Semantic Role Labeling

Introduction

Many slides adapted from Dan Jurafsky
Semantic Role Labeling

Who  did  what  to  whom  at  where?

The police officer detained the suspect at the scene of the crime

Agent  Predicate  Theme  Location
Can we figure out that these have the same meaning?

XYZ corporation *bought* the stock.
They *sold* the stock to XYZ corporation.
The stock was *bought* by XYZ corporation.
The *purchase* of the stock by XYZ corporation...
The stock *purchase* by XYZ corporation...
A Shallow Semantic Representation: Semantic Roles

Predicates (bought, sold, purchase) represent an event. Semantic roles express the abstract role that arguments of a predicate can take in the event.
Semantic Role Labeling

Semantic Roles
Getting to semantic roles

Neo-Davidsonian event representation:

Sasha broke the window
Pat opened the door

$\exists e, x, y \, Breaking(e) \land Breaker(e, Sasha) \land \neg Breaking(e, Pat)$
$\exists e, x, y \, Opening(e) \land Opener(e, Pat) \land \neg Opening(e, Sasha)$

Subjects of break and open: **Breaker** and **Opener**

**Deep roles** specific to each event (breaking, opening)

Hard to reason about them for NLU applications like QA
Thematic roles

- **Breaker** and **Opener** have something in common!
  - Volitional actors
  - Often animate
  - Direct causal responsibility for their events
- Thematic roles are a way to capture this semantic commonality between *Breakers* and *Eaters*.
- They are both **AGENTS**.
- The *BrokenThing* and *OpenedThing*, are **THEMES**.
  - prototypically inanimate objects affected in some way by the action
Thematic roles

• One of the oldest linguistic models
  • Indian grammarian Panini between the 7th and 4th centuries BCE

• Modern formulation from Fillmore (1966, 1968), Gruber (1965)
  • Fillmore influenced by Lucien Tesnière’s (1959) *Éléments de Syntaxe Structurale*, the book that introduced dependency grammar
  • Fillmore first referred to roles as *actants* (Fillmore, 1966) but switched to the term *case*
Thematic roles

- A typical set:

<table>
<thead>
<tr>
<th>Thematic Role</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGENT</td>
<td>The volitional causer of an event</td>
<td>The waiter spilled the soup.</td>
</tr>
<tr>
<td>EXPERIENCER</td>
<td>The experiencer of an event</td>
<td>John has a headache.</td>
</tr>
<tr>
<td>FORCE</td>
<td>The non-volitional causer of the event</td>
<td>The wind blows debris from the mall into our yards.</td>
</tr>
<tr>
<td>THEME</td>
<td>The participant most directly affected by an event</td>
<td>Only after Benjamin Franklin broke the ice...</td>
</tr>
<tr>
<td>RESULT</td>
<td>The end product of an event</td>
<td>The city built a regulation-size baseball diamond...</td>
</tr>
<tr>
<td>CONTENT</td>
<td>The proposition or content of a propositional event</td>
<td>Mona asked “You met Mary Ann at a supermarket?”</td>
</tr>
<tr>
<td>INSTRUMENT</td>
<td>An instrument used in an event</td>
<td>He poached catfish, stunning them with a shocking device...</td>
</tr>
<tr>
<td>BENEFICIARY</td>
<td>The beneficiary of an event</td>
<td>Whenever Ann Callahan makes hotel reservations for her boss...</td>
</tr>
<tr>
<td>SOURCE</td>
<td>The origin of the object of a transfer event</td>
<td>I flew in from Boston.</td>
</tr>
<tr>
<td>GOAL</td>
<td>The destination of an object of a transfer event</td>
<td>I drove to Portland.</td>
</tr>
</tbody>
</table>
Thematic grid, case frame, θ-grid

Example usages of “break”

*John broke the window.*
AGENT THEME

*John broke the window with a rock.*
AGENT THEME INSTRUMENT

*The rock broke the window.*
INSTRUMENT THEME

*The window broke.*
THEME

*The window was broken by John.*
THEME AGENT

**thematic grid, case frame, θ-grid**

**Break:**

AGENT, THEME, INSTRUMENT.

**Some realizations:**

AGENT/Subject, THEME/Object
AGENT/Subject, THEME/Object, INSTRUMENT/PP with
INSTRUMENT/Subject, THEME/Object
THEME/Subject
Diathesis alternations (or verb alternation)

**Doris** gave the book to Cary.  
AGENT  THEME  BENEFICIARY

**Doris** gave Cary the book.  
AGENT  BENEFICIARY  THEME

**Dative alternation**: particular semantic classes of verbs, “verbs of future having” (advance, allocate, offer, owe), “send verbs” (forward, hand, mail), “verbs of throwing” (kick, pass, throw), etc.

**Levin (1993)**: 47 semantic classes (“**Levin classes**”) for 3100 English verbs and alternations. In online resource VerbNet.
Problems with Thematic Roles

Hard to create standard set of roles or formally define them
Often roles need to be fragmented to be defined.

Levin and Rappaport Hovav (2015): two kinds of INSTRUMENTS

intermediary instruments that can appear as subjects
  The cook opened the jar with the new gadget.
  The new gadget opened the jar.

enabling instruments that cannot
  Shelly ate the sliced banana with a fork.
  *The fork ate the sliced banana.
Alternatives to thematic roles

1. **Fewer roles**: generalized semantic roles, defined as prototypes (Dowty 1991)
   - PROTO-AGENT
   - PROTO-PATIENT
   - PropBank

2. **More roles**: Define roles specific to a group of predicates
   - FrameNet
Semantic Role Labeling

The Proposition Bank (PropBank)
PropBank

PropBank Roles

Following Dowty 1991

Proto-Agent

• Volitional involvement in event or state
• Sentience (and/or perception)
• Causes an event or change of state in another participant
• Movement (relative to position of another participant)

Proto-Patient

• Undergoes change of state
• Causally affected by another participant
• Stationary relative to movement of another participant
PropBank Roles

• Following Dowty 1991
  • Role definitions determined verb by verb, with respect to the other roles
  • Semantic roles in PropBank are thus verb-sense specific.

• Each verb sense has numbered argument: Arg0, Arg1, Arg2,...
  Arg0: PROTO-AGENT
  Arg1: PROTO-PATIENT
  Arg2: usually: benefactive, instrument, attribute, or end state
  Arg3: usually: start point, benefactive, instrument, or attribute
  Arg4 the end point

(Arg2-Arg5 are not really that consistent, causes a problem for labeling)
agree.01
Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [Arg0 The group] agreed [Arg1 it wouldn’t make an offer].
Ex2: [ArgM-TMP Usually] [Arg0 John] agrees [Arg2 with Mary] [Arg1 on everything].

fall.01
Arg1: Logical subject, patient, thing falling
Arg2: Extent, amount fallen
Arg3: start point
Arg4: end point, end state of arg1
Ex1: [Arg1 Sales] fell [Arg4 to $25 million] [Arg3 from $27 million].
Ex2: [Arg1 The average junk bond] fell [Arg2 by 4.2%].
Advantage of a ProbBank Labeling

increase.01 “go up incrementally”
Arg0: causer of increase
Arg1: thing increasing
Arg2: amount increased by, EXT, or MNR
Arg3: start point
Arg4: end point

This would allow us to see the commonalities in these 3 sentences:

\[ \text{[Arg0 Big Fruit Co.]} \text{ increased [Arg1 the price of bananas].} \]
\[ \text{[Arg1 The price of bananas] was increased again [Arg0 by Big Fruit Co.]} \]
\[ \text{[Arg1 The price of bananas] increased [Arg2 5%].} \]
Modifiers or adjuncts of the predicate: Arg-M

<table>
<thead>
<tr>
<th>ArgM-TMP</th>
<th>LOC</th>
<th>DIR</th>
<th>MNR</th>
<th>PRP/CAU</th>
<th>REC</th>
<th>ADV</th>
<th>PRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>yesterday evening, now</td>
<td>at the museum, in San Francisco</td>
<td>down, to Bangkok</td>
<td>clearly, with much enthusiasm</td>
<td>because ... , in response to the ruling</td>
<td>themselves, each other</td>
<td>...ate the meat raw</td>
<td></td>
</tr>
</tbody>
</table>
Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.
Analysts have been expecting a GM-Jaguar pact (that would give the US car maker an eventual 30% stake in the British company).
Annotated PropBank Data

- Penn English TreeBank, OntoNotes 5.0.
  - Total ~2 million words
- Penn Chinese TreeBank
- Hindi/Urdu PropBank
- Arabic PropBank

2013 Verb Frames Coverage
Count of word sense (lexical units)

<table>
<thead>
<tr>
<th>Language</th>
<th>Final Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>10,615*</td>
</tr>
<tr>
<td>Chinese</td>
<td>24,642</td>
</tr>
<tr>
<td>Arabic</td>
<td>7,015</td>
</tr>
</tbody>
</table>

From Martha Palmer 2013 Tutorial
Plus nouns and light verbs

Example Noun: Decision

- Roleset: Arg0: decider, Arg1: decision…

- “…[your \text{ARG0}] [decision \text{REL}]
  [to say look I don't want to go through this anymore \text{ARG1}]”

Example within an LVC: Make a decision

- “…[the President \text{ARG0}] [made \text{REL-LVB}]
  the [fundamentally correct \text{ARGM-ADJ}]
  [decision \text{REL}] [to get on offense \text{ARG1}]”

Slight from Palmer 2013
Semantic Role Labeling

Semantic Role Labeling Algorithm
Semantic role labeling (SRL)

• The task of finding the semantic roles of each argument of each predicate in a sentence.

• FrameNet versus PropBank:

   [You] can’t [blame] [the program] [for being unable to identify it]
   COGNIZER TARGET EVALUTEE REASON

   [The San Francisco Examiner] issued [a special edition] [yesterday]
   ARG0 TARGET ARG1 ARGM-TMP
History

• Semantic roles as a intermediate semantics, used early in
  • machine translation (Wilks, 1973)
  • question-answering (Hendrix et al., 1973)
  • spoken-language understanding (Nash-Webber, 1975)
  • dialogue systems (Bobrow et al., 1977)

• Early SRL systems
  Simmons 1973, Marcus 1980:
  • parser followed by hand-written rules for each verb
  • dictionaries with verb-specific case frames (Levin 1977)
Why Semantic Role Labeling

- A useful shallow semantic representation
- Improves NLP tasks like:
  - question answering
    - Shen and Lapata 2007, Surdeanu et al. 2011
  - machine translation
    - Liu and Gildea 2010, Lo et al. 2013
Recall that the difference between these two models of semantic roles is that FrameNet (22.27) employs many frame-specific frame elements as roles, while PropBank (22.28) uses a smaller number of numbered argument labels that can be interpreted as verb-specific labels, along with the more general ARGM labels. Some examples:

(22.27) [You] can't [blame] [the program] [for being unable to identify it]

COGNIZER TARGET EV ALUEE REASON

(22.28) [The San Francisco Examiner] issued [a special edition] [yesterday]

ARG
ARG 0
ARG 1
ARGM - TMP

A simplified semantic role labeling algorithm is sketched in Fig. 22.4. While there are a large number of algorithms, many of them use some version of the steps in this algorithm. Most algorithms, beginning with the very earliest semantic role analyzers (Simmons, 1973), begin by parsing, using broad-coverage parsers to assign a parse to the input string. Figure 22.5 shows a parse of (22.28) above. The parse is then traversed to find all words that are predicates. For each of these predicates, the algorithm examines each node in the parse tree and decides the semantic role (if any) it plays for this predicate. This is generally done by supervised classification. Given a labeled training set such as PropBank or FrameNet, a feature vector is extracted for each node, using feature templates described in the next subsection. A 1-of-N classifier is then trained to predict a semantic role for each constituent given these features, where N is the number of potential semantic roles plus an extra NONE role for non-role constituents. Most standard classification algorithms have been used (logistic regression, SVM, etc). Finally, for each test sentence to be labeled, the classifier is run on each relevant constituent. We give more details of the algorithm after we discuss features.

```
function SEMANTICROLELABEL(words) returns labeled tree

parse ← PARSE(words)
for each predicate in parse do
  for each node in parse do
    featurevector ← EXTRACTFEATURES(node, predicate, parse)
    CLASSIFYNODE(node, featurevector, parse)
```
How do we decide what is a predicate

• If we’re just doing PropBank verbs
  • Choose all verbs
  • Possibly removing light verbs (from a list)

• If we’re doing FrameNet (verbs, nouns, adjectives)
  • Choose every word that was labeled as a target in training data
Semantic Role Labeling

Figure 22.5 Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the path feature NP $S \# VP \# VBD$ for ARG0, the NP-SBJ constituent.

- The headword of the constituent, **Examiner**. The headword of a constituent can be computed with standard head rules, such as those given in Chapter 11 in Fig. 22.5. Certain headwords (e.g., pronouns) place strong constraints on the possible semantic roles they are likely to fill.

- The headword part of speech, **NNP**.

- The path in the parse tree from the constituent to the predicate. This path is marked by the dotted line in Fig. 22.5. Following Gildea and Jurafsky (2000), we can use a simple linear representation of the path, NP $S \# VP \# VBD$.

- The voice of the clause in which the constituent appears, in this case, active (as contrasted with passive). Passive sentences tend to have strongly different linkings of semantic roles to surface form than do active ones.

- The binary linear position of the constituent with respect to the predicate, either before or after.

- The subcategorization of the predicate, the set of expected arguments that appear in the verb phrase. We can extract this information by using the phrase-structure rule that expands the immediate parent of the predicate; VP $! VBD$ NP PP for the predicate in Fig. 22.5.

- The named entity type of the constituent.

- The first words and the last word of the constituent.

The following feature vector thus represents the first NP in our example (recall that most observations will have the value NONE rather than, for example, ARG0, since most constituents in the parse tree will not bear a semantic role):

ARG0: [issued, NP, Examiner, NNP, NP $S \# VP \# VBD$, active, before, VP $! VBD$ NP PP, ORG, The, Examiner]
Features

Headword of constituent
Examiner

Headword POS
NNP

Voice of the clause
Active

Subcategorization of pred
VP -> VBD NP PP

Named Entity type of constit
ORGANIZATION

First and last words of constit
The, Examiner

Linear position, clause re: predicate
before
Path Features

Path in the parse tree from the constituent to the predicate

\[ \text{NP} \rightarrow \text{S} \rightarrow \text{VP} \rightarrow \text{VBD} \]
Final feature vector

- For “The San Francisco Examiner”,
- Arg0, [issued, NP, Examiner, NNP, active, before, VP→VBD NP PP, ORG, The, Examiner, NP↑S↓VP↓VBD ]

- Other features could be used as well
  - sets of n-grams inside the constituent
  - other path features
    - the upward or downward halves
    - whether particular nodes occur in the path
3-step version of SRL algorithm

1. **Pruning**: use simple heuristics to prune unlikely constituents.
2. **Identification**: a binary classification of each node as an argument to be labeled or a NONE.
3. **Classification**: a 1-of-\(N\) classification of all the constituents that were labeled as arguments by the previous stage.
Why add Pruning and Identification steps?

• Algorithm is looking at one predicate at a time
• Very few of the nodes in the tree could possibly be arguments of that one predicate
• Imbalance between
  • positive samples (constituents that are arguments of predicate)
  • negative samples (constituents that are not arguments of predicate)
• Imbalanced data can be hard for many classifiers
• So we prune the very unlikely constituents first, and then use a classifier to get rid of the rest.

- Add sisters of the predicate, then aunts, then great-aunts, etc
- But ignoring anything in a coordination structure

37
A common final stage: joint inference

• The algorithm so far classifies everything locally – each decision about a constituent is made independently of all others
• But this can’t be right: Lots of global or joint interactions between arguments
  • Constituents in FrameNet and PropBank must be non-overlapping.
    • A local system may incorrectly label two overlapping constituents as arguments
  • PropBank does not allow multiple identical arguments
    • labeling one constituent ARG0
    • Thus should increase the probability of another being ARG1
How to do joint inference

• Reranking
  • The first stage SRL system produces multiple possible labels for each constituent
  • The second stage classifier the best global label for all constituents
  • Often a classifier that takes all the inputs along with other features (sequences of labels)
Semantic Role Labeling

Conclusion
Semantic Role Labeling

- A level of shallow semantics for representing events and their participants
  - Intermediate between parses and full semantics
- Two common architectures, for various languages
  - FrameNet: frame-specific roles
  - PropBank: Proto-roles
- Current systems extract by
  - Parsing sentence
  - Finding predicates in the sentence
    - For each one, classify each parse tree constituent