Semantic role labeling

Christopher Potts

CS 244U: Natural language understanding
Feb 2

With diagrams and ideas from Scott Wen-tau Tih, Kristina Toutanova, Dan Jurafsky, Sameer Pradhan, Chris Manning, Charles Fillmore, Martha Palmer, and Ed Loper.

Plan and goals

"There is perhaps no concept in modern syntactic and semantic theory which is so often involved in so wide a range of contexts, but on which there is so little agreement as to its nature and definition, as **THEMATIC ROLE***"* *(Dowty 1991:547)*

Plan and goals

> *"There is perhaps no concept in modern syntactic and semantic theory which is so often involved in so wide a range of contexts, but on which there is so little agreement as to its nature and definition, as* THEMATIC ROLE*"* *(Dowty 1991:547)*

1. Semantic roles as a useful shallow semantic representation
2. Resources for studying semantic roles:
   - FrameNet
   - PropBank
3. Semantic role labeling:
   - *Identification*: which phrases are role-bearing?
   - *Classification*: for role-bearing phrases, what roles do they play?
   - Evaluation
   - Tools
4. Approaches to semantic role labeling:
   - Models
   - Local features
   - Global and joint features

Common high-level roles

Definitions adapted from `http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/WhatIsASemanticRole.htm`

- **Agent**: a person or thing who is the doer of an event
- **Patient/Theme**: affected entity in the event; undergoes the action
- **Experiencer**: receives, accepts, experiences, or undergoes the effect of an action
- **Stimulus**: the thing that is felt or perceived
- **Goal**: place to which something moves, or thing toward which an action is directed.
- **Recipient** (sometimes grouped with Goal):
- **Source** (sometimes grouped with Goal): place or entity of origin
- **Instrument**: an inanimate thing that an Agent uses to implement an event
- **Location**: identifies the location or spatial orientation of a state or action
- **Manner**: how the action, experience, or process of an event is carried out.
- **Measure**: notes the quantification of an event

(Dowty 1991:§3 on how, ahh, extensive and particular these lists can become)

Examples

1. [$_{\textbf{Agent}}$ Doris] caught [$_{\textbf{Theme}}$ the ball] with [$_{\textbf{Instrument}}$ her mitt].
2. [$_{\textbf{Agent}}$ Sotheby's] offered [$_{\textbf{Recipient}}$ the heirs] [$_{\textbf{Theme}}$ a money-back guarantee].
3. [$_{\textbf{Stimulus}}$ The response] dismayed [$_{\textbf{Experiencer}}$ the group].
4. [$_{\textbf{Experiencer}}$ The group] disliked [$_{\textbf{Stimulus}}$ the response].
5. [$_{\textbf{Agent}}$ Kim] sent [$_{\textbf{Theme}}$ a stern letter] to [$_{\textbf{Goal}}$ the company].

Roles and morpho-syntactic diversity

Kim sent Sandy a letter. Kim sent a letter to Sandy. A letter was sent to Sandy by Kim. Sandy was sent a letter by Kim.	**Agent**: Kim, **Theme**: a letter, **Goal**: Sandy
Kim criticized the administration. Kim demanded the resignation. The compromise was rejected by Kim. Kim paid the check.	**Agent**: Kim, **Theme**: *
The storm frightened Sandy. Sandy feared the storm.	**Experiencer**: Sandy, **Stimulus**: the storm

Sam froze the ice cubes.	Jed ate the pizza.	Edith cut the bread easily.
⇕	⇕	⇕
The ice cubes froze.	Jed ate.	The bread cut easily.

Applications

The applications tend to center around places where we want a semantics that abstracts away from syntactic differences:

- Question answering (abstract Q/A alignment)
- Translation (abstract source/target alignment)
- Information extraction (grouping conceptually related events)
- High-level semantic summarization (what role does Obama/Gingrich/Romney typically play in media coverage?)
- . . .

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [Doris] hid [the money] [in the jar].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [Sam] broke [the flowerpot].

Let’s annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\textbf{Agent}}$ Doris] hid [$_{\textbf{Theme}}$ the money] [$_{\textbf{Location}}$ in the jar].
2. [$_{\textbf{Agent}}$ Sam] broke [$_{\textbf{Theme}}$ the flowerpot].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [The flowerpot] broke.

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [The storm] frightened [Sam].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [The speaker] told [a story].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].

Let’s annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [The watch] told [the time].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [$_{\text{Source}}$ The watch] told [$_{\text{Theme}}$ the time]. ???

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [$_{\text{Source}}$ The watch] told [$_{\text{Theme}}$ the time]. ???
7. [Italians] make [great desserts].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [$_{\text{Source}}$ The watch] told [$_{\text{Theme}}$ the time]. ???
7. [$_{\text{Agent}}$ Italians] make [$_{\text{Theme}}$ great desserts].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [$_{\text{Source}}$ The watch] told [$_{\text{Theme}}$ the time]. ???
7. [$_{\text{Agent}}$ Italians] make [$_{\text{Theme}}$ great desserts].
8. [Cookies] make [great desserts].

Let's annotate some data!

- **Agent**
- **Patient/Theme**
- **Experiencer**
- **Stimulus**
- **Goal**
- **Recipient**
- **Source**
- **Instrument**
- **Location**
- **Manner**
- **Measure**

1. [$_{\text{Agent}}$ Doris] hid [$_{\text{Theme}}$ the money] [$_{\text{Location}}$ in the jar].
2. [$_{\text{Agent}}$ Sam] broke [$_{\text{Theme}}$ the flowerpot].
3. [$_{\text{Theme}}$ The flowerpot] broke.
4. [$_{\text{Stimulus}}$ The storm] frightened [$_{\text{Experiencer}}$ Sam].
5. [$_{\text{Agent}}$ The speaker] told [$_{\text{Theme}}$ a story].
6. [$_{\text{Source}}$ The watch] told [$_{\text{Theme}}$ the time]. ???
7. [$_{\text{Agent}}$ Italians] make [$_{\text{Theme}}$ great desserts].
8. [$_{\text{Source?}}$ Cookies] make [$_{\text{Pred?}}$ great desserts]. ???

Challenges and responses

Challenges (from Dowty 1991:§3)

- Roles are hard to define/delimit.
- It can be hard to know which meaning contrasts are role-related and which belong to other domains, especially
  - lexical influences that subdivide roles very finely;
  - conceptual domains that cross-cut role distinctions;
  - information structuring

Responses

- Dowty (1991): argue forcefully for a tiny set of very general roles.
- PropBank: adopt a small set of roles as a matter of convenience, or to change the subject.
- FrameNet: different roles sets for different semantic domains, with some abstract connections between domains.

A brief history of semantic roles

1. Common in descriptive grammars dating back to the origins of linguistics, where they are used to informally classify predicates and case morphology.
2. Fillmore (1968) proposes an abstract theory of Case to capture underlying semantic relationships that affect/guide syntactic expression.
3. Syntacticians seek to discover patterns in how thematic (theta) roles are expressed syntactically (linking theory), and in how roles relate to each other and to other properties (e.g., animacy).
4. In linguistics, lexical semantics is currently a thriving area in which one of the central concerns is to find systematic connections between different argument realizations (Levin and Rappaport Hovav 2005).
5. Early SRL systems based on rule sets designed for specific texts (Simmons 1973; Riesbeck 1975).
6. The FrameNet project (Baker et al. 1998; Fillmore and Baker 2001) continues the research line begun by Fillmore.
7. Gildea and Jurafsky (2000, 2002) are among the very first to use resources like FrameNet to train general-purpose SRL systems.
8. PropBank (Palmer et al. 2005) provides comprehensive annotations for a section of the Penn Treebank, facilitating experiments of the sort that dominate NLP currently.

PropBank 1 (Palmer et al. 2005)

- A subset of the Wall Street Journal section of the Penn Treebank 2:
  - the version number is important; v1 and v3 will be misaligned in places
  - the subdirectory is `combined/wsj/`, which contains subdirectories of `.mrg` files
- 112,917 annotated examples (all centered around verbs)
- 3,257 unique verbs
- Core arguments numbered; peripheral arguments labeled
- Contains only verbs and their arguments
- Stand-off annotations:
  - `data/prop.txt`: one example per row, indexed to the Treebank files
  - `data/verbs.txt`: the list of verbs (by type)
  - `data/vloc.txt`: format

    `filename tree_no string_index verb_lemma`
  - `data/frames`: directory containing verbal frame files (XML)

PropBank frames and labels

Frame: `increase.01`

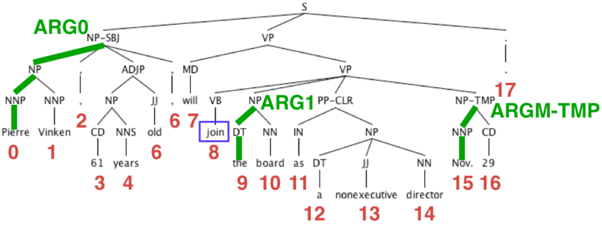
- name: go up incrementally
- vncls: 45.4 45.6
- ARG0 causer of increase (vntheta: Agent)
- ARG1 thing increasing (vntheta: Patient)
- ARG2 amount increased by, EXT or MNR (vntheta: Extent)
- ARG3 start point (vntheta: –)
- ARG4 end point (vntheta: –)

Examples

1. [$_{\text{ARG0}}$ The Polish government] [$_{\text{rel}}$ increased] [$_{\text{ARG1}}$ home electricity charges] [$_{\text{ARG2-EXT}}$ by 150%].
2. [$_{\text{ARG1}}$ The nation's exports] [$_{\text{rel}}$ increased] [$_{\text{2-EXT}}$ 4%] [$_{\text{4-2}}$ to $50.45 billion].
3. [$_{\text{ARG1}}$ Output] will be [$_{\text{2-MNR}}$ gradually] [$_{\text{rel}}$ increased] .

Example

First row of prop.txt	
Field	Value
wsj-filename	wsj/00/wsj_0001.mrg
sentence	0
terminal	8
tagger	gold
frameset	join.01
inflection	vf--a
proplabel	0:2-ARG0
proplabel	7:0-ARGM-MOD
proplabel	8:0-rel
proplabel	9:1-ARG1
proplabel	11:1-ARGM-PRD
proplabel	15:1-ARGM-TMP



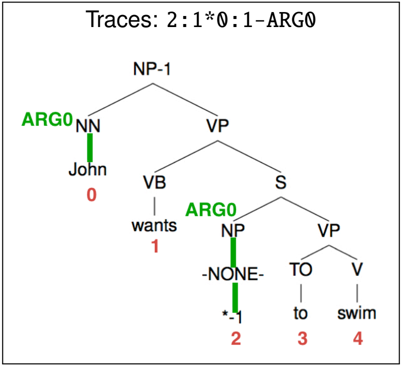
(only a subset of the ARG's labeled to avoid clutter)

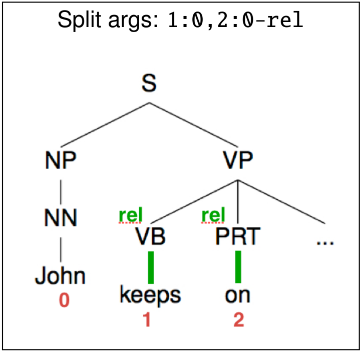
rel (verb) inflection fields ('-' means no value)		
1.	form:	i=inf g=gerund p=part v=finite
2.	tense:	f=future p=past n=present
3.	aspect:	p=perfect o=prog. b=both perfect & prog.
4.	person:	3=3rd person
5.	voice:	a=active p=passive

Label	
rel	the verb
ARGA	causative agent
ARGM	adjuncts
ARG0	generally subj
ARG1	generally dobj
ARG2	generally iobj
⋮	

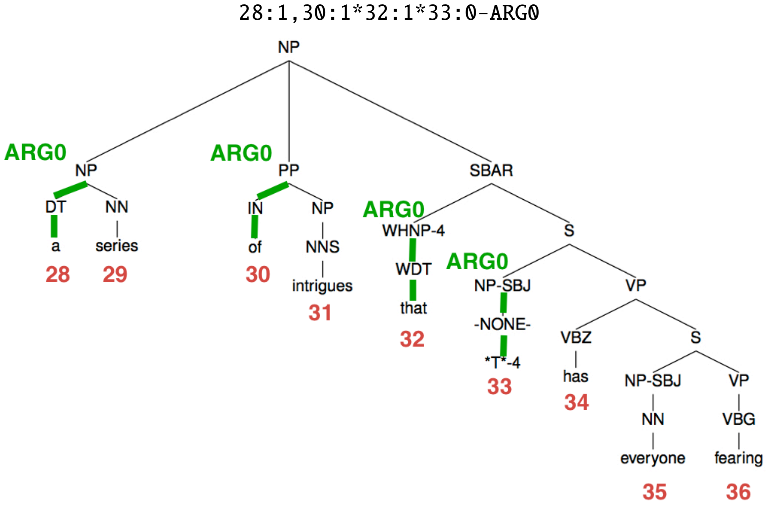
Label	
EXT	extent
DIR	direction
LOC	location
TMP	temporal
REC	reciprocal
PRD	predication
NEG	negation
MOD	modal
ADV	adverbial
MNR	manner
CAU	cause
PNC	purpose not cause
DIS	discourse

Trace paths and discontinuity





Trace chains and discontinuity combined



PropBank tools

- Web browser:
  http://verbs.colorado.edu/verb-index/index.php
- Stanford JavaNLP:
  http://nlp.stanford.edu/software/framenet.shtml
- Python NLTK:
  http://nltk.sourceforge.net/corpus.html#propbank-corpus
  http://nltk.googlecode.com/svn/trunk/doc/api/nltk.corpus.reader.propbank-module.html

NLTK interface to PropBank: example level

```
>>> import nltk.data; nltk.data.path = ['/path/to/penn-treebank2/'] + nltk.data.path
>>> from nltk.corpus import propbank
>>> pb = propbank.instances()
>>> len(pb)
112917
>>> len(propbank.verbs())
3257
########## Grab the first sentence, the one we looked at before:
>>> i0 = pb[0]
>>> i0.fileid
'wsj_0001.mrg'
>>> i0.sentnum
0
>>> i0.wordnum
8
>>> i0.inflection.tense
'f'
>>> i0.inflection.aspect
'_'
>>> i0.inflection.person
'_'
>>> i0.inflection.voice
'a'
>>> i0.roleset
'join.01'
>>> i0.arguments
((PropbankTreePointer(0, 2), 'ARG0'), (PropbankTreePointer(7, 0), 'ARGM-MOD'), \
(PropbankTreePointer(9, 1), 'ARG1'), (PropbankTreePointer(11, 1), 'ARGM-PRD'), \
(PropbankTreePointer(15, 1), 'ARGM-TMP')
```

NLTK interface to PropBank: example level (continued)

```
>>> i0.tree.pprint()
'(S
  (NP-SBJ
    (NP (NNP Pierre) (NNP Vinken))
    (, ,)
    (ADJP (NP (CD 61) (NNS years)) (JJ old))
    (, ,))
  (VP
    (MD will)
    (VP
      (VB join)
      (NP (DT the) (NN board))
      (PP-CLR (IN as) (NP (DT a) (JJ nonexecutive) (NN director)))
      (NP-TMP (NNP Nov.) (CD 29))))
  (. .))'

>>> inst.predicate.select(i0.tree)
Tree('VB', ['join'])

>>> i0.arguments[0][0].select(i0.tree).pprint()
'(NP-SBJ
  (NP (NNP Pierre) (NNP Vinken))
  (, ,)
  (ADJP (NP (CD 61) (NNS years)) (JJ old))
  (, ,))'
```

NLTK interface to PropBank: frame level

```
>>> from nltk.etree import ElementTree

>>> j = propbank.roleset('join.01')
>>> j
<Element 'roleset' at 0x3b781a0>

>>> ElementTree.tostring(j)
<roleset id="join.01" name="attach" vncls="22.1-2">
<roles>
  <role descr="agent, entity doing the tying" n="0">
    <vnrole vncls="22.1-2" vntheta="Agent" /></role>
  <role descr="patient, thing(s) being tied" n="1">
    <vnrole vncls="22.1-2" vntheta="Patient1" /></role>
  <role descr="instrument, string" n="2">
    <vnrole vncls="22.1-2" vntheta="Patient2" /></role>
</roles>

<example name="straight transitive">
...

>>> for r in j.findall('roles/role'): print 'ARG' + r.attrib['n'], r.attrib['descr']
ARG0 agent, entity doing the tying
ARG1 patient, thing(s) being tied
ARG2 instrument, string
```

A more advanced example: argument number and theta role alignment

```
from collections import defaultdict
from operator import itemgetter
import nltk.data; nltk.data.path = ['/path/to/penn-treebank2/'] + nltk.data.path
from nltk.corpus import propbank

def role_iterator():
    for verb in iter(propbank.verbs():
        index = 1
        while True:
            roleset_id = '%s.%s' % (verb, str(index).zfill(2))
            try:
                for role in propbank.roleset(roleset_id).findall('roles/role'):
                    yield role
                index += 1
            except ValueError:
                break

def view_arg_theta_alignment(n):
    counts = defaultdict(int)
    for role in role_iterator():
        if role.attrib['n'] == n:
            counts[role.attrib['descr']] += 1
    # View the result, sorted from most to least common theta role:
    for vtheta, count in sorted(counts.items(), key=itemgetter(1), reverse=True):
        print vtheta, count, round(float(count) / sum(counts.values()), 2)
```

Argument number and theta role alignment: examples

view_theta_alignment('0')			view_theta_alignment('1')			view_theta_alignment('2')		
causer	96	0.023	utterance	77	0.017	instrument	93	0.04
speaker	66	0.016	path	41	0.009	hearer	61	0.026
agent, causer	46	0.011	entity in motion	26	0.006	benefactive	53	0.023
causal agent	45	0.011	thing hit	25	0.006	EXT	42	0.018
entity in motion	41	0.01	victim	22	0.005	attribute	40	0.017
giver	35	0.008	commodity	21	0.005	source	36	0.015
causer, agent	31	0.007	impelled agent	21	0.005	destination	32	0.014
cause, agent	29	0.007	experiencer	19	0.004	attribute of arg1	29	0.012
creator	29	0.007	thing given	19	0.004	instrument, if separate from arg0	26	0.011
agent	20	0.005	topic	17	0.004	impelled action	22	0.009
thinker	19	0.005	thing changing	17	0.004	listener	21	0.009
cutter	19	0.005	Logical subject, patient, thing falling	17	0.004	end state	20	0.009
agent, hitter - animate only!	18	0.004	thing in motion	17	0.004	instrument, thing hit by or with	19	0.008
builder	17	0.004	food	16	0.004	location	19	0.008
describer	16	0.004	construction	15	0.003	EXT, amount fallen	18	0.008
Agent	15	0.004	subject	14	0.003	recipient	17	0.007
⋮			⋮			⋮		
2,454 vtheta types			2,842 vtheta types			1,125 vtheta types		

Dependency relations and PropBank core semantic roles

Dep	ARG0	ARG1	ARG2	ARG3	ARG4
nsubj	32,564	13,034	995	42	1
dobj	340	16,416	971	79	9
iobj	4	65	195	24	1
pobj	53	246	14	0	0

PropBank summary

Virtues

- Full gold-standard parses.
- Full coverage of a single collection of documents — one of the most heavily annotated document collections in the world.
- Different levels of role granularity.

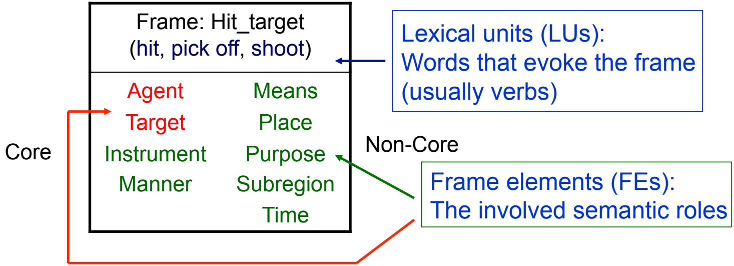
Limitations

- ARG2-5 overloaded. FrameNet (and VerbNet) both provide more fine-grained role labels
- WSJ too domain specific and too financial.
- Only verbs are covered; in language, nouns and adjs also have role arguments.

FrameNet

Data source: `https://framenet.icsi.berkeley.edu/fndrupal/current_status`

- Database of over 12,379 lexical units (7,963 full annotated).
- 1,135 distinct semantic frames (1,020 lexical; 115 non-lexical).
- 188,682 annotation sets (162,643 lexicographic; 26,039 full text).
- The ‘net’ part: words are related in numerous ways via their frames.



[$_{\text{Agent}}$ *Kristina*] ***hit*** [$_{\text{Target}}$ *Scott*] [$_{\text{Instrument}}$ *with a baseball*] [$_{\text{Time}}$ *yesterday*].

Background ideas (see Ruppenhofer et al. 2006)

Theoretical assumptions

- Word meanings are best understood in terms of the semantic/conceptual structures (frames) which they presuppose.
- Words and grammatical constructions that evoke frames and their elements.

Goals

- To discover and describe the frames that support lexical meanings.
- To provide names for the relevant elements of those frames
- To describe the syntactic/semantic valence of the words that fit the frames.
- To base the whole process on attestations from a corpus.

The focus is on the frames and their connections. Role labeling is necessary but secondary.

Example domains and frames

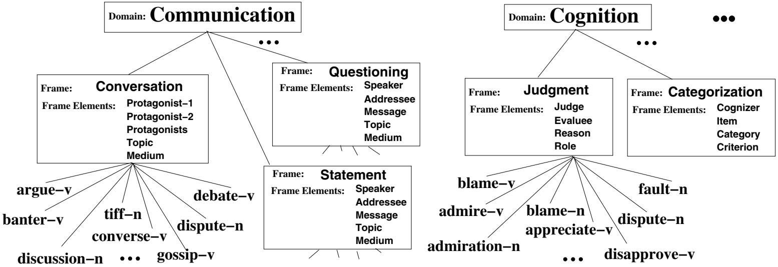


Figure 1
Sample domains and frames from the FrameNet lexicon.

(From Gildea and Jurafsky 2002:249)

From strings to frames

FrameNet Data Search for risk

Frame search results: Closest match is risk

Risk_scenario, Risky_situation

Lexical unit search results: Closest match is risk

Lexical Unit	Frame
risk.n	Daring
risk.n	Run_risk
risk.n	Risky_situation
risk.n	Being_at_risk
risk.v	Daring
risk.v	Run_risk
riskily.adv	Risky_situation
risky.a	Risky_situation

From strings to frames

Daring

Lexical Unit Index

Definition:

An Agent performs some Action which is considered imprudent. This frame is distinct from Attempt in that the danger that the Agent puts themselves in by performing the Action is profiled. The danger is not spelled out, but generally the Action has a possibility or likelihood of causing social or physical harm to the Agent.

I RISKED taking another look.

From strings to frames

FEs:

Core:

Action [Act]	This FE denotes the Action taken by the Agent.
Agent [age] **Semantic Type:** Sentient	The individual that performs the Action, resulting in danger to themself. I do n't think you should HAZARD the ascent in the dark , that 's all .

Non-Core:

Manner [man] **Semantic Type:** Manner	Any description of the risking action which is not covered by more specific FEs, including secondary effects (quietly, loudly), and general descriptions comparing events (the same way). In addition, it may indicate salient characteristics of an Agent that also affect the action (presumptuously, coldly, deliberately, eagerly, carefully). If you carelessly CHANCE going back there, you deserve what you get.
Place [pla] **Semantic Type:** Locative_relation	The location where the Agent risks doing the Action. Back in his room, he RISKED peeking inside.
Purpose [Purp] **Semantic Type:** State_of_affairs	The purpose for which the Action is performed. He RISKED a look back to check for pursuit.
Reason [Reas] **Semantic Type:** State_of_affairs	The Reason the Agent takes the risk. She RISKED a disguise since they already had a photo of her.
Time [tim] **Semantic Type:** Time	The time at which the Agent dares to perform the Action. At dawn, the Captain CHANCED opening the hatch.

From strings to frames

Frame-frame Relations:

Inherits from: Intentionally_act
Is Inherited by:
Perspective on:
Is Perspectivized in:
Uses:
Is Used by:
Subframe of:
Has Subframe(s):
Precedes:
Is Preceded by:
Is Inchoative of:
Is Causative of:
See also:

From strings to frames

Lexical Units:

chance.n, chance.v, dare.v, hazard.v, risk.n, risk.v, venture.v

Created by 664 on 07/10/2002 06:18:42 PDT Wed

Lexical Unit	LU Status	Lexical Entry Report	Annotation Report	Annotator ID	Created Date
chance.n	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:22:52 PDT Wed
chance.v	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:20:20 PDT Wed
dare.v	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:21:09 PDT Wed
hazard.v	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:20:55 PDT Wed
risk.n	**Created**	Lexical entry	Annotation	664	07/10/2002 06:22:08 PDT Wed
risk.v	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:21:53 PDT Wed
venture.v	**Finished_Initial**	Lexical entry	Annotation	664	07/10/2002 06:21:32 PDT Wed

From strings to frames

Lexical Entry

chance.n

Frame: Daring

Definition:

COD: a possibility of something happening.

Support(s): get, take
Controller(s): welcome

Frame Elements and Their Syntactic Realizations

The Frame Elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realization(s)
Action	(6)	DEN.-- (1) DNI.-- (4) VPto.Dep (1)
Agent	(7)	NP.Ext (5) DNI.-- (2)
Manner	(1)	AVP.Dep (1)

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns	
6 TOTAL	Action	Agent
(1)	DEN --	NP Ext
(2)	DNI --	DNI --
(2)	DNI --	NP Ext
(1)	VPto Dep	NP Ext
1 TOTAL	Agent	Manner
(1)	NP Ext	AVP Dep

From strings to frames

Annotation

Lexical Entry Daring

chance.n

Frame Element	Core Type
Action	Core
Agent	Core
Manner	Peripheral
Place	Peripheral
Purpose	Peripheral
Reason	Extra-Thematic
Time	Peripheral

Turn Colors On

- added
  1. and uh pittosporum that used to be the kind of things we could plant all the time you [$_{\text{Agent}}$you] [$_{\text{Manner}}$really] [take]$^{\text{Supp}}$ a *CHANCE*$^{\text{Target}}$ with them freezing
  2. On the positive side , most say the acting is great , and though [$_{\text{Agent}}$the film] `` does n't [take]$^{\text{Supp}}$ enormous [$_{\text{Action}}$*CHANCES*$^{\text{Target}}$] , `` it is nevertheless `` extremely satisfying `` (Denby , The New Yorker) . Slate 's Edelstein is more positive than most , praising the `` deliciously resonant dual setting : a Catskills summer community to which middle - class Jews from the city migrate to swim and eat and play mah-jongg , and the gathering hippies at nearby Woodstock . ``
  3. He can hardly bring himself to turn away , and sneaks back for another fix whenever [$_{\text{Agent}}$he] [gets]$^{\text{Supp}}$ the *CHANCE*$^{\text{Target}}$.[$_{\text{Action}}$DNI]
  4. so you really not making a sizeable profit so it 's not really lucrative to take *CHANCES*$^{\text{Target}}$ like that you know even though it it does exist[$_{\text{Agent}}$DNI][$_{\text{Action}}$DNI]
  5. i i do n't think it 's anything wrong for a doctor to refuse to i do n't care about the Hippocratic oath i do n't think they should have to treat a patient with AIDS if they do n't want to you know why take a *CHANCE*$^{\text{Target}}$ like that doctor in New York that got infected from a patient and you know she ended up i think[$_{\text{Agent}}$DNI][$_{\text{Action}}$DNI]
  6. uh-huh well the best thing about it is that you can uh try something if you do n't like it shoot move on to something else that 's the way i 've uh looked at the whole thing here [$_{\text{Agent}}$i] 'll [take]$^{\text{Supp}}$ a *CHANCE*$^{\text{Target}}$ if i do n't like it i 'll go someplace else do something different[$_{\text{Action}}$DNI]
  7. [$_{\text{Agent}}$They] [welcome]$^{\text{Ctrlr}}$ the *CHANCE*$^{\text{Target}}$ [$_{\text{Action}}$to belong] , to become self - sufficient , to regain their self - esteem and confidence .

Annotator ID(s): 571, 976

Full text annotations

From `https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=fulltextIndex`

- American National Corpus Texts
- AQUAINT Knowledge-Based Evaluation Texts
- LUCorpus-v0.3
- Miscellaneous
- Texts from Nuclear Threat Initiative website, created by Center for Non-Proliferation Studies
- Wall Street Journal Texts from the PropBank Project

Gildea and Jurafsky (2000, 2002) FrameNet experiment format

From their training set:

```
body/action/arch.v.ar:<S TPOS="30621249"> <C TARGET="y"> Arch/VVB </C> <C FE="Agt" PT="CNI">
(TOP (S (VP (VP (VB Arch) (NP (PRP$ your) (NN back)) (ADVP (ADVP (RB as) (JJ high)) (SBAR (I

body/action/arch.v.ar:<S TPOS="67141515"> <T TYPE="Canonical"> </T> She/PNP snatched/VVD Bus
(TOP (S (S (NP (PRP She)) (VP (VBD snatched) (NP (NN Buster)) (PP (IN from) (NP (PRP$ his) (

...

body/action/bat.v.ar:<S TPOS="77171143"> <C TYPE="Blend"> </C> <C FE="Agt"> The/AT0 receptic
(TOP (S (NP (DT The) (NN receptionist)) (VP (VBD had) (VP (ADVP (RB obviously)) (VBN recogni

body/action/bat.v.ar:<S TPOS="69048344"> Did/VDD <C FE="Agt"> saints/NN2 </C> ever/AV0 <C TA
(TOP (SQ (VBD Did) (NP (NNS saints)) (ADVP (RB ever)) (VP (VP (VB bat) (NP (PRP$ their) (NNS

...

body/action/bend.v.ar:<S TPOS="25399472"> <C FE="Agt"> You/PNP </C> may/VM0 <C TARGET="y"> b
(TOP (S (NP (PRP You)) (VP (MD may) (VP (VB bend) (NP (DT the) (JJR lower) (NN arm)) (NP (DT
```

FrameNet summary

Virtues

- Many levels of analysis.
- Different parts of speech (not just verbs).
- Diverse document collection.
- A rich lexical resource, not just for SRL.

Limitations (some addressed by the new full-text annotations)

- Example sentences are chosen by hand (non-random).
- Complete sentences not labeled
- No gold-standard parses or other annotations.
- A work in progress with sometimes surprising gaps.

Other corpora

- FrameNets in other languages:
  https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages
- VerbNet:
  http://verbs.colorado.edu/~mpalmer/projects/verbnet.html
- NomBank (extends PropBank with NP-internal annotations):
  http://nlp.cs.nyu.edu/meyers/NomBank.html
- Korean PropBank:
  http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2006T03
- Chinese Propbanks:
  http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2005T23
  http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2008T07
- CoNNL-2005 shared task PropBank subset (tabular format):
  http://www.lsi.upc.edu/~srlconll/soft.html
- Senseval 3 SRL (FrameNet subset):
  http://www.clres.com/SensSemRoles.html
- SemEval 2007 (FrameNet, NomBank, PropBank, Arabic)
  http://nlp.cs.swarthmore.edu/semeval/tasks/index.php

SRL tasks

Identification: which phrases are role-bearing?

- Necessary for real-world tasks, where phrases are unlikely to be identified as role-bearing.
- Role-bearing phrases need not be constituents, or even necessarily contiguous, making the search space enormous (2^n for n words, though most candidates will be absurd).

Classification: for role-bearing phrases, what roles do they play?

- Highly dependent on the underlying role set.
- Also a very large search space: $\approx 20^m$ for m arguments, assuming 20 candidate labels.

Evaluation: very involved and tricky to get right

- In identification, how do we score overlap/containment/subsumption?
- Should classification scores be influenced by identification errors?
- Are some argument-tyles more important than others?
- Are some mis-classifications worse than others?

Evaluation in Toutanova et al. 2008:§3.2

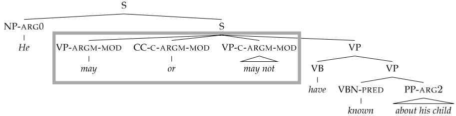
CORE: only core args

CoarseARGM: adjuncts all collapsed to ARGM

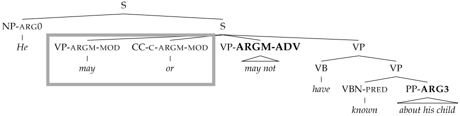
ALL: all args

Argument ID: classify word sets as role-bearing or not; all labels mapped to ARG or NONE

Argument Cls: assign roles to role-bearing phrases



(a) Correct labeling.



(b) Guessed labeling.

Location	Labeling
Correct	{0}-ARG0, {1, 2, 3, 4}-ARGM-MOD, {7, 8, 9}-ARG2
Guessed	{0}-ARG0, {1, 2}-ARGM-MOD, {3, 4}-ARGM-ADV {7, 8, 9}-ARG3

(c) Labelings of spans.

Task	CORE		COARSEARGM		ALL	
	F1	Acc.	F1	Acc.	F1	Acc.
Argument ID	100.0	100.0	57.1	0.0	57.1	0.0
Argument CLS	50.0	0.0	50.0	0.0	50.0	0.0
Argument ID&CLS	50.0	0.0	28.6	0.0	28.6	0.0

(d) Scoring measures.

Figure 2
Argument-based scoring measures for the guessed labeling.

Acc: whole frame accuracy

tp: gold ≠ NONE & guess = gold
fp: guess ≠ NONE & guess ≠ gold
fn: gold ≠ NONE & guess ≠ gold

p: tp / (tp + fp)
r: tp / (tp + fn)
F1: (2*p*r) / (p+r)

Evaluation in Toutanova et al. 2008:§3.2

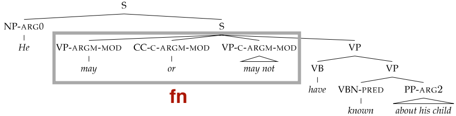
CORE: only core args

CoarseARGM: adjuncts all collapsed to ARGM

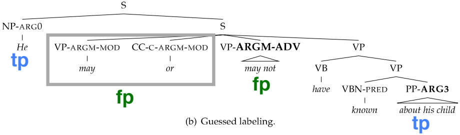
ALL: all args

Argument ID: classify word sets as role-bearing or not; all labels mapped to ARG or NONE

Argument Cls: assign roles to role-bearing phrases



(a) Correct labeling.



(b) Guessed labeling.

Location	Labeling
Correct	{0}-ARG0, {1, 2, 3, 4}-ARGM-MOD, {7, 8, 9}-ARG2
Guessed	{0}-ARG0, {1, 2}-ARGM-MOD, {3, 4}-ARGM-ADV {7, 8, 9}-ARG3

(c) Labelings of spans.

Task	CORE		COARSEARGM		ALL	
	F1	Acc.	F1	Acc.	F1	Acc.
Argument ID	100.0	100.0	57.1	0.0	57.1	0.0
Argument CLS	50.0	0.0	50.0	0.0	50.0	0.0
Argument ID&CLS	50.0	0.0	28.6	0.0	28.6	0.0

(d) Scoring measures.

Figure 2
Argument-based scoring measures for the guessed labeling.

Acc: whole frame accuracy

tp: gold ≠ NONE & guess = gold
fp: guess ≠ NONE & guess ≠ gold
fn: gold ≠ NONE & guess ≠ gold

p: tp / (tp + fp)
r: tp / (tp + fn)
F1: (2*p*r) / (p+r)

p = tp / (tp + fp) = 2 / (2 + 2)
r = tp / (tp + fn) = 2 / (2 + 1)
f1 = (2 * 0.5 * 0.67) / (0.5 + 0.67) = 0.571

CoNNL evaluation (Carreras and Màrquez 2005)

- Distributed as a Perl script from
  `http://www.lsi.upc.edu/~srlconll/soft.html`
- Essentially the same as the ARGUMENT ID&CLS metric of Toutanova et al. 2008: "For an argument to be correctly recognized, the words spanning the argument as well as its semantic role have to be correct."
- Verbs are excluded from the evaluation, since they are generally the targets.
- For CoNNL, co-indexed arguments are treated as separate arguments

  [ARG1 The deregulation] of railroads [R-ARG1 that] [PRED began] enabled shippers to bargain for transportation.

  whereas for Toutanova et al. they are treated as single C- related constituents to be assigned a single role:

  [ARG1 The deregulation] of railroads [C-ARG1 that] [PRED began] enabled shippers to bargain for transportation.

Tools (pause here for demos)

SwiRL: http://www.surdeanu.name/mihai/swirl/

(Surdeanu and Turmo 2005; Surdeanu et al. 2007)

```
The glass broke .

( S1 0
    ( S 1
        ( NP 1 { B-A1-2 }
            ( DT 0 The the O  )
            ( NN 1 glass glass O  ) )
        ( VP 0
            ( VBD 2 broke break O  ) )
        ( . 3 . . O  ) ) )

DT     (S1(^S(NP*      "the"         0 (A1*
NN     ^*)             "glass"       0 *)
VBD    (^VP^*)         "break"       1 *
.      *))             "."           0 *
```

Illinois: http://cogcomp.cs.illinois.edu/demo/srl/

```
The glass broke.

The      breaker [A0]   (S1 (S (NP (DT the)
glass                              (NN glass))
broke    V: break              (VB (VBD broke))
.                              (. .)))
```

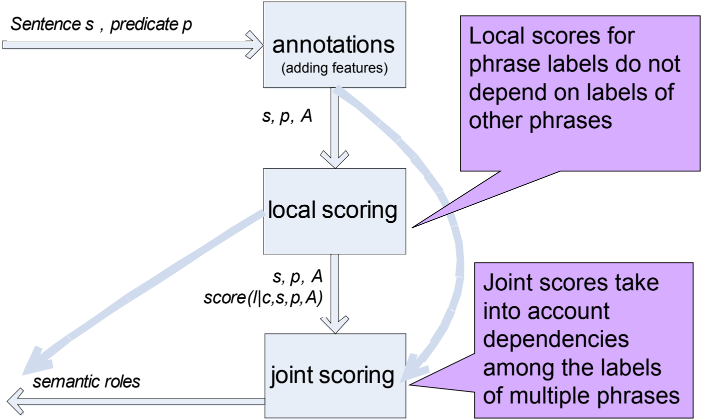
Approaches to SRL

Many different kinds of models have been used for SRL:

- Gildea and Jurafsky (2002): direct Bayesian estimates using rich morpho-synactic features
- Pradhan et al. (2004): SVMs with very rich features
- Punyakanok et al. (2004, 2005): systems of hand-built, categorical rules with an integer linear programming solver
- Shallow morph-syntactic features (CoNNL-2005 systems)
- Toutanova et al. (2008): inter-label dependencies (discussed extensively here)

For many additional references, see Yih and Toutanova 2007.

Basic architecture



(From Yih and Toutanova 2007)

Local classifiers

Definition (Local SRL classifier)

- t: a tree
- v: a target predicate node in t
- L: a mapping from nodes in t to semantic roles (including NONE)
- $Id(L)$: the mapping that is just like L except all non-NONE values are ARG

The probability of L is given by

$$P_{SRL}^{LOCAL}(L|t,v) = \prod_{n_i \in t} P_{ID}\Big(Id(l_i)|t,v\Big) \times \prod_{n_i \in t} P_{CLS}\Big(l_i|t,v,Id(l_i)\Big)$$

For classification, pick the L that maximizes this product.

- Toutanova et al. (2008:§4) train MaxEnt models for each term in the product and then multiply the predicted distributions together to obtain $P_{SRL}^{LOCAL}(L|t,v)$. The feature sets are the same for both models.
- Because the maximal labeling could involve overlapping spans and role assignments, they develop a dynamic programming algorithm that memoizes scores moving from the leaves to the root (§4.2). The gains are modest, though.

Baseline features

Standard Features (Gildea and Jurafsky 2002)
PHRASE TYPE: Syntactic Category of node
PREDICATE LEMMA: Stemmed Verb
PATH: Path from node to predicate
POSITION: Before or after predicate?
VOICE: Active or passive relative to predicate
HEAD WORD OF PHRASE
SUB-CAT: CFG expansion of predicate's parent
Additional Features (Surdeanu et al. 2003; Pradhan et al. 2004)
FIRST/LAST WORD
LEFT/RIGHT SISTER PHRASE-TYPE
LEFT/RIGHT SISTER HEAD WORD/POS
PARENT PHRASE-TYPE
PARENT POS/HEAD-WORD
ORDINAL TREE DISTANCE: Phrase Type concatenated with length of PATH feature
NODE-LCA PARTIAL PATH: Path from constituent to lowest common ancestor with predicate
PP PARENT HEAD WORD: If parent is a PP, parent's head word
PP NP HEAD WORD/POS: For a PP, the head Word / POS of its rightmost NP
Selected Pairs (Xue and Palmer 2004)
PREDICATE LEMMA & PATH
PREDICATE LEMMA & HEAD WORD
PREDICATE LEMMA & PHRASE TYPE
VOICE & POSITION
PREDICATE LEMMA & PP PARENT HEAD WORD

Figure 3
Baseline features.

(Toutanova et al. 2008:172)

Handling displaced constituents (Toutanova et al. 2008:§4.1)

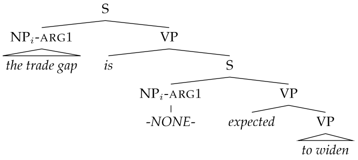
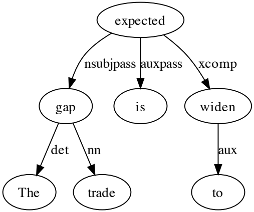


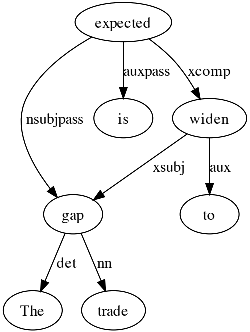
Figure 4
Example of displaced arguments.

Numerous errors caused by displaced constituents. Response is to have a feature MISSING SUBJECT and a PATH feature, so that the model establishes the associations.

Basic Stanford dependencies



Collapsed Stanford dependencies



Joint model features (Toutanova et al. 2008:174–176)

Task	CORE		COARSEARGM		ALL	
	F1	Acc.	F1	Acc.	F1	Acc.
ID	92.3	83.7	92.4	79.0	92.4	79.0
CLS	98.0	96.8	98.1	95.9	95.7	90.8
ID&CLS	**90.5**	**81.4**	90.6	76.2	**88.4**	**72.3**

(a) Summary performance results.

Correct	Guessed ARG0	ARG1	ARG2	ARG3	ARG4	ARG5	ARGM	NONE	F-Measure
ARG0	2912	22	1	0	0	0	4	248	91.7
ARG1	69	3964	15	1	1	0	12	302	91.8
ARG2	7	25	740	3	2	0	9	151	82.4
ARG3	1	5	3	83	1	0	5	36	70.3
ARG4	0	1	3	0	63	0	0	7	88.1
ARG5	0	0	0	0	0	5	0	0	100.0
ARGM	0	7	10	0	0	0	2907	322	91.0
NONE	173	248	87	15	2	0	204	–	–

(b) COARSEARGM confusion matrix.

Correct	Guessed ADV	CAU	DIR	DIS	EXT	LOC	MNR	MOD	NEG	PNC	PRD	REC	TMP	CORE	NONE	F_1
ADV	295	3	0	13	3	10	35	0	0	5	0	0	20	1	51	71.3
CAU	0	48	0	1	0	2	3	0	0	2	0	0	3	0	6	81.4
DIR	0	0	40	0	0	0	6	0	0	0	0	0	1	2	25	61.1
DIS	13	0	0	214	0	3	2	0	0	0	0	0	8	0	31	79.9
EXT	2	0	0	1	17	0	5	0	0	0	0	0	0	2	5	63.0
LOC	4	0	0	2	0	251	3	0	0	2	1	0	8	1	45	77.5
MNR	17	0	5	0	2	12	196	0	0	0	0	0	4	5	66	65.8
MOD	0	0	0	0	0	0	0	453	0	0	0	0	0	0	2	99.4
NEG	0	0	0	0	0	0	0	0	200	0	0	0	0	0	2	99.0
PNC	4	2	0	0	0	1	0	0	0	59	0	0	5	3	26	64.8
PRD	1	0	1	0	0	0	0	0	0	1	1	0	0	1	0	28.6
REC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.0
TMP	23	0	0	4	0	11	3	0	1	1	0	0	874	2	61	88.7
CORE	4	0	2	2	0	0	6	0	0	7	0	0	9	7927	744	92.3
NONE	28	0	9	28	0	41	30	3	1	5	0	0	59	525	–	–

(c) Modifier arguments confusion matrix.

Figure 7
Performance measures for local model using all local features and enforcing the non-overlapping constraint. Results are on Section 23 using gold standard parse trees.

- Higher precision than recall.
- Most mistakes involve NONE. (Not surprising to me; I am often surprised at what does and doesn't get role-labeled.)
- Few CORE ARG labels are swapped.
- More MODIFIER labels are swapped.
- Few CORE ARG/MODIFIER swaps.

Joint model (Toutanova et al. 2008:§5)

1. Use the local model to generate the top *n* non-overlapping labeling functions L, via a variant of the dynamic programming algorithm used to ensure non-overlap (§4.2).
2. Use a MaxEnt model to re-rank the top *n* labeling sequences via values $P^r_{SRL}(L|t, v)$.
3. Obtain final scores:

Definition (Joint model scoring)

$$P_{SRL}(L|t, v) = \left(P^{LOCAL}_{SRL}(L|t, v)\right)^{\alpha} \times P^r_{SRL}(L|t, v)$$

where α is a tuntable parameter (they used 1.0)

4. Classification: pick the L that maximizes this scoring function.

Joint model features (Toutanova et al. 2008:§5.2)

- All the features from the local models.
- *Whole Label Sequence* features of arbitrary length:

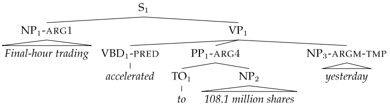


Figure 9
An example tree from Propbank with semantic role annotations, for the sentence *Final-hour trading accelerated to 108.1 million shares yesterday*.

Basic: [voice:active, ARG1, PRED, ARG4, ARGM-TMP]
Lemma: [voice:active, lemma:accelerate, ARG1, PRED, ARG4, ARGM-TMP]
Generic: [voice:active, ARG, PRED, ARG, ARG]
POS: [voice:active, NP-ARG0, PRED, NP-ARG1, PP-ARG2]
POS+lemma: [voice:active,lemma:offer, NP-ARG0, PRED, NP-ARG1, PP-ARG2]

- *Repetition features*: POS-annotated features indicating when the same ARG occurs multiple times.

Joint model results (Toutanova et al. 2008:§5.4)

- LOCAL: the local model and results given above
- JOINTLOCAL: a joint model using only the LOCAL features
- LABELSEQ: a joint model using only the LOCAL features and the whole labels sequence features
- ALLJOINT: a joint model using the LOCAL features, the whole labels sequence features, and the repetition features

Model	# Features	CORE		COARSEARGM		ALL	
		F1	Acc.	F1	Acc.	F1	Acc.
LOCAL	5,201K	90.5	81.4	90.6	76.2	88.4	72.3
JOINTLOCAL	2,193K	90.9	82.6	91.1	78.3	88.9	74.3
LABELSEQ	2,357K	92.9	86.1	92.6	81.4	90.4	77.0
ALLJOINT	2,811K	94.0	87.6	93.4	82.7	**91.2**	**78.3**

Figure 10
Performance of local and joint models on ID&CLS on Section 23, using gold-standard parse trees. The number of features of each model is shown in thousands.

- The pattern of errors for the Joint models is broadly the same as for the Local models, though there are notable points of improvement (p. 183).
- Toutanova et al. (2008:§6) show that the Joint-model approach is robust for automatic (and therefore error-ridden) parses as well.

Conclusions

- Semantic roles are distinct from syntactic roles.
- Semantic roles capture usefully abstract semantic information (despite the challenges of assigning them).
- SRL reached a peak of popularity around 2005-2006, and it is currently on the wane, but this is probably just because system performance is still not great.
- There are many SRL models, but a lot of commonalities in the underlying feature sets.
- Even if we manage to do complete and accurate semantic composition (stay tuned for Bill, Percy Liang, and Richard Socher!) SRL will remain valuable where a coarse-grained semantics is called for.

References I

Baker, Collin F.; Charles J. Fillmore; and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of COLING-ACL*, 86–90. Montreal: Association for Computational Linguistics.

Carreras, Xavier and Luís Màrquez. 2005. Introduction to the CoNLL-2005 shared task: Semantic role labeling. In *Proceedings of CoNLL*, 152–164. Ann Arbor, MI.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3):547–619.

Fillmore, Charles J. 1968. The case for Case. In Emmon Bach and Robert T. Harms, eds., *Universals in Linguistic Theory*, 1–88. New York: Holt, Rinehart, and Winston.

Fillmore, Charles J. and Collin F. Baker. 2001. Frame semantics for text understanding. In *Proceedings of the WordNet and Other Lexical Resources Workshop*, 59–64. Pittsburgh, PA: Association for Computational Linguistics.

Gildea, Daniel and Daniel Jurafsky. 2000. Automatic labeling of semantic roles. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, 512–520. Hong Kong: Association for Computational Linguistics. doi:\bibinfo{doi}{10.3115/1075218.1075283}. URL `http://www.aclweb.org/anthology/P00-1065`.

Gildea, Daniel and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics* 28(3):245–288.

Levin, Beth and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.

Palmer, Martha; Paul Kingsbury; and Daniel Gildea. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics* 31(1):71–105.

Pradhan, Sameer S.; Wayne H. Ward; Kadri Hacioglu; James H. Martin; and Dan Jurafsky. 2004. Shallow semantic parsing using support vector machines. In Daniel Marcu Susan Dumais and Salim Roukos, eds., *HLT-NAACL 2004: Main Proceedings*, 233–240. Boston, Massachusetts, USA: Association for Computational Linguistics.

Punyakanok, Vasin; Dan Roth; and Scott Wen-tau Yih. 2005. The necessity of syntactic parsing for semantic role labeling. In *Proceedings of IJCAI*, 1117–1123. Acapulco, Mexico.

References II

Punyakanok, Vasin; Dan Roth; Scott Wen-tau Yih; Dav Zimak; and Tuancheng Tu. 2004. Semantic role labeling via generalized instances over classifiers. In *Proceedings of CoNLL*, 130–133. Boston.

Riesbeck, Christopher K. 1975. Conceptual analysis. In Roger C. Schank, ed., *Conceptual Information Processing*. North-Holland and Elsevier.

Ruppenhofer, Josef; Michael Ellsworth; Miriam R. L. Petruck; Christopher R. Johnson; and Jan Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. Berkeley, CA: International Computer Science Institute.

Simmons, Robert F. 1973. Semantic networks: Their computation and use for understanding English sentences. In Roger Schank and Kenneth Mark Colby, eds., *Computer Models of Thought and Language*, 61–113. San Francisco: W. H. Freeman and Company.

Surdeanu, Mihai; Sanda Harabagiu; John Williams; and Paul Aarseth. 2003. Using predicate-argument structures for information extraction. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, 8–15. Sapporo, Japan: Association for Computational Linguistics. doi:\bibinfo{doi}{10.3115/1075096.1075098}. URL http://www.aclweb.org/anthology/P03-1002.

Surdeanu, Mihai; Luís Màrquez; Xavier Carreras; and Pere R. Comas. 2007. Combination strategies for semantic role labeling. *Journal of Artificial Intelligence Research* 29:105–151.

Surdeanu, Mihai and Jordi Turmo. 2005. Semantic role labeling using complete syntactic analysis. In *Proceedings of CoNLL 2005 Shared Task*.

Toutanova, Kristina; Aria Haghighi; and Christopher D. Manning. 2008. A global joint model for semantic role labeling. *Computational Linguistics* 34(2):161–191.

Xue, Niawen and Martha Palmer. 2004. Calibrating features for semantic role labeling. In *Proceeedings of EMNLP*, 88–94. Barcelona, Spain.

Yih, Scott Wen-tau and Kristina Toutanova. 2007. Automatic semantic role labeling. Tutorial at AAAI-07, URL http://research.microsoft.com/apps/pubs/default.aspx?id=101987.