## Computational Lexical Semantics

Martha Palmer

#### **Vilem Mathesius Lecture Series 21**

Charles University, Prague

December, 2006





Complete representation of real world knowledge -Natural Language Understanding?



Only build useful representations for small vocabularies

 Major impediment to accurate Machine Translation, Information Retrieval and Question Answering



Ask Jeeves – A Q/A, IR ex.

What do you call a successful movie? Blockbuster

- Tips on Being a Successful Movie Vampire ... I shall call the police.
- Successful Casting Call & Shoot for ``Clash of Empires'' ... thank everyone for their participation in the making of yesterday's movie.
- Demme's casting is also highly entertaining, although I wouldn't go so far as to call it successful. This movie's resemblance to its predecessor is pretty vague...
- VHS Movies: Successful Cold Call Selling: Over 100 New Ideas, Scripts, and Examples from the Nation's Foremost Sales Trainer.



## Ask Jeeves – filtering w/ POS tag

What do you call a successful movie?

- Tips on Being a Successful Movie Vampire ... I shall call the police.
- Successful Casting Call & Shoot for `Clash of Empires' thank everyone for their participation in the making of yesterday's movie.
- Demme's casting is also highly entertaining, although I wouldn't go so far as to call it successful. This movie's resemblance to its predecessor is pretty vague...
- VHS Movies: Successful Cold Call Selling: Over 100 New Ideas, Scripts, and Examples from the Nation's Foremost Sales Trainer.



### Filtering out "call the police"

Different senses,

- different syntax,
- different kinds of participants,
- different types of propositions.





#### Outline

- Linguistic Theories of semantic representation
  - Case Frames Fillmore FrameNet
  - Lexical Conceptual Structure Jackendoff LCS
  - Proto-Roles Dowty PropBank
  - English verb classes (diathesis alternations) -Levin - VerbNet
  - **•** Talmy, Levin and Rappaport
- Manual Semantic Annotation
- Automatic Semantic annotation



The Case for Case Charles J. Fillmore in E. Bach and R.T. Harms, eds. Universals in Linguistic Theory, 1-88. New York: Holt, Rinehart and Winston.

#### Thanks to Steven Bethard



#### Case Theory

Case relations occur in deep-structure
 Surface-structure cases are derived

A sentence is a verb + one or more NPs

- Each NP has a deep-structure case
  - A(gentive)
  - I(nstrumental)
  - D(ative)
  - F(actitive)
  - L(ocative)
  - O(bjective)

Subject is no more important than Object

Subject/Object are surface structure



#### Case Selection

#### Noun types

- Different cases require different nouns
- □ E.g.  $N \rightarrow$  [+animate]/<sup>A,D</sup>[X\_Y]

#### Verb frames

- Verbs require arguments of particular cases
- E.g.
  - sad [ \_\_D]
  - give [ \_\_\_O+D+A]
  - open [ \_\_O(I)(A)]



#### Case Theory Benefits

#### Fewer tokens

- Fewer verb senses
- E.g. cook [ \_\_O(A)] covers
  - Mother is cooking the potatoes
  - The potatoes are cooking
  - Mother is cooking
- Fewer types
  - "Different" verbs may be the same semantically, but with different subject selection preferences
  - E.g. like and please are both [ \_\_O+D]
- Only noun phrases of the same case may be conjoined
  - \*John and a hammer broke the window
  - The car broke the window with a fender



#### Case Theory Drawbacks

- How can a handful of cases cover every possible type of verb argument?
  - Is an agent always animate? Always volitional?
  - Is an instrument always an artifact?
- What are the mapping rules from syntax to semantics?



#### FrameNet

- Baker, Collin F., Charles J. Fillmore, and John B. Lowe. (1998) The Berkeley FrameNet project. In *Proceedings of COLING/ACL-98*, pages 86--90, Montreal.
- Fillmore, Charles J. and Collin F. Baker. (2001). Frame semantics for text understanding. In the *Proceedings of NAACL WordNet and Other Lexical Resources Workshop* Pittsburgh, June.



#### Introducing FrameNet

Thanks to Chuck Fillmore and Collin Baker

In one of its senses, the verb *observe* evokes a frame called **Compliance**: this frame concerns people's responses to norms, rules or practices.

The following sentences illustrate the use of the verb in the intended sense:

• Our family **observes** the Jewish dietary laws.

- □ You have to **observe** the rules or you'll be penalized.
- □ How do you **observe** Easter?
- □ Please **observe** the illuminated signs.



#### FrameNet

FrameNet records information about English words in the general vocabulary in terms of

- the frames (e.g. Compliance) that they evoke,
- 2. the **frame elements** (semantic roles) that make up the components of the frames (in **Compliance**, Norm is one such frame element), and
- a. each word's valence possibilities, the ways in which information about the frames is provided in the linguistic structures connected to them (with observe, Norm is typically the direct object).



#### The FrameNet Product

The FrameNet database constitutes

- a set of frame descriptions
- a set of corpus examples annotated with respect to the frame elements of the frame evoked by each lexical unit
- Iexical entries, including definitions and displays of the combinatory possibilities of each lexical unit, as automatically derived from the annotations
- a display of frame-to-frame relations, showing how some frames are elaborations of others, or are components of other frames.



### Frame Elements for Compliance

The frame elements that figure in the Compliance frame are called

- Norm (the rule, practice or convention)
- Protagonist (the person[s] reacting to the Norm)
- Act (something done by the Protagonist that is evaluated in terms of the Norm)
- State\_of\_affairs (a situation evaluated in terms of the Norm)



You do a whole frame for just *observe*?No. There are other Compliance words too.

V - adhere, comply, conform, follow, heed, obey, submit, ...;

#### AND NOT ONLY VERBS

- N adherence, compliance, conformity, obedience, observance, ...;
- A compliant, obedient, ...;
- PP in compliance with, in conformity to, ...;

AND NOT ONLY WORDS FOR POSITIVE RESPONSES TO NORMS

- V break, disobey, flout, transgress, violate ,...;
- N breach, disobedience, transgression, violation,...;

PP - in violation of, in breach of, ...



#### Tagging Compliance sentences



- Are we finished with the verb *observe*?
- No. This verb has several other meanings too.
- In the Perception\_active frame we get the uses seen in observing children at play, observing an ant colony, sharing frame membership with watch, attend, listen to, view & pay attention.
- In a Commenting frame, observe and observation share frame membership with remark & comment.



#### Lexical Unit

Our unit of description is not the word (or "lemma") but the **lexical unit** (Cruse 1986), – <u>a</u> <u>pairing of a word with a sense</u>. In our terms this is <u>the pairing of a word with a single frame</u>.

The lexical unit - roughly equivalent to a word in a synset - is the unit in terms of which important generalizations about lexical relations, meanings and syntactic behavior can best be formulated.



#### LUs and V-N relationships

- Note that the nouns based on observe are
  - observance in the Compliance frame,
  - observation in the Perception\_active frame
- Similarly, the nouns based on *adhere* are
  - □ adherence in the Compliance frame,
  - adhesion in the Attachment frame.
- When we need to be precise we show the framespecific sense of a lemma (the full name of an LU) with a dotted expression:
  - Compliance.observe, Attachment.adhere, etc.



#### words, frames, lexical units



2 lexical units sharing same form: Compliance.*observe*, Perception.*observe* 



#### words, frames, lexical units



2 lexical units sharing the same form: Compliance.*adhere*, Attachment.*adhere* 



The study of polysemy concerns membership in different frames





#### Different LU, Different Valence

**Compliance**.*observe* generally has an NP as its direct object.

**Perception**.*observe* has these patterns:

- □ NP: Observe the clouds overhead.
- NP+Ving: I observed the children playing.
- wh-clause: Observe what I'm doing.
- that-clause: We observed that the process terminated after an hour.
- **Comment**.*observe* occurs frequently with a quoted comment:
  - "That was brilliant," he observed snidely.



## Lexical-units: Wrap-up

Lexical units are the entities with respect to which we define

- meanings
- grammatical behavior
- semantic relations with other entities
- morphological relations with other entities

In short, there aren't interesting things to say about the verb *observe* in general, but only about the individual lexical units that happen to have the form *observe*.



Jackendoff: Lexical Conceptual Structures from Jackendoff, R.S., Towards an Explanatory Semantic Representation, *Linguistic Inquiry*, 7:1, pp. 89-150, 1976.



Semantic Decomposition

# Markers HORSE the red horse RED

Functions
 SEE(x,y) the man saw the (red) horse
 SEE(x,HORSE)
 SEE(THE MAN,THE HORSE)
 SEE(X1, Y1)
 (What is the value? predicates? )



#### Five Semantic Functions

- GO
- BE
- STAY
- LET
- CAUSE



#### GO – Change of location

The train traveled from Detroit to Cincinatti. The hawk flew from its next to the ground. An apple fell from the tree to the ground. The coffee filtered from the funnel into the cup.

GO (x,y,z) THROUGH THE AIR/DOWNWARD

#### THEME GOES FROM SOURCE, TO GOAL

#### Mapping from Syntax to Semantics





#### BE – Stationary location

Max is in Africa. The vine clung to the wall. The dog is on the left of the cat. The circle contains/surrounds the dot?

BE(x,y) THEME IS AT LOCATION

#### BE (THE DOG, LEFT OF (THE CAT))



#### STAY – Durational stationary location

The bacteria stayed in his body. Stanley remained in Africa. Bill kept the book on the shelf.

STAY(x,y) THEME IS AT LOCATION for a duration

#### STAY (STANLEY, AFRICA) (for two years)



Locational modes: POSIT, POSS, ID

## The train traveled from Detroit to Cincinatti. $\begin{bmatrix} GO(x,y,z) \\ POSIT \end{bmatrix}$

Harry gave the book to the library. GO (x,y,z) POSS The book belonged to the library..



Locational modes: POSIT, POSS, ID

#### The bacteria stayed in his body. STAY (x,z) POSIT

The library kept the book. STAY (x,z) POSS



#### Locational modes: POSIT, POSS, ID

\**The coach changed from a handsome young man to a pumpkin.* [GO<sub>IDENT</sub> (x,y,z)]

Princess Mia changed from an ugly duckling into a swan. [GO<sub>IDENT</sub> (x,y,z)] Universal grammar?


Causation and Permission: CAUSE and LET

# *The rock fell from the roof to the ground.* [GO<sub>POSIT</sub> (x,y,z)]

# *Linda* lowered the rock from the roof to the ground. [CAUSE (a, GO<sub>POSIT</sub> (x,y,z))]

*Linda* dropped the rock from the roof to the ground. [LET (a, GO<sub>POSIT</sub> (x,y,z))]



#### INSTRUMENTS

*Linda* lowered the rock from the roof to the ground with a cable.

CAUSE (a, GO<sub>POSIT</sub> (x,y,z))

Instruments only occur with causation. CAUSE always has an *event* second argument.



#### Lexical Conceptual Structure

concept	POSIT	POSS	IDENT
GO	go	receive	become
motional	fall	inherit	change
BE	be	have	be
punctual	contain	own	seem
STAY	stay	keep	stay
durational	remain		remain
CAUSE(a,GO)	bring, take	obtain, give	make,elect
CAUSE(a,STAY)	keep, hold	keep, retain	keep
LET(a,GO)	drop,release	accept,	CSLP
LET(a,BE)	leave, allow	fritter, permit	leav

Rules of inference

#### CAUSE(a, event) -> event.



# Machine Translation: Interlingual Methods

# Bonnie J. Dorr, Eduard H. Hovy, Lori S. Levin

Thanks to Les Sikos



- What is Machine Translation (MT)?
  - Automated system
  - Analyzes text from Source Language (SL)
  - Produces "equivalent" text in Target Language (TL)
  - Ideally without human intervention





- Three main methodologies for Machine Translation
  - Direct
  - Transfer
  - Interlingual



Three main methodologies for Machine
Translation



Figure 1: The Vauquois Triangle for MT

•

Three main methodologies for Machine Translation Interlingua Semantic Semantic Direct Decomposition Composition Transfer Semantic Semanino C..... 9<sub>-</sub> mania Semantic Androis \_\_\_\_ Interlingual Seniacie Svntactic i a mine inc Transf-r Word Word Structure Structure Direct Morphological Morphological Generation Analysis Target Text Source Text

Figure 1: The Vauquois Triangle for MT

Ŧ

Three main methodologies for Machine
 Translation



Figure 1: The Vauquois Triangle for MT

Ŧ

#### Interlingua

 Single underlying representation for both SL and TL

#### which ideally

- Abstracts away from language-specific characteristics
- Creates a "language-neutral" representation
- Can be used as a "pivot" representation in the translation



- Cost/Benefit analysis of moving up the triangle
  - Benefit
    - Reduces the amount of work required to traverse the gap between languages
  - Cost
    - Increases amount of analysis
      - Convert the source input into a suitable pre-transfer representation
    - Increases amount of synthesis
      - Convert the post-transfer representation into the final target surface form



- Two major advantages of Interlingua method
  - The more target languages there are, the more valuable

an Interlingua becomes





- Two major advantages of Interlingua method
  - 2. Interlingual representations can also be used by NLP systems for other multilingual applications



Figure 2: Use of Interlingua in Multiple Applications



Sounds great, but...due to many complexities

- Only one interlingual MT system has ever been made operational in a commercial setting
  - KANT (Nyberg and Mitamura, 1992, 2000; Lonsdale *et al.*, 1995)
- Only a few have been taken beyond research prototype



#### Loss of Stylistic Elements

#### Because representation is independent of syntax

- Generated target text reads more like a paraphrase
- Style and emphasis of the original text are lost
- Not so much a failure of Interlingua as incompleteness
  - Caused by a lack of understanding of discourse and pragmatic elements required to recognize and appropriately reproduce style and emphasis
  - In some cases it may be an advantage to ignore the author's style
    - Outside the field of artistic texts (poetry and fiction) syntactic form of source text is superfluous



#### Loss of Stylistic Elements

Current state of the art

 It is only possible to produce reliable interlinguas between language groups (e.g., Japanese – Western European) within specialized domains



#### Linguistic Divergences

#### Structural differences between languages

- Categorical Divergence
  - Translation of words in one language into words that have different parts of speech in another language
    - To be jealous
    - Tener celos (To have jealousy)



- Conflational Divergence
  - Translation of *two or more words* in one language *into* one word in another language
    - To kick
    - Dar una patada (Give a kick)



- Structural Divergence
  - Realization of verb arguments in *different syntactic configurations* in different languages
    - To enter the house
    - Entrar en la casa (Enter in the house)



- Head-Swapping Divergence
  - Inversion of a structural-dominance relation between two semantically equivalent words
    - To run in
    - Entrar corriendo (Enter running)



- Thematic Divergence
  - Realization of verb arguments that reflect *different* thematic to syntactic *mapping* orders
    - I like grapes
    - *Me gustan uvas (To-me please grapes)*



- Linguistic Divergences may be the norm rather than the exception
  - Differences in MT architecture (direct, transfer, interlingual) are crucial for resolution of cross-language divergences
    - Interlingua approach takes advantage of the compositionality of basic units of meaning to resolve divergences



#### For example:

To kick – Dar una patada (Give a kick)

- Conflational divergence can be resolved by mapping English *kick* into two components before translating into in Spanish
  - Motional component (movement of the leg)
  - Manner component (a kicking motion)



KANT system (Nyberg and Mitamura, 1992)

- Only interlingual MT system that has ever been made operational in a commercial setting
  - Caterpillar document workflow (mid-90s)
- Knowledge-based system
- Designed for translation of technical documents written in Caterpillar Technical English (CTE) to French, Spanish, and German
- Controlled English no pronouns, conjunctions,...



- Pangloss project (Frederking et al., 1994)
  - Ambitious attempt to build rich interlingual expressions
  - Uses humans to augment system analysis
  - Representation includes a set of frames for representing semantic components, each of which
    - Are headed by a unique identifier
    - And have a separate frame with aspectual information (duration, telicity, etc.)
  - Some modifiers are treated as scalars and represented by numerical values



- Mikrokosmos (Mahesh and Nirenburg, 1995) / OntoSem (Nirenburg and Raskin, 2004)
  - Focus is to produce semantically rich Text-Meaning Representations (TMRs) of text
  - TMRs use a language-independent metalanguage also used for static knowledge resources
  - TMRs aimed at the most difficult problems of NLP
     Disambiguation, reference resolution
  - Goal is to populate a fact repository with TMRs as a language-independent search space for questionanswering and knowledge-extraction applications



#### PRINCITRAN (Dorr & Voss, 1996)

 Approach assumes an interlingua derived from lexical semantics and predicate decomposition
 Jackendoff 1983, 1990; Levin & Rappaport-Hovav 1995a, 1995b
 Has not complicated, but rather facilitated, the identification and construction of systematic relations at the interface between each level



- Motivation for Non-Uniform Approach
  - German: Der Berg liegt im Suden der Stadt
  - Ambiguous in English:
    - The mountain lies in the south of the city
    - The mountain lies **to** the south of the city
  - In other words, the German phrase maps to two distinct representations



Using Default knowledge in the KR

- Mountains are physical entities, typically distinct and external to cities
- System chooses second translation
  - → The mountain lies to the south of the city
- Using specific facts in the KR
  - □ A particular mountain is in the city
  - System overrides default knowledge and chooses first translation

The mountain lies in the south of the city



 The need to translate such sentences accurately is a clear case of where general as well as specific real-world knowledge should assist in eliminating inappropriate translations
 Knowledge Representational level, not the Interlingual level, provides this capability in this model



- Lexical Conceptual Structure (LCS)
  - Used as part of many MT language pairs including ChinMT (Habash et al., 2003a)
    - Chinese-English
  - Also been used for other natural language applications
    - Cross-language information retrieval



 Lexical Conceptual Structure (LCS)
 Approach focuses on linguistic divergences
 For example – Conflational divergence Arabic: The reporter caused the email to go to Al-Jazeera in a sending manner.
 English: The reporter emailed Al-Jazeera.



#### LCS representation

(event cause
 (thing[agent] reporter+)
 (go loc
 (thing[theme] email+)
 (path to loc
 (thing email+)
 (position at loc (thing email+) (thing[goal] aljazeera+)))
 (manner send+ingly)))



#### LCS representation

(event fails)
(thing[agent] reporter+)
loc
(thing[theme] email+)
(path to loc
 (thing email+)
 (position at loc (thing email+) (thing[goal] aljazeera+)))
(manner send+ingly)))

Primary components of meaning are the top-level conceptual nodes cause and conceptual nodes



# Current Efforts LCS representation

(event cause	
(thing[agen	(reporter+)
loc	
(thing[	heme email+)
(path to	
(th	ing email+)
(pe	osition at loc (thing email+) (thing[
(manne	er send+ingly)))

- Primary components of meaning are the top-level conceptual nodes cause and go
- These are taken together with their arguments
  - Each identified by a semantic role (agent, theme, goal)

And a modifier (manner) send+ingly


# LCS as an interlingua?

- Jackendoff wasn't trying to capture all of meaning – just the semantics that corresponds to syntactic generalizations
- Ch-of-loc, causation, states, ... are very fundamental. If we don't get anything else, we should get at least these
- LCS highlights just these relations not bad for an interlingua, but what about those stylistic things, etc?



## Current Efforts

- Approximate Interlingua (Dorr and Habash, 2002)
  - Depth of knowledge-based systems is approximated
  - Taps into the richness of resources in one language (often English)
  - This information is used to map the sourcelanguage input to the target-language output



## Current Efforts

Approximate Interlingua (Dorr and Habash, 2002)
Focus on linguistic divergences but with fewer

- knowledge-intensive components than in LCS
- Key feature
  - Coupling of basic argument-structure information with some, but not all, components the LCS representation
  - Only the top-level primitives and semantic roles are retained
- This new representation provides the basis for generation of multiple sentences that are statistically pared down – ranked by TL constraints



## Current Efforts

#### Approximate Interlingua representation:



Check top-level conceptual nodes for matches

- Check unmatched thematic roles for 'conflatability'
  - Cases where semantic roles are absorbed into other predicate positions
- Here there is a relation between the conflated argument EMAIL<sub>N</sub> and EMAIL<sub>V</sub>

