
Case Grammar Semantic Role Labeling

Nancy McCracken

Semantics of events in sentences

- In a sentence, a **verb and its semantic roles** form a **proposition**; the verb can be called the predicate and the roles are known as arguments.

*When Disney **offered** to **pay** Mr. Steinberg a premium for his shares, the New York investor didn't **demand** the company also **pay** a premium to other shareholders.*

Example semantic roles for the verb “pay” (using verb-specific roles)

When [_{payer} Disney] offered to [_v **pay**] [_{recipient} Mr. Steinberg] [_{money} a premium] for [_{commodity} his shares], the New York investor ...

CASE Grammar

- **Fillmore, Charles (1968) “*The Case for Case.*”**
 - A response to Chomsky’s disregard for any semantics
 - “A semantically justified syntactic theory”
- Given a sentence, it is possible to say much more than this NP is the subject and this NP is the object
- Chomsky’s Transformational Grammar would reduce active & passive versions of the same deep structure, but doesn’t go far enough to reveal why this is possible semantically
 - *A crowbar could open that door easily.*
 - *That door could be opened easily with a crowbar.*

CASE Grammar

- Focuses on conceptual events
 - for each event or situation, there is a limited number of roles/cases which people or objects play in the situation
 - roles reflect ordinary human judgments about:
 - Who did the action?
 - Who / what was it done to?
 - What was it done with?
 - Where was it done?
 - What was the result?
 - When was it done?

Syntactic structure vs. semantic structure

- Syntactic similarities hide semantic dissimilarities
 - *We baked every Saturday morning.*
 - *The pie baked to a golden brown.*
 - *This oven bakes evenly.*
 - 3 subject NPs perform very different roles in regard to *bake*
- Syntactic dissimilarities hide semantic similarities
 - John_{agent} broke the window_{theme}.
 - John_{agent} broke the window_{theme} with a rock_{instrument}.
 - The rock_{instrument} broke the window_{theme}.
 - The window_{theme} broke.
 - The window_{theme} was broken by John_{agent}.

Cases (aka Thematic Roles or Theta Roles)

- Some of Fillmore's original set of roles still in use as general descriptors of roles
 - **Agentive (A)**
 - the instigator of the action, an animate being
 - *John opened the door.*
 - *The door was opened by John.*
 - **Instrumental (I)**
 - the thing used to perform the action, an inanimate object
 - *The key opened the door.*
 - *John opened the door with the key.*
 - **Locative (L)**
 - the location or spatial orientation of the state or action of the verb
 - *It's windy in Chicago.*
- Other original roles not typically used
 - **Dative (D), Neutral (N), Objective (O), Factitive (F)**

Verb-specific Roles

- Difficult to fit many verbs and roles into the general thematic roles
 - Many general sets are proposed; not uniform agreement
 - Generalized semantic roles now often called
 - Proto roles: Proto-agent, proto-patient, etc.
 - Or theta roles
- Verb-specific roles are proposed in systems
 - PropBank annotates the verbs of Penn Treebank
 - Extended with NomBank for nominalizations
 - FrameNet annotates the British National Corpus
 - Uses domains of semantically similar verbs called frames.

Propbank

- Propbank is a corpus with annotation of semantic roles, capturing the **semantic role structure of each verb sense**
 - Funded by ACE to Martha Palmer and Mitch Marcus at U Penn
- Each verb sense has a **frameset**, listing its possible semantic roles
 - Argument notation uses numbers for the annotation
 - First sense of accept (accept.01)
 - Arg0: acceptor
 - Arg1: thing accepted
 - Arg2: accepted-from
 - Arg3: attribute
- The frameset roles are standard across all syntactic realizations in the corpus of that verb sense
 - Each verb has a frameset file describing the args as above
 - Example texts are also given

Roles consistent with VerbNet

- Propbank builds on VerbNet to assign more specific roles.
- VerbNet is one extension of Levin's verb classes, giving semantic roles from about 20 possible roles
 - Agent, Patient, Theme, Experiencer, etc.
 - Similar to the theta roles
- Each class consists of a number of synonymous verbs that have the same semantic and syntactic role structure in a frame
- Whenever possible, the Propbank argument numbering is made consistent for all verbs in a VerbNet class.
 - There is only 50% overlap between Propbank and VerbNet verbs.
- Example from frameset file for “explore”, which has a VN class:

```
<roleset id="explore.01" name="explore, discover new places or things" vncls="35.4">  
<roles> <role descr="explorer" n="0">  
    <vnrole vncls="35.4" vntheta="Agent"/></role>  
    <role descr="thing (place, stuff) explored" n="1">  
        <vnrole vncls="35.4" vntheta="Location"/></role>  
</roles>
```

Semantic Role Notation for Propbank

- The first two numbered arguments correspond, approximately, to the **core case roles**:
 - Arg0 – Prototypical Agent
 - Arg1 – Prototypical Patient or Theme
 - Remaining numbered args are verb specific case roles, Arg2 through Arg5
- Another large groups of roles are the **adjunctive roles** (which can be applied to any verb) and are annotated as ArgM with a suffix:

– ArgM-LOC – location	ArgM-CAU - cause
– ArgM-EXT – extent	ArgM-TMP - time
– ArgM-DIR – direction	ArgM-PNC – purpose
– ArgM-ADV – general purpose adverbial	ArgM-MNR - manner
– ArgM-DIS – discourse connective	ArgM- NEG – negation
– ArgM-MOD – modal verb	

Adjunctive and additional arguments

- Example of adjunctive arguments
 - Not all core arguments are required to be present
 - See Arg2 in this example.
 - Arguments can be phrases, clauses, even partial words.

*When Disney **offered** to **pay** Mr. Steinberg a premium for his shares, the New York investor didn't **demand** the company also **pay** a premium to other shareholders.*

Example of Propbank annotation (on demand):

[_{ArgM-TMP} When Disney offered to pay Mr. Steinberg a premium for his shares], [_{Arg0} the New York investor] did [_{ArgM-NEG} n' t] [_v **demand**] [_{Arg1} the company also pay a premium to other shareholders].

Where for **demand**, Arg0 is “asker”, Arg1 is “favor”, Arg2 is “hearer”

Prepositional phrases and additional args

- Arguments that occur as the head of a prepositional phrase are annotated as the whole phrase
 - Consistent with other ArgM's that are prepositional phrases

[_{Arg1} Its net income] [_v declining] [_{ArgM-EXT} 42%] [_{Arg4} to \$121 million] [_{ArgM-TMP} in the first 9 months of 1989]

- Additional arguments are
 - ArgA – causative agents
 - C-Arg* - a continuation of another arg (mostly for what is said)
 - R-Arg* - reference to another arg (mostly for “that”)

Propbank Annotations

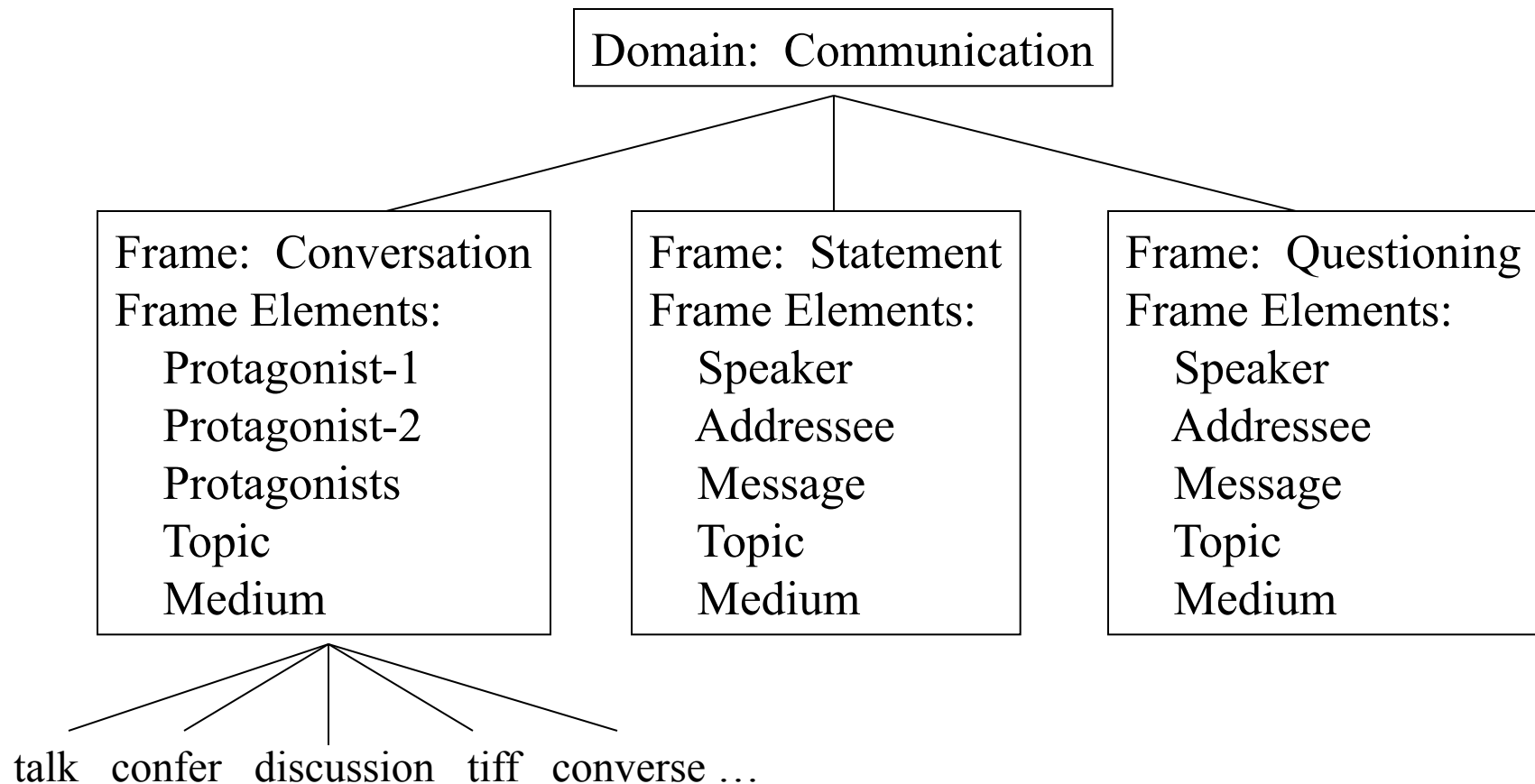
- **Framesets** were created by looking at sample sentences containing each verb sense.
 - ~ 4500 frames (in 3314 framesets for each verb)
- Corpus is primarily newswire text from Penn Treebank
 - Annotated the Wall Street Journal section, and, more recently, the “Brown” corpus
 - Verbs and semantic role annotations added to the parse trees
- Annotators are presented with **roleset descriptions** of a verb and the (gold) **syntactic parses** of a sentence in Treebank, and they annotate the roles of the verb.
 - Lexical sampling – annotated on a verb-by-verb basis.
 - ~40,000 sentences were annotated
- Interannotator agreement
 - Identifying argument and classifying role: 99%
 - kappa statistic of .91 overall and .93 if ArgM’s excluded

FrameNet

- Project at International Computer Science Institute with Charles Fillmore
 - <http://framenet.icsi.berkeley.edu/>
- Similar goal to document the syntactic realization of arguments of predicates in the English language
- Starts from semantic frames (e.g. Commerce) and defines frame elements (e.g. Buyer, Goods, Seller, Money)
- Annotates example sentences chosen to illustrate all possibilities
 - But recent release includes 132,968 sentences
 - British National Corpus

Example of FrameNet frames

- Semantic frames are related by topic domain



Comparison of FrameNet and Propbank

- FrameNet semantic roles are consistent for semantically related verbs (not just synonyms as in the VerbNet subset of PropBank)
- Commerce examples:

FrameNet annotation:

[_{Buyer} Chuck] *bought* [_{Goods} a car] [_{Seller} from Jerry][_{Payment} for \$1000].
[_{Seller} Jerry] *sold* [_{Goods} a car] [_{Buyer} to Chuck] [_{Payment} for \$1000].

Propbank annotation:

[_{Arg0} Chuck] *bought* [_{Arg1} a car] [_{Arg2} from Jerry][_{Arg3} for \$1000].
[_{Arg0} Jerry] *sold* [_{Arg1} a car] [_{Arg2} to Chuck] [_{Arg3} for \$1000].

Frame for buy:

Arg0: buyer
Arg1: thing bought
Arg2: seller
Arg3: price paid
Arg4: benefactive

Frame for sell:

Arg0: seller
Arg1: thing sold
Arg2: buyer
Arg3: price paid
Arg4: benefactive