- *John may come in now.* (permission)
- *John must come in now.* (obligation)

- Dynamic

- The conditioning factors are internal. It relates to ability or willingness, which comes from the individual concern.

- *John can speak Hindi.* (ability)
- *John will do it for you.* (willingness)

Modal Verbs

- Modals (English)
 - will, shall, may, can, must, would, should, might, could
 - Ought (to), need (to), dare (to)
- Modal verbs are used in all four types of modality
 - Epistemic
 - Evidential
 - Deontic
 - Dynamic

Some Formal Differences

- Deontic MUST has negative mustn't and needn't, but epistemic MUST has no morphologically related negative.
 - John must / mustn't come for the meeting. (deontic)
 - John must /*mustn't be in his office. (epistemic)
- MAY and MUST followed by HAVE always epistemic.
 - John may/must have been in his office.
- MAY is replaceable by CAN only in deontic use
 - You may go now. (deontic)
 - You can go now. (deontic)
 - He may/*can be in his office. (epistemic)
- When MUST refers to the future, it is deontic.
 - John must come for the meeting. (deontic)
 - John must be in his office. (epistemic)

Resources for Semantic Processing

Resources for Semantic Processing

- LCS Resources (Dorr, 2001)
 - http://www.umiacs.umd.edu/~bonnie/LCS_Database_Documentation.html
- WordNet (George A. Miller, Christiane Fellbaum and others)
 - <http://wordnet.princeton.edu/>
- VerbNet (Martha Palmer and others)
 - <http://www.cis.upenn.edu/group/verbnet/>
- FrameNet (Charles Fillmore, Collin Baker and others)
 - <http://framenet.icsi.berkeley.edu/>
- Propbank (Martha Palmer and others)
 - <http://www.cis.upenn.edu/~ace/>

LCS Resources

- LCS: A compositional abstraction with language-independent properties (Jackendoff, 1990, 1996).
- An LCS captures the semantics of a lexical item through a combination of semantic structure and semantic content.
- LCS Database: Organized into semantic classes of English Verbs (Levin, 1993)
- Hand-tagged with WordNet senses using the 1.6 version of WordNet (Fellbaum, 1998).
- This representation has been used as the interlingua of several projects such as UNITRAN (Dorr, 1993) and MILT (Dorr, 1997).

WordNet

- A large lexical database with information about words and concepts (Miller *et.al.*, 1993)
- A resource used to identify shallow semantic features attached to lexical items
- Words are organized in synonym sets called synsets
- Each synset represents a concept
- Wide coverage of content words
- Does not encode syntactic or semantic behavior of verbs

VerbNet

- A verb lexicon compatible with WordNet
- Based on Levin's Verb classification (Levin, 1993)
- Associates the semantics of a verb with its syntactic frames
- Combines *Thematic roles* with subcategorization frames and selectional restrictions

FrameNet

- Based on *Frame Semantics*
- Defines Sentence level ontology
- Many primitives and several semantic roles
- Based on corpus analysis
- Still under development
- Coverage: about 7500 verbs

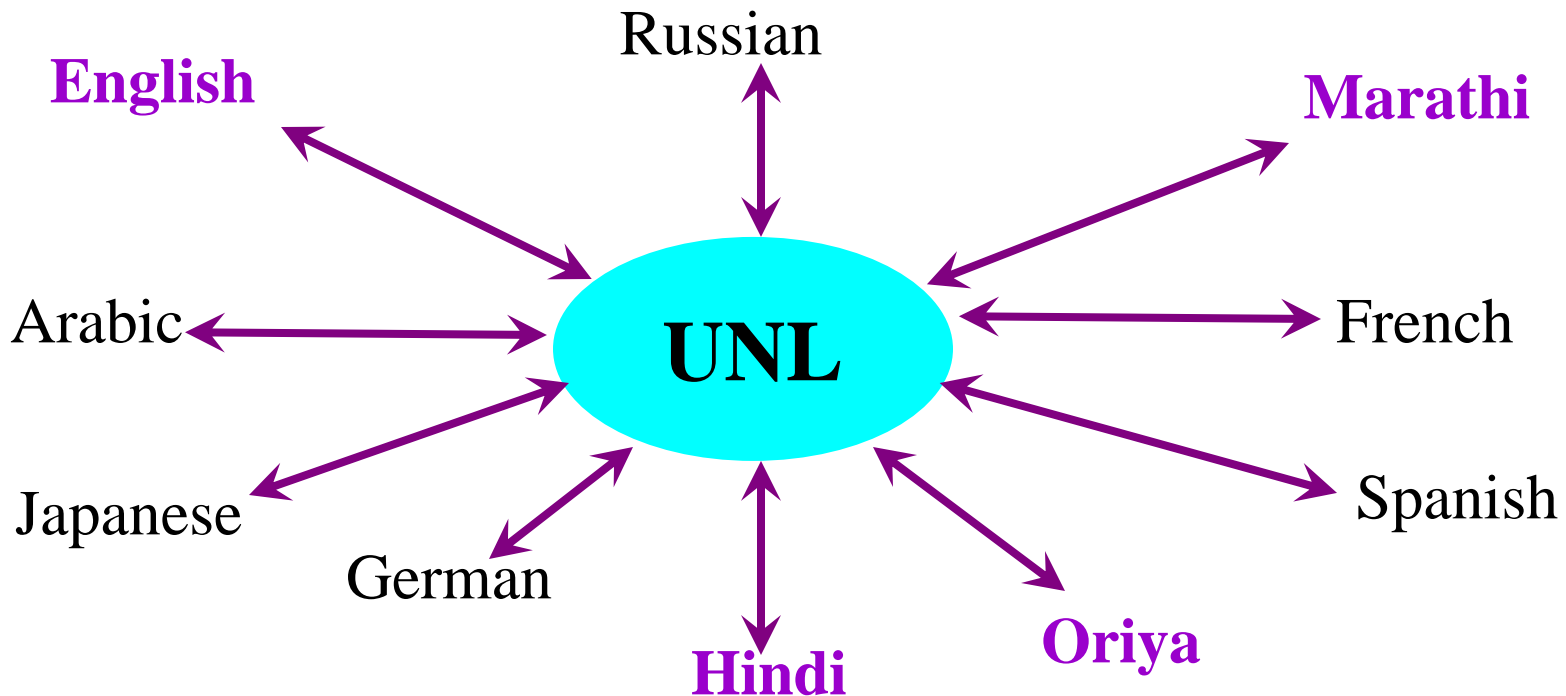
Proposition Bank

- Annotation of the Penn Treebank with predicate-argument information
- Useful for automatic learning of the surface realization of each argument

Universal Networking Language (UNL)

<http://www.undl.org>

The UNL Project

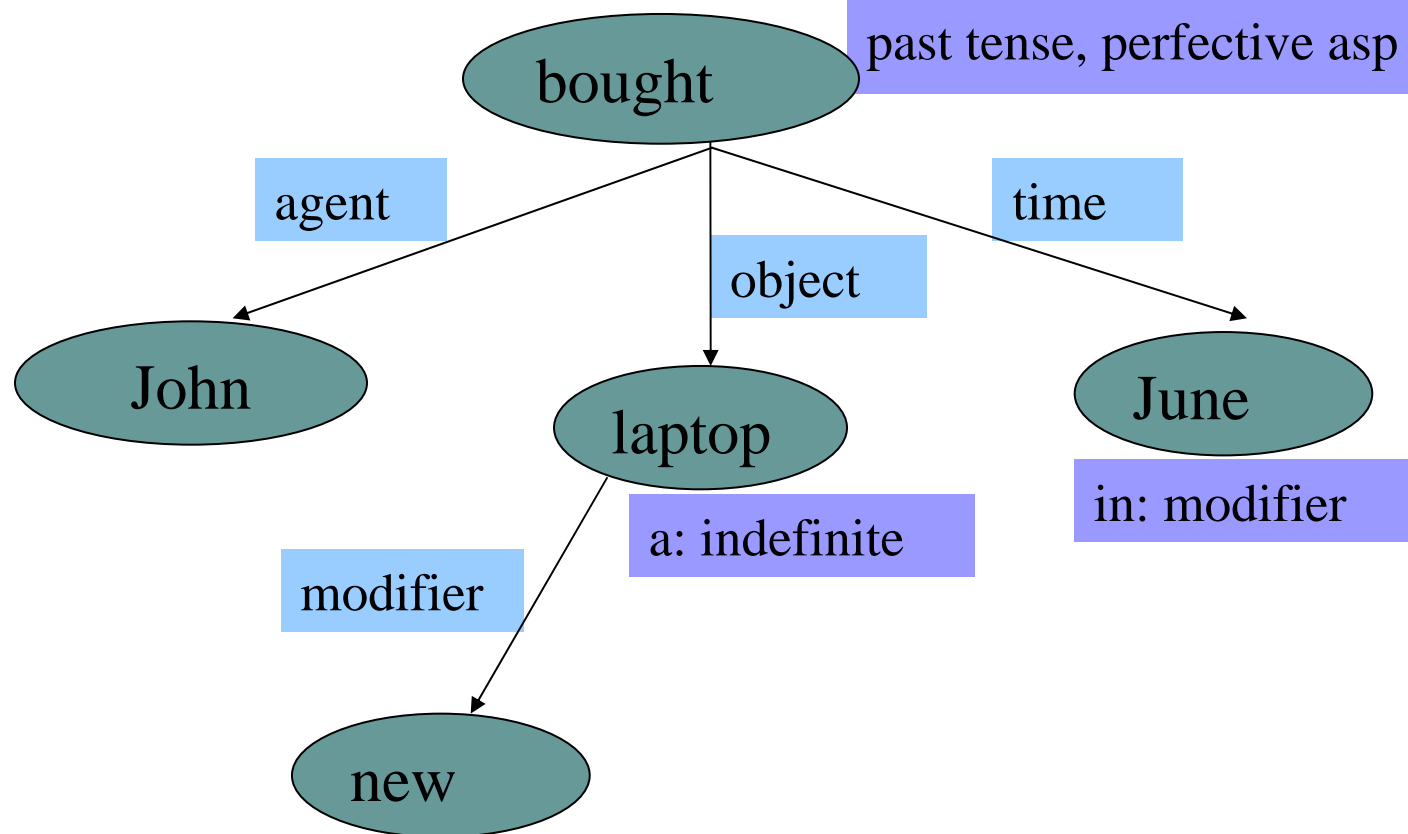


- Language independent meaning representation.
- Our goal : Robust and scalable UNL generation for multilingual meaning-based search engines.

Universal Networking Language

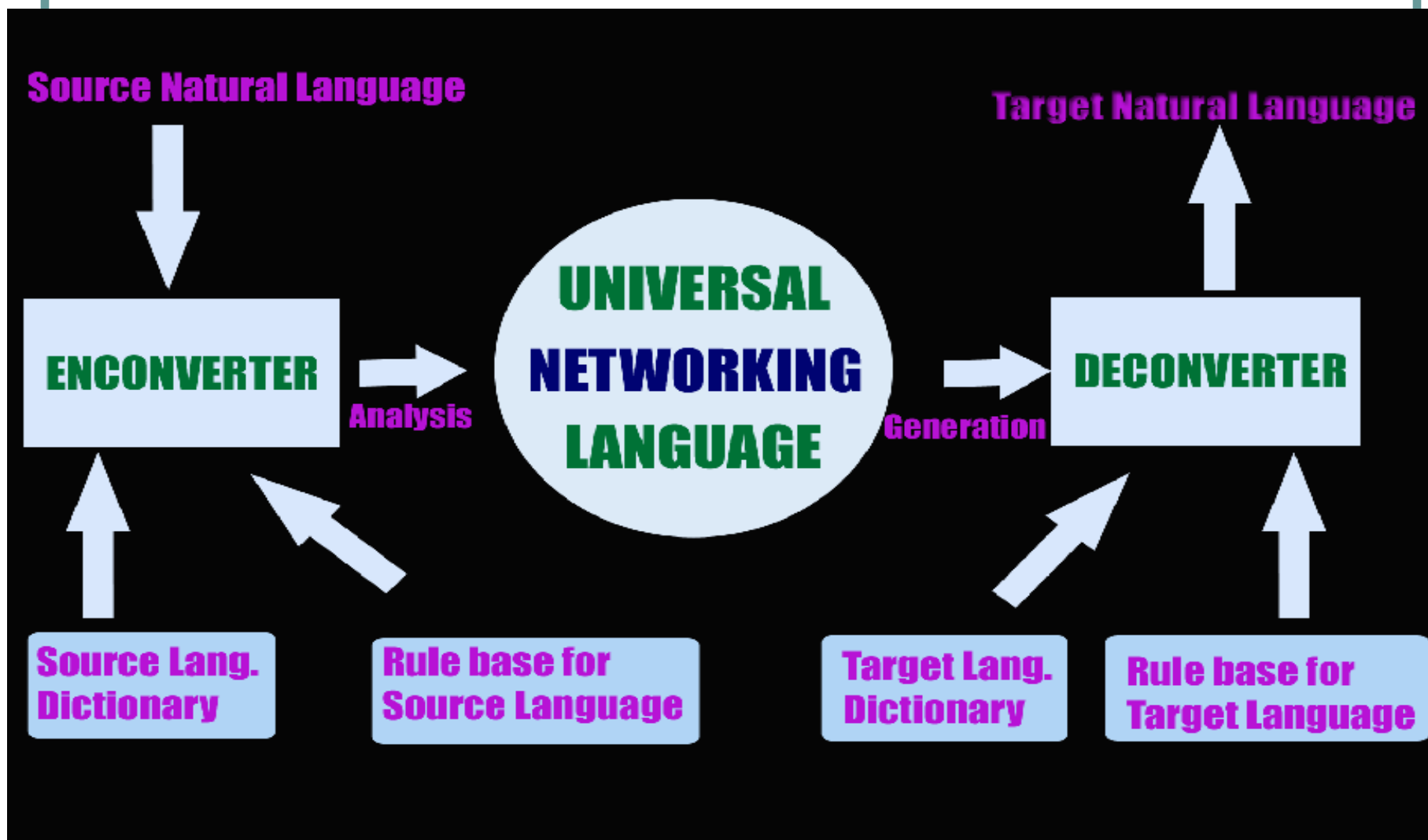
- The UNL System
- UNL: The Framework
- The Enconversion Process

Knowledge Representation



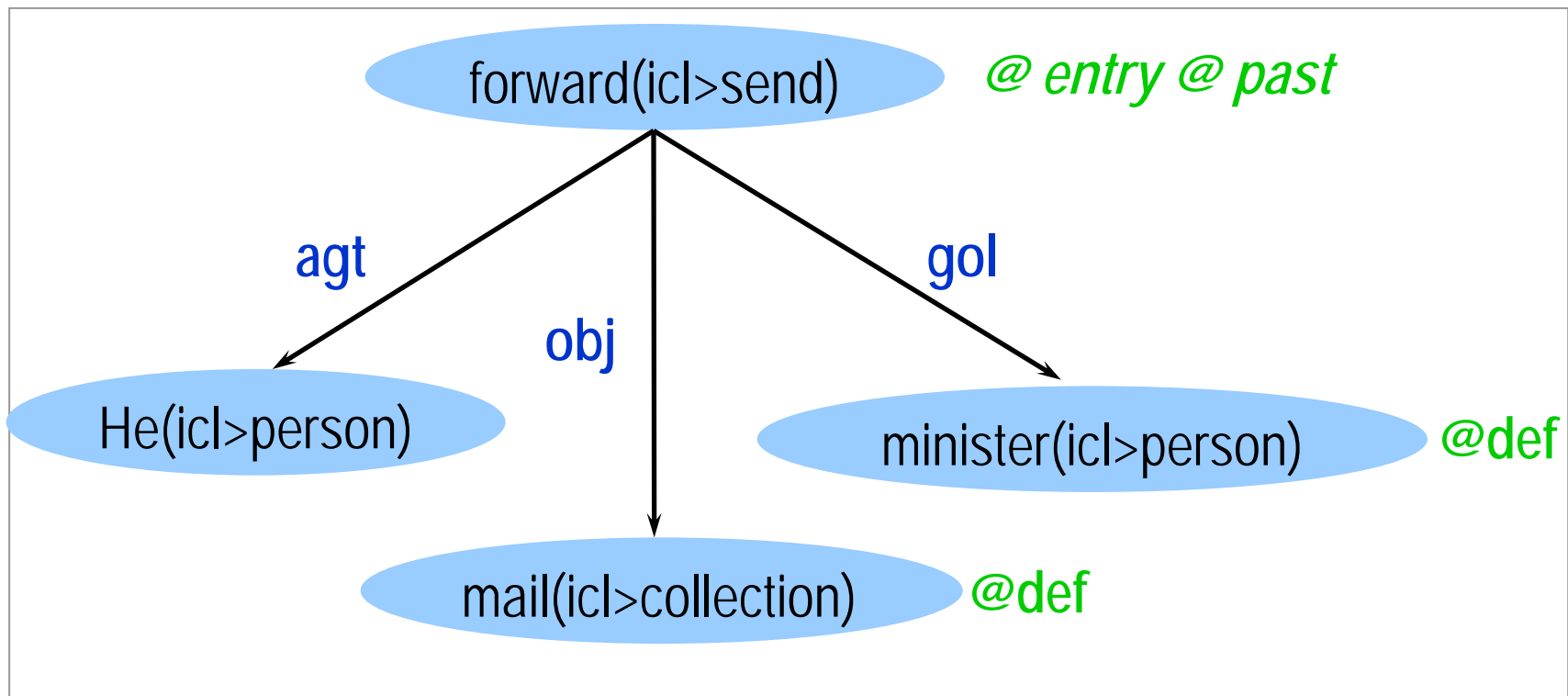
John had bought a new laptop in 2004

The UNL System: An Overview



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). **@def**)

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

The UNL Framework

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

Universal Word (UW)

- What is a Universal Word (UW)?
- What are the features of a UW?
- How to create UWs?

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 UNL Lexicon

He forwarded the mail to the minister.

Content words:

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

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

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

↑
Headword

↑
Universal Word

↑
Attributes

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)’

UNL Relations

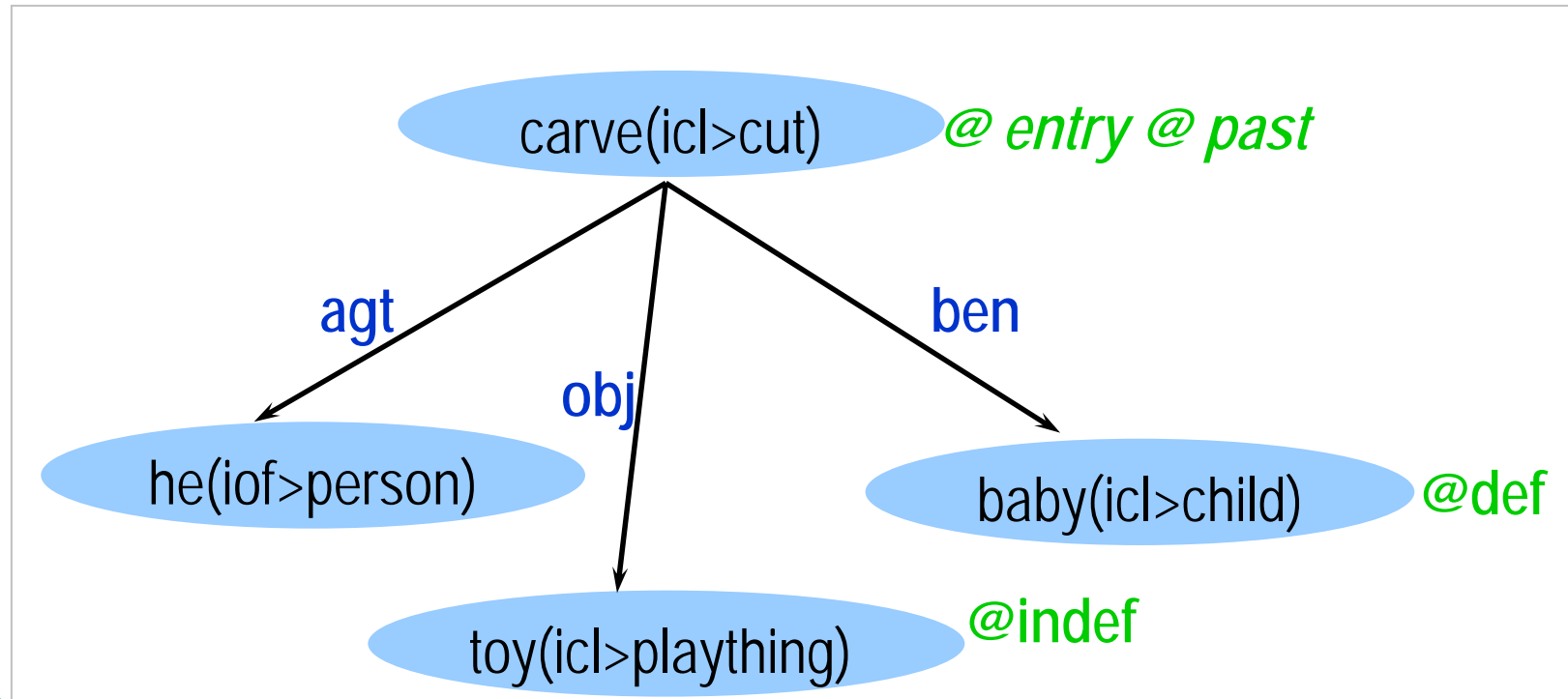
- Express how concepts(UWs) constitute a sentence
- Represented as strings of 3 characters or less
- A set of 41 relations, specified in UNL (e.g., *agt*, *aoj*, *ben*, *gol*, *obj*, *plc*, *src*, *tim*,...)
- Refer to a semantic role between two lexical items in a sentence
 - John *has* composed *this* poem.

AGT / AOJ / OBJ

- AGT (Agent) : *Agt defines a thing which initiates an action*
 - John broke the window.
agt (break. @entry. @past, John)
- AOJ (Thing with attribute): *Aoj defines a thing which is in a state or has an attribute*
 - This flower is beautiful.
aoj (beautiful. @entry, flower)
- OBJ (Affected thing): *Obj defines a thing in focus which is directly affected by an event or state*
 - He blamed John for the accident
obj (blame. @entry. @past, John)

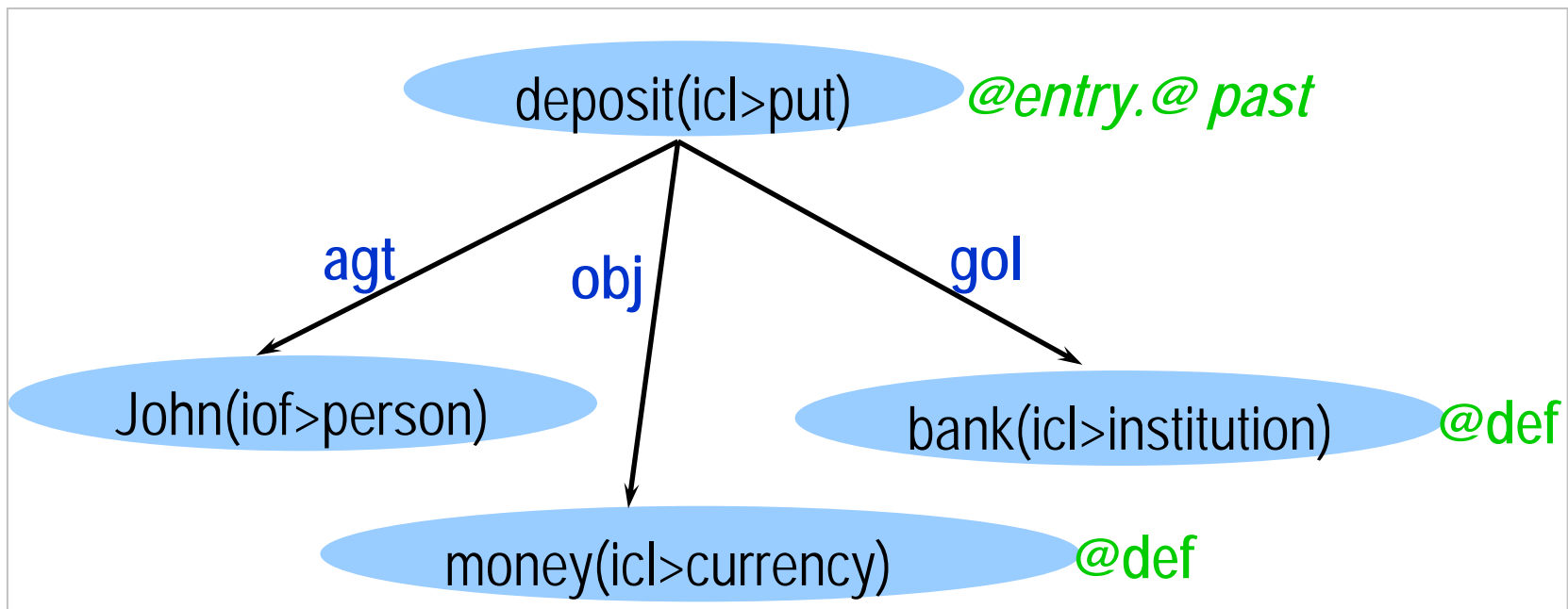
BEN (UNL Graph)

- BEN (Beneficiary) : *ben* defines a not directly related beneficiary or victim of an event or state
 - He carved a toy for the baby



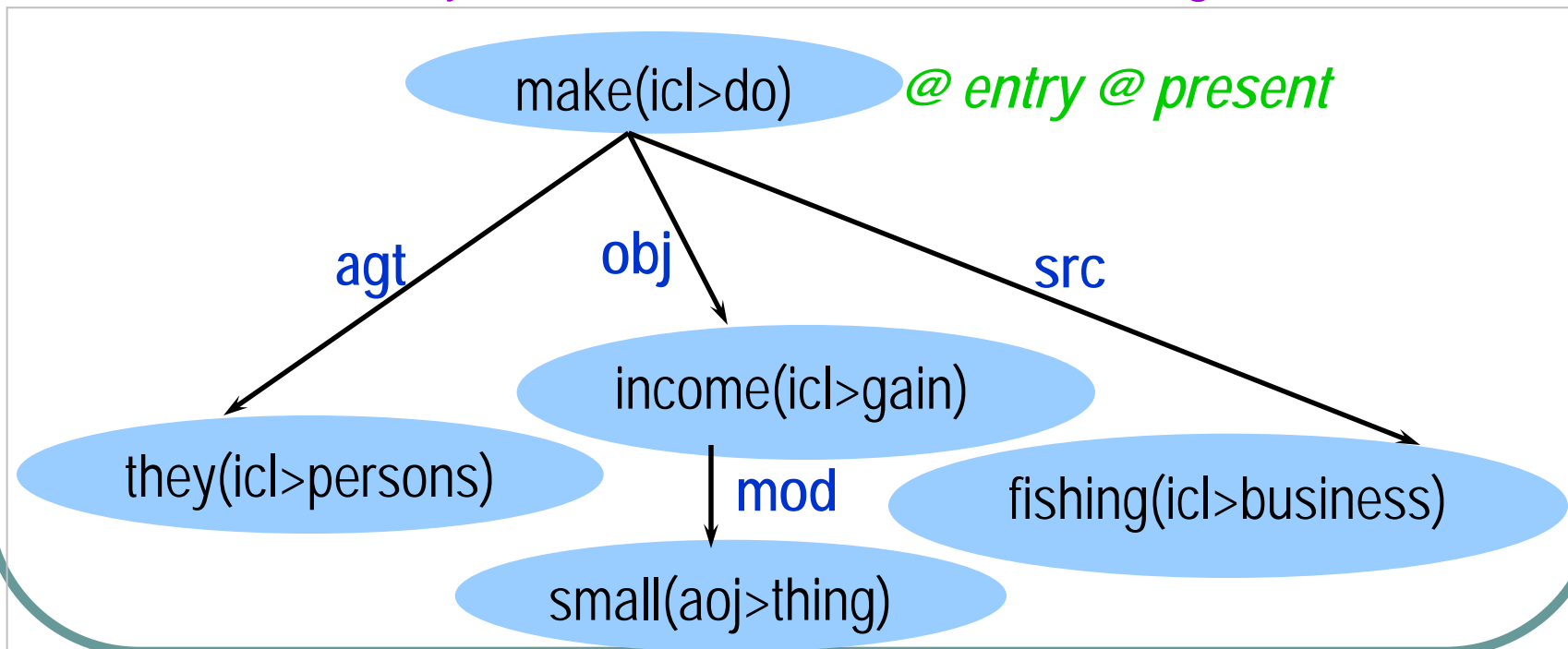
GOL

- GOL (Goal : final state): *gol* defines the final state of an object or the thing finally associated with an object of an event
 - I deposited the money in the bank



SRC

- SRC (Source : initial state): *src* defines the initial state of object or the thing initially associated with object of an event
 - They make a small income from fishing



PUR/RSN

- PUR (Purpose or objective): *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 (Reason): *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 / TMF

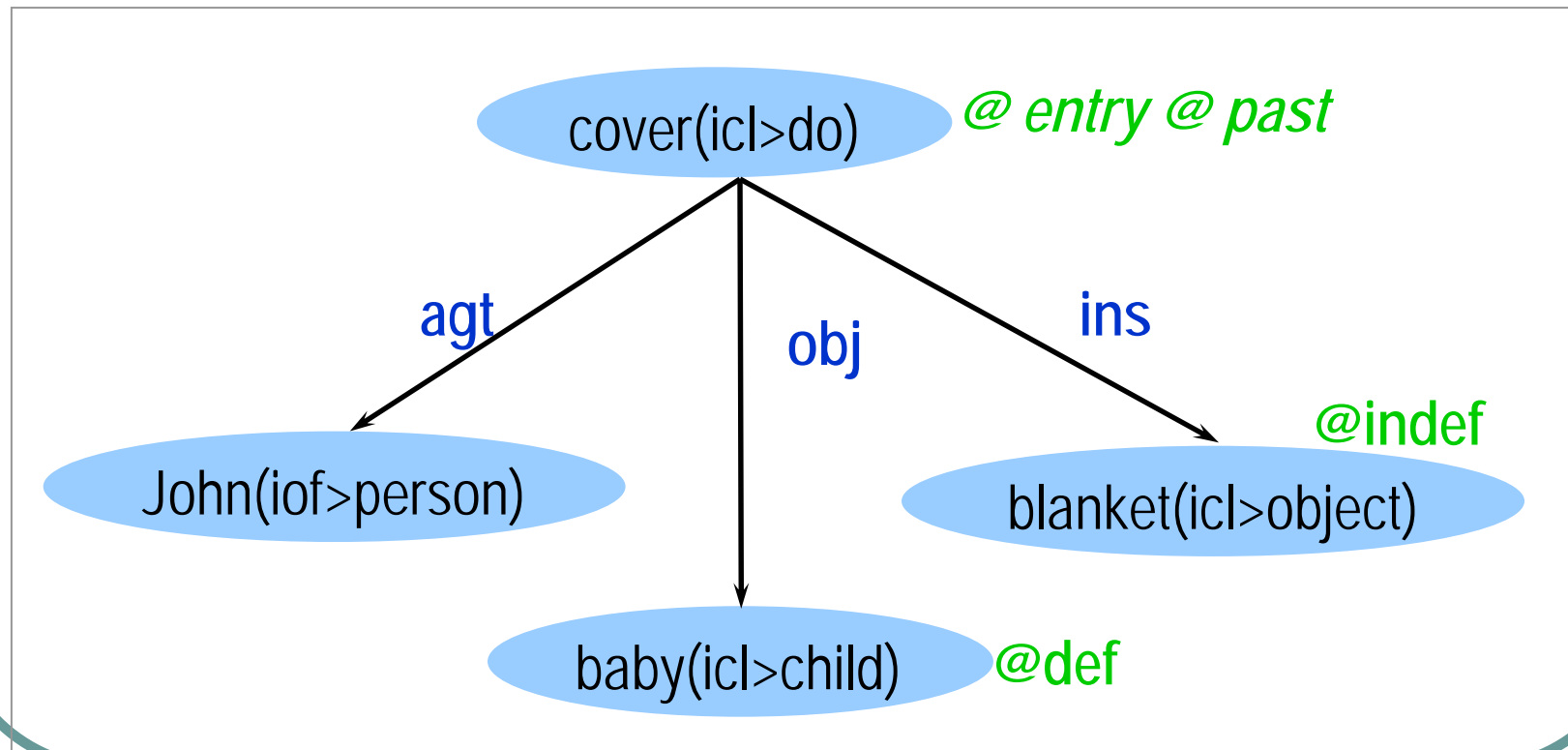
- TIM (Time): *tim* **defines the** time an event occurs or a state is true.
 - I wake up at noon.
agt (wake_up.@entry, I)
tim (wake_up.@entry, noon(icl>time))
- TMF (Initial time): *tmf* defines a time an event starts.
 - The meeting started from morning.
obj (start.@entry.@past, meeting.@def)
tmf (start.@entry.@past, morning(icl>time))
- TMT (Final time): *tmt* defines a time an event ends
 - The meeting continued till evening.
obj (continue.@entry.@past, meeting.@def)
tmt (continue.@entry.@past, evening(icl>time))

PLC / PLF / PLT

- PLC (Place): *plc* defines the place an event occurs or a state is true or a thing exists
 - I met him in India.
plc (meet. @entry, India)
- PLF (Initial place): *plf* defines the place an event begins or a state becomes true
 - Participants come from the whole world.
plf (come. @entry, world)
- PLT (Final place): *plt* defines the place an event ends or a state becomes false
 - We will go to Delhi.
agt (go. @entry. @future, we)
plt (go. @entry. @future, Delhi)

INS (UNL Graph)

- *INS* (Instrument): *ins* defines the instrument to carry out an event
 - John covered the baby with a blanket



Attributes

- Play the role of bridging the conceptual world and the real world in the UNL expressions
- Show how and when the speaker views what is said and with what intention, feeling, and *so on*.
- Seven types:
 - Time with respect to the speaker
 - Aspects
 - Speaker's view of reference
 - Speaker's emphasis, focus, topic, etc.
 - Convention
 - Speaker's attitudes
 - Speaker's feelings and viewpoints

Tense and Aspect

- *He went there yesterday.*
- The past tense is normally expressed by *@past*
 - {unl}
agt(go(icl>move).@entry.@past, he)
... {/unl}
- *He is playing chess now.*
- The progressive aspect is expressed by *@progress*
 - {unl}
obj (play(icl>compete).@entry.@present.@progress,
chess(icl>game))
... {/unl}

Speaker's view of reference

- **@def** (Specific concept (already referred))
 - The house on the corner is for sale.
 - house@def
- **@indef** (Non-specific class)
 - There is a book on the desk
 - book@indef
- **@not** is always attached to the UW which is negated.
 - He didn't come.
 - agt (come.@entry.@past.@not, he)

Speaker's emphasis

- **@emphasis**

- John his name is.

mod (name, he)

aoj (John.@emphasis.@entry, name)

- **@entry** denotes the entry point or main UW of an UNL expression

UNL Knowledge Base (UNLKB)

- What is the UNL Knowledge Base?
 - A semantic network comprising every directed binary relation between UWs
 - Categorized according to the role of a concept to other concepts
- How to define the UWs in the UNL Knowledge-Base?
 - Nominal concept (Abstract / Concrete)
 - Verbal concept (DO / OCCUR / BE)
 - Adjective concept
 - Adverbial concept

Nominal Concept: Abstract thing

```
abstract thing{(icl>thing)}
  culture(icl>abstract thing)
    civilization(icl>culture{>abstract thing})
  direction(icl>abstract thing)
    east(icl>direction{>abstract thing})
  duty(icl>abstract thing)
    mission(icl>duty{>abstract thing})
      responsibility(icl>duty{>abstract thing})
        accountability{(icl>responsibility>duty)}
  event(icl>abstract thing,{icl>time>abstract thing})
    meeting(icl>event{>abstract thing,icl>group>abstract thing})
      conference(icl>meeting{>event})
```

Nominal Concept: Concrete thing

```
concrete thing{(icl>thing,icl>place>thing)}
  building(icl>concrete thing)
    house(icl>building{>concrete thing})
  substance(icl>concrete thing)
    cloth(icl>substance{>concrete thing})
      cotton(icl>cloth{>substance})
    liquid(icl>substance{>concrete thing})

  beverage(icl>food,icl>liquid>substance})
    coffee(icl>beverage{>food})
```

Verbal Concept: *do*

do({icl>do,}agt>thing,gol>thing,obj>thing)

express({icl>do()}agt>thing,gol>thing,obj>thing{}})

state(icl>express(agt>thing,gol>thing,obj>thing))

explain(icl>state(agt>thing,gol>thing,obj>thing))

add({icl>do()}agt>thing,gol>thing,obj>thing{}})

change({icl>do()}agt>thing,gol>thing,obj>thing{}})

convert(icl>change(agt>thing,gol>thing,obj>thing)

classify({icl>do()}agt>thing,gol>thing,obj>thing{}})

divide(icl>classify(agt>thing,gol>thing,obj>thing))

Verbal Concept: *occur* and *be*

- occur({icl>occur,}gol>thing,obj>thing)
melt({icl>occur({gol>thing,obj>thing{}})})
divide({icl>occur({gol>thing,obj>thing{}})})
arrive({icl>occur({obj>thing{}})})
- be({icl>be,}aoj>thing{,^obj>thing})
exist({icl>be({aoj>thing{}})})
born({icl>be({aoj>thing{}})})

How to define the UWs in UNLKB?

- In order to distinguish among the verb classes headed by 'do', 'occur' and 'be', the following features are used:

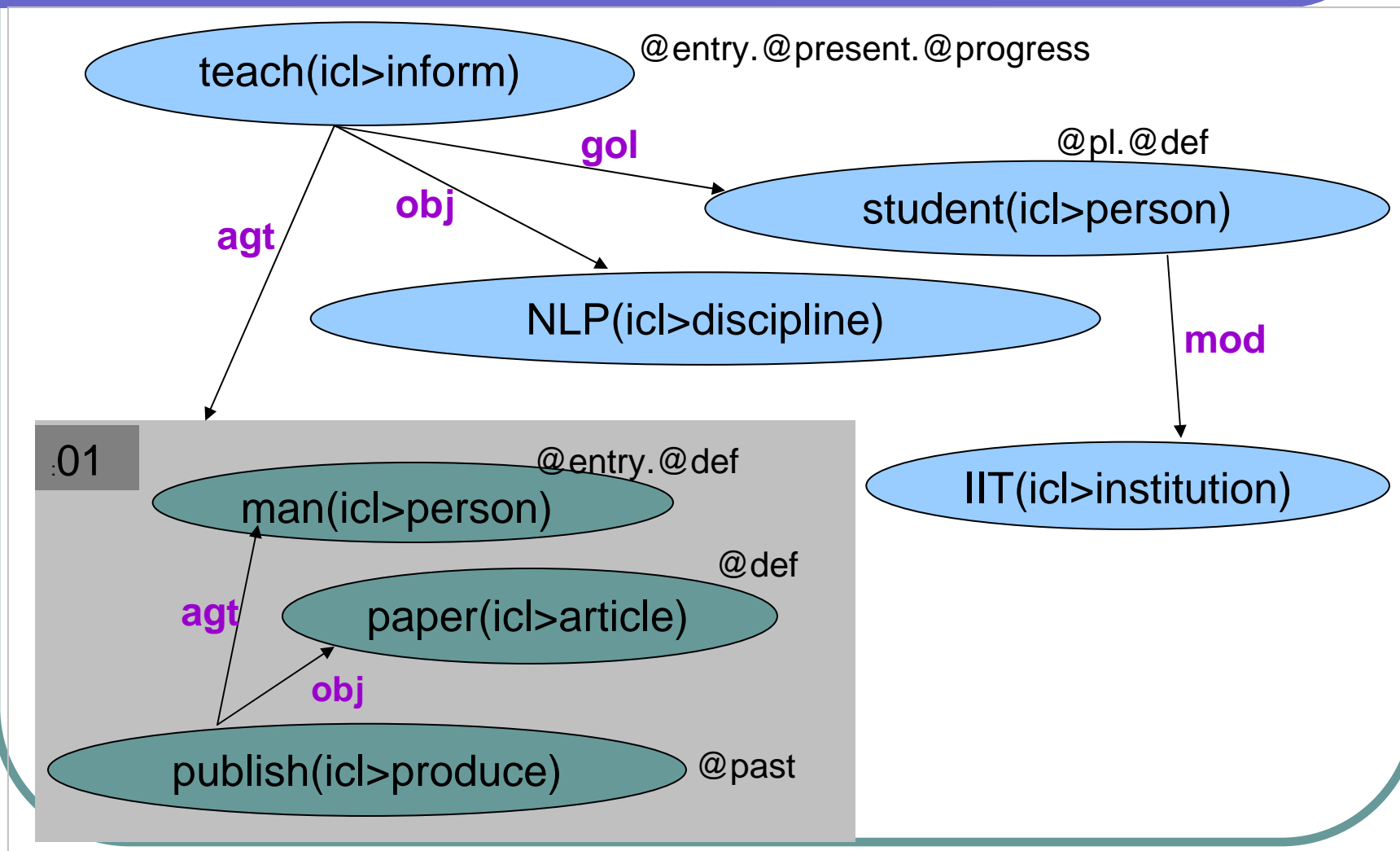
UW	[need an agent]	[need an object]	English
'do'	+	+	"to kill"
'occur'	-	+	"to fall"
'be'	-	-	"to know"

How to define the UWs in UNLKB?

- The verbal UWs (do, occur, be) also take some pre-defined semantic cases, as follows:

UW	Pre-defined Cases	English
'do'	takes necessarily agt>thing	"to kill"
'occur'	takes necessarily obj>thing	"to fall"
'be'	takes necessarily aoj>thing	"to know"

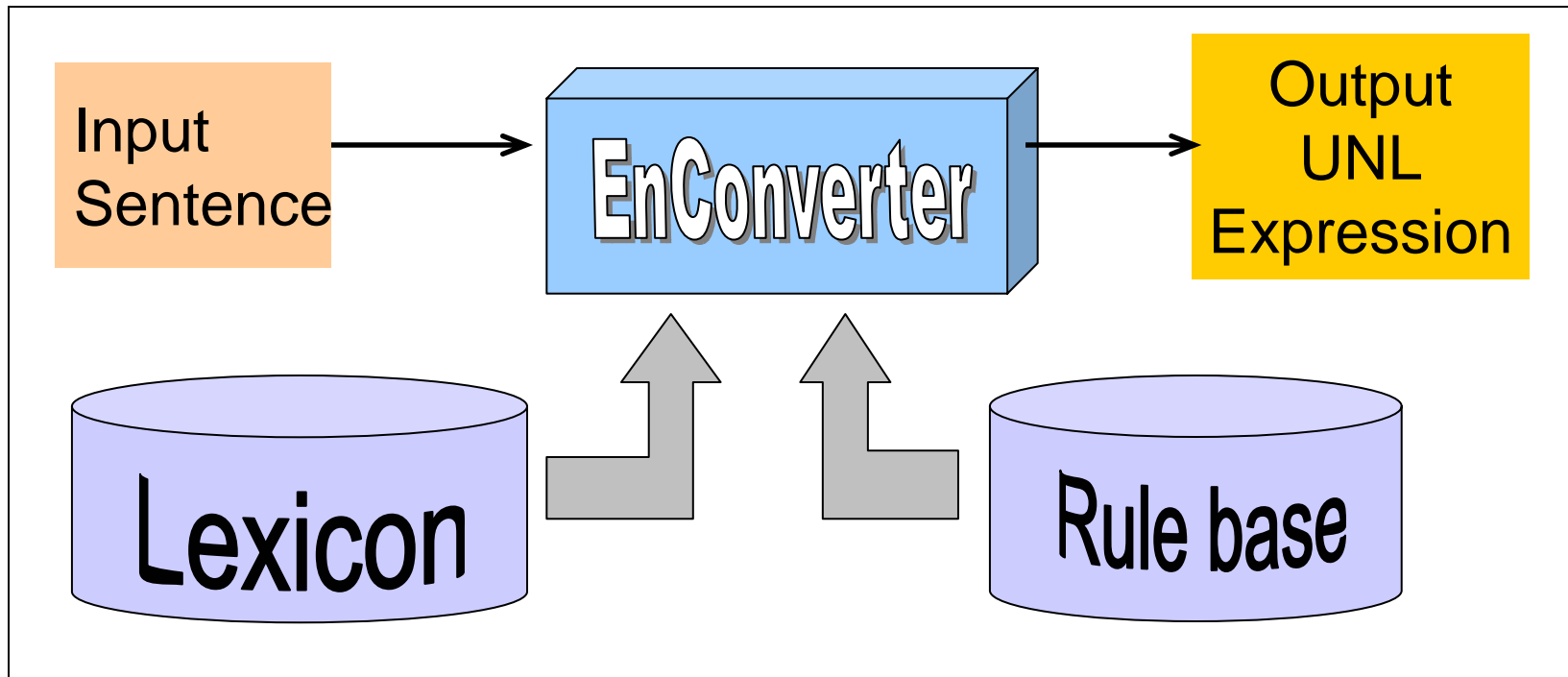
Hypergraph



Semantically Relatable Sets

- UNL: The language independent knowledge representation
 - They selected him for his honesty.
 - agt (select(icl>choose).@entry, they)
 - obj (select(icl>choose).@entry, he)
 - rsn (select(icl>choose).@entry, honesty(icl>morality))
 - mod (honesty(icl>morality), he)
- Semantically Relatable Sets: Building Blocks for Representing Semantics
 - (CW, CW) → R (UW₁, UW₂)
 - (CW, FW, CW) → R (UW₁, UW₂)
 - (FW, CW) → @attributes

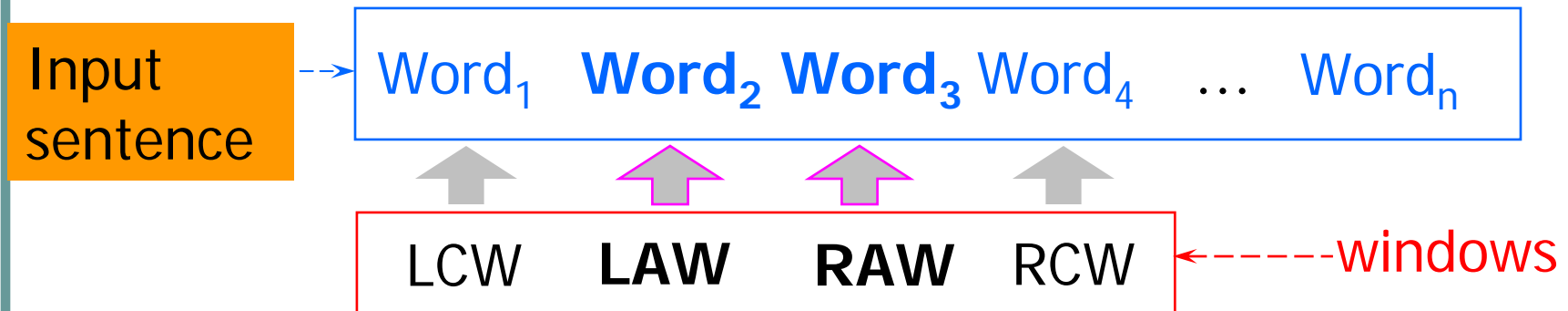
The EnConversion Process



The EnConverter

- The EnConverter
 - that works like a multi-headed Turing Machine
- Simultaneously does *Morphological*, *Syntactic* and *Semantic* processing
- Operations: Heads move over the input *to and fro*
 - Movement of heads
 - Addition of two nodes
 - Deletion of a node
 - Creating relation between two nodes
 - Adding dynamically inferred attributes to node

The Analyzer Machine



- Analysis windows - ***Two in number***
 - Left Analysis Window (LAW)
 - Right Analysis Window (RAW)
- Condition windows - ***Many in number***
 - Left Condition Window (LCW)
 - Right Condition Window (RAW)

The Lexicon

He forwarded the mail to the minister.

Content words:

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

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

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

↑
Headword

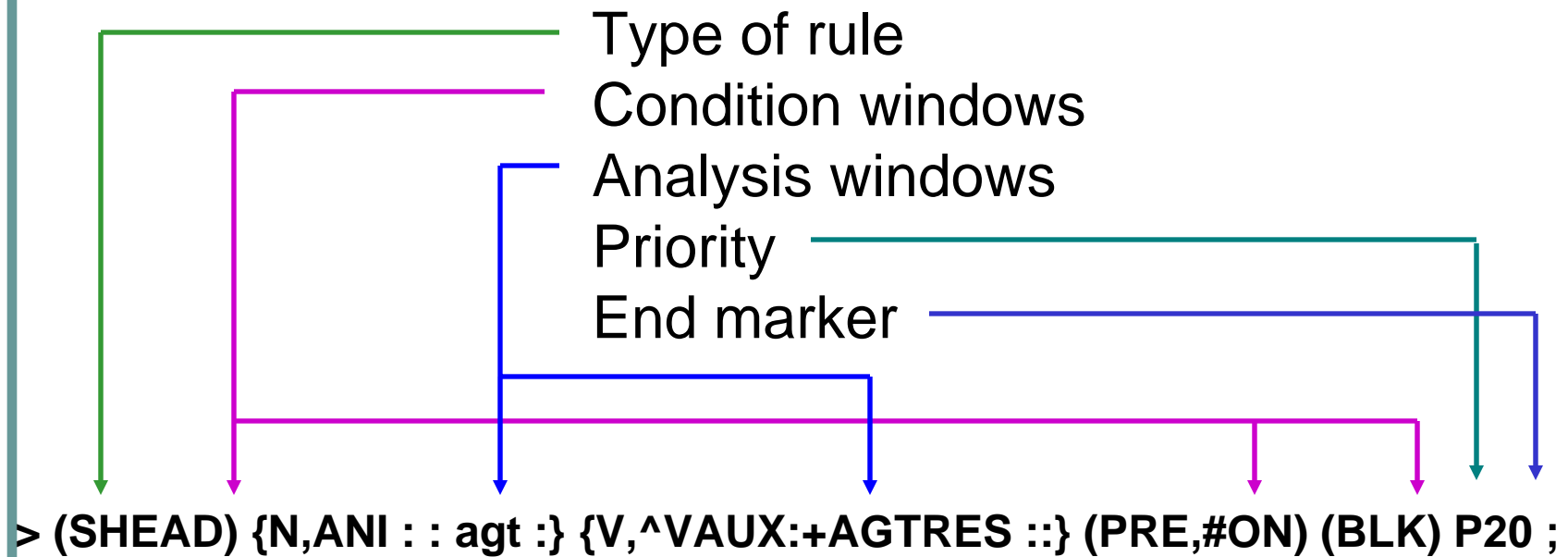
↑
Universal Word

↑
Attributes

Rule base

- If (condition) then action
 - Conditions : matching attributes or the HW or UW
 - Actions
 - Create relation
 - Add attributes
 - Move windows
- Priority
 - Indicated by number 1 to 255
 - General conditions – Low priority
 - (N) (V)
 - Priority is 20
 - Specific conditions – High priority
 - (N,ANIMT) (V,VOA)
 - Priority is 30

Rule format



The Condition and Analysis windows

- Condition windows
 - Format: (attribute₁, attribute₂, ...)
 - *E.g.*, (N, TIME)
- Analysis windows
 - Format: { conditions:+/- attributes:relation: }
 - If conditions are matched then do the action
 - *i.e.*, to add(+)/remove(-) its attributes
 - Or create the specified relation with the other analysis window
 - Or move the heads according to the type of rule

Right Shift Rule

$R\{V, VOA::: \} \{N, ANIMT, PRSN::: \} (PRE, \#OF) P60;$
;Right shift

IF

the left analysis window is on a *verb* (V) which is also a verb of action (VOA)

AND

the right analysis window is on a noun (N) which is animate ($ANIMT$) and a person ($PRSN$)

AND

the preposition-*of* follows the noun (N) indicated by ($PRE, \#OF$)

THEN

shift right (indicated by R at the start of the rule)

Right Shift Rule

Before application of the rule

... **asked** **director** of

$R\{V, VOA::: \} \{N, ANIMT, PRSN::: \} (PRE, \#OF) P60;$

After application of the rule

... **asked** **director** **of**

A Rule for a Semantic Relation

<{V,VOA,:::}{N,TIME,DAY,ONRES,PRERES::tim:}P25;

;Create relation between V and N2, after resolving the preposition preceding N2

IF

the left analysis window is on a verb(V) which is verb of action (VOA)

AND

the right analysis window is on a noun (N) and has TIME, DAY attribute for which the preceding preposition (on) has been processed and deleted

THEN

set up the tim relation between V and N₂. (indicated by < at the start of the rule)

A Rule for a Semantic Relation

Before application of rule

... came monday

`<{VRB,VOA,:::}{N,TIME,DAY,ONRES,PRERES::tim:}P25;`

After application of rule

... came ...

Conclusion and Future Direction

- In IIT Bombay, we have a long experience of generating semantic structures from text.
- It is our finding that the available semantic resources (albeit with some mistakes) can be used in a UNL generator.
- Our goal: Robust and Scalable UNL Generation
- This robust & scalable intermediate step is useful for many large applications, like
 - meaning-based search engines
 - keyword extraction
 - question answering and
 - text summarization systems
 - MT system
- We need semantically annotated corpora.

Sources and Suggested Readings

- Comrie, Bernard. 1995. *Aspects*. CUP.
- Comrie, Bernard. 1987. *Tense*. CUP.
- Dorr, Bonnie J. 1997. "Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation," *Machine Translation*, 12:4, pp. 271--322, 1997.
- Dorr, Bonnie J. 2001. LCS Database. University of Maryland.
- Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*, MIT Press.
- Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. The MIT Press.
- Jackendoff, R. 1990. *Semantic Structures*. The MIT Press, Cambridge, Mass.
- Jackendoff, R. 1997. *Semantics and Cognition*. The MIT Press, Cambridge, Mass.
- Jurafsky, D. and J. H. Martin. 2000. *Speech and Language Processing*. Prentice-Hall, New Jersey.

Sources and Suggested Readings

- Levin, Beth. 1993. "English Verb Classes and Alternations: A Preliminary Investigation", University of Chicago Press, Chicago, IL, 1993.
- Miller, G. A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. 1993. Introduction to WordNet: An On-line Lexical Database. <http://wordnet.princeton.edu/>
- Mohanty, R. K., A. Dutta, P. Bhattacharyya. 2005. Semantically Relatable Sets: Building Blocks for Representing Semantics. <http://www.cse.iitb.ac.in/~rkm/papers>
- Palmer, F.R. 2001. *Mood and Modality*. CUP
- Smith, S. Carlota. 1997. *The Parameter of Aspects*. Kluwer, Dordrecht.
- Uchida, H., M. Zhu and T. Della Senta. 1999. *UNL: A Gift for a Millennium*. The United Nations University, Tokyo. <http://www.unl.ias.unu.edu/publications/index.html>
- UNDL Foundation. 2003. The Universal Networking Language (UNL), Specifications, Version 3 Edition 3.



Thank You