An Empirical View on Semantic Roles Part II

Katrin Erk
Sebastian Pado
Saarland University
ESSLLI 2006

Structure

1. History of Semantic Roles
2. Contemporary Frameworks
3. Difficult Phenomena (from an empirical perspective)
4. Role Semantics vs. Formal Semantics
5. Cross-lingual aspects

Background

- Early 1990s: Empirical turn in computational linguistics
  - Increasing focus on *data*
    - Validation of theories
    - Data-driven learning of statistical models
- Required: annotated training data
  - Parts of Spech: BNC
  - Syntax: Penn Treebank

What about a corpus with (role) semantics?

Methodological issues

- Exhaustiveness
  - Annotation has to be **broad-coverge**
  - How to handle controversial cases? (Cf. parts 1 and 3)
- Consistency
  - Intuitions have to be operationalised in the form of annotation guidelines
- Direction of inquiry
  - Bottom-up: data-driven
  - Top-down: theory-driven

Goals

- Framework for lexical semantics
  - Describe (and model) meaning of predicates
- Semantic role labelling: Annotate free text with semantic roles
  - Replace grammatical categories like SUBJ, OBJ with semantically motivated categories

Empirical / NLP-oriented twist on 70s goals

What we will look at

- Three Phenomena from part 1:
  - Do analyses generalise over alternations?
    - "Uniform basis" for data acquisition
  - Do analyses provide semantic properties?
    - "Computing the meaning"
  - How regular is the linking these analyses provide?
    - Suitability for computational modelling: Required for automatic processing of free text for NLP purposes

The three main frameworks

- Currently: three important frameworks with large annotated corpora
  1. "Praguian roles"
     - Tectogrammatical (Semantic) layer of Functional Generative Description (FGD)
     - Corpus: Prague Dependency Treebank (Czech)
  2. PropBank
     - Surface-oriented role framework
     - Corpus: Penn Treebank
  3. Frame Semantics
     - Usage-oriented theory of predicate meaning
     - "Corpus": FrameNet examples

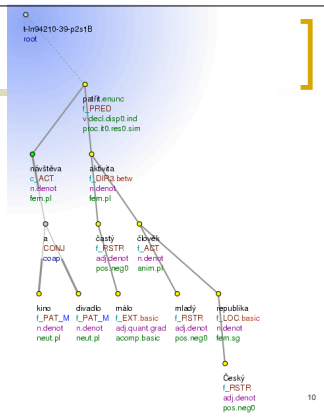
Functional Generative Description

- Dependency-based theory of language
  - Top-down approach
- Stratified structure:
  1. Surface syntax
  2. Analytical structure (=surface dependencies)
  3. **Tectogrammatical structure**
     - "Literal meaning of sentence"
     - Interface between linguistics (FDG) and interpretation/discouse
     - **Semantic role-like representation**

The Prague Dependency Treebank

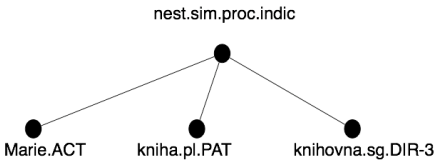
- 1M words
- Language: Czech
- Genre: Newspaper (60%), newswire and magazine (20% each)
- Specification of tectogrammatical level:
  - "Deep" trees
    - Every node = one content word
  - Roles (called functors) form part of node label
    - More detailed information provided by "grammatemes"

Example



Example

Marie	nese	knihy	do	knihovny
Marie	is carrying	the books	to	the library



Functor classification

- Inner participants vs. free modifiers:
  - Inner participants (Arguments)
    - **May not** occur more than once
    - **Prototypically** obligatory
    - „Semantically vague“
    - Occur with limited class of predicates
  - Free modifiers (Adjuncts)
    - **May** occur more than once
    - **Prototypically** optional
    - „Semantically homogeneous“
    - Occur with all predicates

Inner Participants (IPs)

- 5 IPs: Actor, Addressee, Effect, Origin, Patient
- Syntacto-semantic motivation
  - Verbs with one IP (Nominative): Actor
  - Verbs with two IPs (Nom, Acc): Actor, Patient
  - More than two: semantic considerations
- Semantic vagueness: Theory of „shifting“
  - Actors assume semantic properties in context of specific predicate

Free Modifiers (FMs)

- About 70
  - Temporal, Manner, Regard, Extent, Norm, Criterion, Substitution, Accompaniment, etc. pp.
  - Mostly realised by specific prepositional phrases
  - Well-defined semantic contribution

IPs vs. FMs

- Dichotomy between IPs and FMs problematic
  - IPs:
    - **May not** occur more than once, **Prototypically** obligatory
    - „Semantically vague“, Occur with limited class of predicates
  - FMs:
    - **May** occur more than once, **Prototypically** optional
    - „Semantically homogeneous“, Occur with all predicates
- Third class of functors: „quasi-valency complements“
  - May not occur more than once, but are semantically homogeneous
  - Example: Intent

Praguian roles and alternations

- Do alternations obtain the same analysis?
  - Only lexically unspecific alternations:
    - [Pojist'ovna.ACT] zaplatila [vyrobcum.ADDR] [ztraty.PAT]
      "[The insurance company] covered [producers'] [losses]"
    - [Vyrobci.ADDR] dostali [od pojist'ovny.ACT] [zaplaceny ztraty.PAT]
      "[The producers] got covered [from the insurance company] [the losses]."
  - Not lexically specific alternations:
    - Martin.ACT nastrikal barvu.PAT na zed'.DIR3
      "Martin sprayed paint on the wall."
    - Martin.ACT nastrikal zed'.PAT barvou.MEANS
      "Martin sprayed the wall with paint."
  - However: This information present in VALLEX (valency lexcion for Czech)

Praguian roles and semantic properties

- How strongly do Prague roles model semantic properties?
  - Dichotomy between IPs and FMs
    - IPs provide only very weak, general properties
      - "Shifting" allows stronger verb-specific interpretation: but largely theoretic account
    - FMs semantically defined
      - However, event-unspecific information

Computational Modelling

- Main task: automatic assignment of tectogrammatical functors
  - Input: analytical (surface dependency) structure
  - Output: tectogrammatical structure
- Modelling in two steps:
  - Structural changes: delete non-content words
  - Classification: Assign functor to each node
- Results: Simple ML approaches can yield F-Scores around 80-85% (Zabokrtsky 2002)

Praguian roles: Summary

- Status of functors differs from classical roles
  - Functor assignment verb sense-specific
    - Alternations explicable by reference to mappings in valency lexicon
  - Syntax-driven assignment of Inner Participants
    - Stronger semantic characterisation only through shifting
- Tectogrammatical description entrenched in FGD
- Czech not widely investigated language

Merit of PDT widely recognised, but limited impact

PropBank

- Initiative to add exhaustive role-semantic layer to Penn TreeBank (Wall Street Journal)
  - "Proposition Bank"
- About 1 M words
- ~4000 predicates (verbs only)
  - NomBank: ongoing project to annotate nouns as well (over 90% of nouns in corpus completed)
- "Practical", surface-oriented annotation framework

Annotation process

- Two step process:
  1. "Framing": Development of "frame files" by a linguist
     - Bottom-up approach
     - Contain **sense distinctions** for predicates
     - Contain definition of "role set" for each sense
     - Available online: http://www.cs.rochester.edu/~gildea/PropBank/Sort/
  2. Annotation
     - Each verb annotated separately
     - "Flat trees"

Verb senses

- Verb senses are separated generally if they take different numbers of arguments
  - decline.01 “go down incrementally”
    - Arg1: entity going down
    - Arg2: amount gone down
    - Arg3: start point
    - Arg4: end point
  - decline.02: “reject”
    - Arg0: agent
    - Arg1: rejected thing
- Results in coarse-grained sense distinctions (average 1.4 senses / verb)

Role sets: Arguments

- Arguments vs. Adjuncts:
  - Arguments
    - Verb sense-specific
    - Can occur at most once
    - Identified by index number plus *verb sense-specific* “mnemonic”
    - Criteria for index numbers:
      - Arg0: “proto-agent” (Dowty)
      - Arg1: “proto-patient”
      - Rest: none (though consistent within Levin Class)

decline.02: “reject”
Arg0: agent
Arg1: rejected thing

Role sets: Adjuncts

- Arguments vs. Adjuncts:
  - Adjuncts/Modifiers
    - Universal
    - Can occur any number of times
    - ARGM-X: 11 subtypes
      - ARGM-LOC: Location
      - ARGM-EXT: Extent
      - ARGM-NEG: Negation (?)

Example

[Its net income $_{ARG1}$] *declined*
[42% $_{ARG2}$] to [$121 million $_{ARG4}$]
[in the first 9 months of 1989 $_{ARGM-TMP}$]

PropBank roles and alternations

- PropBank roles generalise over alternations
  - Roles defined on “canonical realisation”

Standard: [Peter $_0$] gave [Mary $_2$] [the book $_1$]
Alternation: [Peter $_0$] gave [the book $_1$] [to Mary $_2$]

- Roles might or might not transfer well across predicates

[Peter $_0$] sold [the book $_1$] [to John $_2$]
[John $_0$] bought [the book $_1$] [from Peter $_2$]

PropBank roles and semantic properties

- Roles have a twofold nature
  - Identified by *universal* index number plus *verb sense-specific* “mnemonic”
- Universal meaning aspect:
  - For ARG-0 and ARG-1 (Dowty’s proto-roles)
    - Provides prototypical properties for ARG-0 and ARG-1
  - Nothing for higher ARGs
- Verb sense-specific meaning aspect:
  - Provides fine-grained specification of role
    - However, “no theoretical standing” (Palmer et al. 2005)

Computational Modelling

- Main task: Assign role labels
  - Input: Syntactic structure
  - Output: list of role labels / NONE
- CoNLL shared tasks 2004/2005
  - Best systems around 80% F-Score (automatically generated input)
  - With “gold standard” input up to 90%
- Properties of the task:
  - Most important: syntactic path, predicate, parts of speech
  - Linking between syntax (grammatical functions) and PropBank roles rather straightforward

Cross-lingual activities

- Proposition Bank for Chinese
- Similar methodology to PropBank
  - On top of Penn Chinese Treebank
- Similar methodology:
  - Coarse-grained verb senses
  - Twofold role definitions
- Is the data comparable across languages?
  - ARG0/1 yes, syntactically motivated roles: open

“Practical annotation”

- PropBank places emphasis on simple, consistent annotation
- Annotation of “what is there”
  - No annotation of unrealised arguments
  - No annotation of non-literal phenomena
    - “[That] goes [too far]”: simply go.06 “proceed”
  - No role generalisations across senses
- Rationale: These phenomena cannot be annotated reliably; can be induced from the data in subsequent steps

PB: Relevance

- Advantages:
  - English
  - Additional layer on standard dataset
    - Gold standard syntax (Treebank)
    - Interaction between syntax and semantics
- Disadvantages:
  - Unrepresentative corpus
    - Syntactic structure: newspaper style
    - Domain vocabulary.
      Most frequent ARG-1 of “rise”: “stocks”

Frame Semantics

- “Semantics of understanding” (Fillmore 1985)
  - Goal: characterise the “relation between linguistic texts and the process and products of their interpretation”
- Observation: Foreign language learning proceeds *scenario*-driven
  - “Monday” or “fortnight” only comprehensible through background knowledge about time organisation

Frame Semantics

- Central concept: Frame
  - A conceptual structure which provides the background and motivation for the existence of words in the language and for their use in discourse“
  - (Rough) similarity to schemata/frames in KI and gestalt in cognitive psychology
- Claim: Meaning of predicate can be modelled by reference to its frame
  - More specifically, **frame** = **prototypical situation**
  - Request, Statement

Frame Semantics

- Claim 2: The arguments of a predicate can be described by reference to the **relevant participants and objects** in that situation
  - „Frame elements“ = semantic roles
  - Frame **Request**: Speaker, Message, Medium
- Model of predicate-argument structure on **cognitive basis**
  - Consequence: **Semantic roles are frame-specific**

FrameNet

- Project in Berkeley since (1997), head: C. Fillmore
  - Goal: Construction of a frame-semantic lexicon for the English “core vocabulary”
    - For each predicate, list all appropriate frames
    - For each frame, list the frame elements
    - Provide annotated example sentences
- Current coverage: http://framenet.icsi.berkeley.edu
  - ~700 frames
  - ~7500 lemmas (V, N, Adj, some Preps and MWEs)
  - ~9000 senses (polysemy ~ 1.2)
  - ~130 000 example sentences

FrameNet: The construction

- Problem: no “a priori” inventory of frames
- Lexicographic “bootstrapping” approach:
  - Frame definition **interleaved** with predicate description
- Procedure:
  - Find **predicate groups / clusters** with
    1. common meaning (same semantic properties) and
    2. common linguistic expressiveness (same set of realisable “core” roles)
  - Define frame (potentially in contrast to existing frames)
- Tension between cognitive and linguistic criteria
  - Tries to strike a compromise between top-down and bottom-up
  - Cf. part 5

Frame Definition: Example

Frame: COMMITMENT		
Def.	A Speaker makes a commitment to an Addressee to carry out some future action. This may be an action desirable (promise) or undesirable (threaten) to the Addressee.	
Frame El.s	SPEAKER	The Speaker is the person who commits him/herself to do something
	ADDRESSEE	The Speaker's commitment can be made to an Addressee
	MESSAGE	An expression of the commitment made by the Speaker.
	TOPIC	The topic about which the Speaker makes a promise.
	MEDIUM	Medium is the physical entity or channel used to transmit the Message.
FEEs	consent.v, covenant.n, covenant.v, oath.n, vow.n, pledge.n, pledge.v, promise.n, promise.v, swear.v, threat.n, threaten.v, undertake.v	
Ann.	[Democratic audiences]$_{Speaker}$ had to **consent** [to this approach]$_{Message}$. [The politicians]$_{Speaker}$ made vague **promises** [about independence]$_{Topic}$. ['I'll be back , "]$_{Message}$ [he]$_{Speaker}$ **threatened**.	

37

Frame-to-frame relations

- There is an incomplete hierarchy that links frames (and their roles)
- Inheritance: "Specialisation" (all roles inherited)
  - **Placing** inherits from **Transitive_action**
- Uses: Cognitive background
  - **Placing** uses **Motion**
- Subframe: relates events to subevents
  - **Placing** is subframe of **Cause_motion**
- Is causative/inchoative of: Relates alternations
  - **Change_position_on_scale** is inchoative of **Cause_change_of_scalar_position**

38

Annotation style

- No prior syntactic analysis
- One frame at a time
  - "Flat trees"
- Example:
  - [The occupants $_{Agent}$] jumped out and began to LOAD [packages $_{Theme}$] [into a waiting truck $_{Goal}$].

39

Role Types

- Core roles ("Arguments"): Speaker, Message, Addressee
  - Can only occur once
  - Have to be realisable by each predicate (or be incorporated)
  - Are frame-specific
- Peripheral roles ("Adjuncts"): Time, Location
  - Can occur with any frame
  - Can occur more than once
- Extrathematic roles: Beneficiry, Degree
  - Can occur with many frames
  - Can occur only once
- Note: Some roles are "core" in some frames, but non-core in others
  - Example: Location is core in Motion frames

Criteria for Role Definitions

- Most roles defined by their semantic properties
  - **Statement.Speaker**: "The speaker is the person making the statement"
- Sometimes, ontological considerations ("semantic type")
  - Actor vs. Cause
- Sometimes, syntactic considerations
  - To account for "reciprocal alternations"
    - [Car1 $_{Impactor}$] collided [with Car2 $_{Impactee}$].
    - [Car1 and Car2 $_{Impactors}$] collided.
  - "Excludes"/"Implies" role-to-role relations.
- Same role name across frames indicates similarity, but only in a loose sense

Frame-semantic roles for alternations

- For semantically defined roles: Same Analysis
  - [Peter $_{Seller}$] sold [the book $_{Goods}$] [to John $_{Buyer}$]
  - [The book $_{Goods}$] was sold [to John $_{Buyer}$] [by Peter $_{Seller}$]
- For syntactically defined roles: Role-to-role relations.
- Some alternations evoke different frames: requires frame-to-frame relations
  - [The temperature $_{Item}$] increases. (Inchoative)
  - [The sun $_{Cause}$] increases [the temperature $_{Item}$]. (Causative)

Frame-semantic roles and semantic properties

- Mid-grained level of semantic characterisation
  - Definition of roles at frame level
- (Naturally) not as detailed as verb-specific definitions
  - Judgment: ADDRESSEE is judged either positively or negatively
- Problems:
  - Incomplete frame hierarchy
  - Whole area of nonliteral usages (cf. part 3)

Computational Modelling

- Two-step procedure
  1. Assign frame to predicate (similar to sense disambiguation)
  2. Assign role labels to syntactic nodes (similar to PropBank)
- Modelling mostly concentrated on step 2
  - Gildea and Jurafsky 2000/2002, SENSEVAL 3 Track
  - Results:
    - Best F-Scores 70 -- 75 (automatically generated input)
    - Somewhat more difficult than PropBank
- Problem with step 1: Incompleteness of FrameNet
  - Naïve modelling as classification presupposes complete sense inventory for each predicate

Cross-lingual activities

- FrameNet initiatives for
  - German (SALSA, Saarbruecken)
  - Japanese (JFN, Keio University)
  - Spanish (SFN, Barcelona)
- Conceptual nature of frames / frame elements allows re-use of most frames
  - Differences in lexicalisation patterns (cf. part 5)

Summary

Differences and Commonalities: Definitions

- Frameworks differ in the emphasis on prior (theoretical) assumptions
  - Prague (linguistics) > FrameNet (cognition) > PropBank
- All frameworks distinguish “central” from “not-so-central” roles
  - Difference: two vs. three categories
- “Not-so-central” roles can be defined on semantic grounds
  - But they are not so central
- Central roles: different approaches
  - Continuum in the use of syntactic and semantic criteria
    - Syntax < Prague < PropBank < FrameNet < Semantics
  - Even FrameNet cannot completely get rid of syntactically motivated distinctions

Differences and Commonalities: Phenomena

- Alternations:
  - More semantically oriented role definitions lead to stronger generalisations
- Semantic properties:
  - PDT and PropBank offer general (vague) and verb-specific (unformalised) roles
  - FrameNet attempts to provide “middle ground” by defining roles per situation
    - Still middle ground

Challenges (I): Modelling

- Performance for role assignment (as classification task) comparable for all frameworks (75-85% F-Score)
  - Caveat: current strategy is evaluation on held-out datasets from same corpus
- Challenge: provide accurate analysis for free text
  - Must address **incompleteness** on many levels: Unseen words, unseen senses, unseen constructions, etc.

Challenges (II): Application

- Most important for NLP is characterisation of semantic properties
  - Answer questions like “does X imply Y”?
  - Information access etc.
- At the same time, most difficult problem
  - “AI-complete”
  - All frameworks fall short (specific characterisations are not formalised - shifting, “mnemonics”, natural language, …)
- Challenge: Demonstrate that semantic roles can provide a clear benefit for NLP

Promising direction: templates

- Template: representation for information extraction

Presentation:

Presenter:
Date:
Time:
Place:
Title:

  - Typically filled by pattern matching
  - Very domain-specific
- Semantic roles as domain-independent generalisation of templates?

References: Prague

- Functional Generative Description:
  P. Sgall, E. Hajicova and J. Panevova: The Meaning of the sentence in its semantic and pragmatic aspects. Dordrecht: Reidel (1986).
- Tectogrammatic roles:
  E. Hajicova: Dependency-based underlying-structure tagging of a very large Czech corpus. T.A.L. vol. 41 no.1 (2000).
- Inner participants vs. free modifiers:
  M. Lopatkova and J. Panevova: Recent developments in the theory of valency in the light of the Prague Dependency Treebank. Insight into Slovak and Czech Corpus Linguistic. Veda Bratislava (2005).
- Automatic functor assignment:
  Z. Zabokrtsky and P. Sgall and S. Dzeroski: A Machine Learning Approach to Automatic Functor Assignment in the Prague Dependency Treebank. Proceedings of LREC 2002.
- Website: http://ufal.mff.cuni.cz/pdt2.0/

References: PropBank

- The corpus:
  M. Palmer and D. Gildea and P. Kingsbury: The Proposition Bank: An Annotated Corpus of Semantic Roles. Computational Linguistics 31(1), 2005.
- Modelling:
  D. Gildea and M. Palmer: The Necessity of Syntactic Parsing for Predicate Argument Recognition. Proceedings of ACL 2002.
- Modelling:
  X. Carreras and L. Marquez: Proceedings of the CoNLL shared task 2004/2005: Semantic Role Labelling.
- Website:
  http://www.cs.rochester.edu/~gildea/PropBank/Sort/

References: Frame Semantics

- Frame Semantics:
  C. Fillmore: Frames and the semantics of understanding. Quaderni di Semantics VI(2), 1985.
- FrameNet:
  C. Baker, C. Fillmore, and J. Lowe: The Berkeley FrameNet project. In Proceedings of COLING/ACL 1998.
- "The Book": FrameNet: Theory and Practice. In-depth discussion of frame construction and annotation. Can be found on the website.
- Modelling:
  D. Gildea and D. Jurafsky: Automatic labelling of semantic roles. Computational Linguistics 28(3), 2002.
- Website: http://framenet.icsi.berkeley.edu/