Homework 2 Report

Min/Max 45.00 / 100.00
Average 80.50
Median 86.25

• A lot of people failed to correctly implement HMM. Their output sequences are wrong.

• For the second question, a lot of people failed to produce the two parses for the blind test sentences, or their probabilities for those parses are incorrect.

• TA has office hours tomorrow

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greater than 100</td>
<td>0</td>
</tr>
<tr>
<td>90 - 100</td>
<td>7</td>
</tr>
<tr>
<td>80 - 89</td>
<td>3</td>
</tr>
<tr>
<td>70 - 79</td>
<td>2</td>
</tr>
<tr>
<td>60 - 69</td>
<td>2</td>
</tr>
<tr>
<td>50 - 59</td>
<td>0</td>
</tr>
<tr>
<td>40 - 49</td>
<td>2</td>
</tr>
<tr>
<td>30 - 39</td>
<td>0</td>
</tr>
<tr>
<td>20 - 29</td>
<td>0</td>
</tr>
<tr>
<td>10 - 19</td>
<td>0</td>
</tr>
<tr>
<td>0 - 9</td>
<td>0</td>
</tr>
<tr>
<td>Less than 0</td>
<td>0</td>
</tr>
</tbody>
</table>
In contrast, Homework 1

<table>
<thead>
<tr>
<th>Minimum Value</th>
<th>75.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Value</td>
<td>100.00</td>
</tr>
<tr>
<td>Average</td>
<td>96.69</td>
</tr>
<tr>
<td>Median</td>
<td>100.00</td>
</tr>
</tbody>
</table>

90 - 100: 14
80 - 89: 1
70 - 79: 1

Project Team Update

- 4 three-person teams
- 2 two-person teams
Midterm Notes

- See course homepage

Semantic Role Labeling

Introduction
Motivation: 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...

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
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.

More specific: buyer | More general: proto-agent

Semantic Role Labeling

Semantic Roles
Getting to semantic roles

Reified first order logic event representation:

Sasha broke the window
Pat opened the door

\[ \exists e, x, y \, \text{Breaking}(e) \land \text{Breaker}(e, \text{Sasha}) \land \text{BrokenThing}(e, y) \land \text{Window}(y) \]
\[ \exists e, x, y \, \text{Opening}(e) \land \text{Opener}(e, \text{Pat}) \land \text{OpenedThing}(e, y) \land \text{Door}(y) \]

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 Openers.
- 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>
Example usages of “break”  (AGENT, THEME, INSTRUMENT)

John broke the window

John broke the window with a rock.

The rock broke the window.

The window broke.

The window was broken by John.

<table>
<thead>
<tr>
<th>Thematic Role</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGENT</td>
<td>The volitional causer of an event</td>
</tr>
<tr>
<td>EXPERIENCER</td>
<td>The experient of an event</td>
</tr>
<tr>
<td>FORCE</td>
<td>The non-volitional causer of the event</td>
</tr>
<tr>
<td>THEME</td>
<td>The participant most directly affected by an event</td>
</tr>
<tr>
<td>RESULT</td>
<td>The end product of an event</td>
</tr>
<tr>
<td>CONTENT</td>
<td>The proposition or content of a propositional event</td>
</tr>
<tr>
<td>INSTRUMENT</td>
<td>An instrument used in an event</td>
</tr>
<tr>
<td>BENEFICIARY</td>
<td>The beneficiary of an event</td>
</tr>
<tr>
<td>SOURCE</td>
<td>The origin of the object of a transfer event</td>
</tr>
<tr>
<td>GOAL</td>
<td>The destination of an object of a transfer event</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
INSTRUMENT/Subject, THEME/Object
THEME/Subject
Example usages of “give” 

(AGENT, THEME, ?)

Doris gave the book to Cory.

Doris gave Cory the book.

<table>
<thead>
<tr>
<th>Thematic Role</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGENT</td>
<td>The volitional causor of an event</td>
</tr>
<tr>
<td>EXPERIENCER</td>
<td>The experiencer of an event</td>
</tr>
<tr>
<td>FORCE</td>
<td>The non-volitional causor of the event</td>
</tr>
<tr>
<td>THEME</td>
<td>The participant most directly affected by an event</td>
</tr>
<tr>
<td>RESULT</td>
<td>The end product of an event</td>
</tr>
<tr>
<td>CONTENT</td>
<td>The proposition or content of a propositional event</td>
</tr>
<tr>
<td>INSTRUMENT</td>
<td>An instrument used in an event</td>
</tr>
<tr>
<td>BENEFICIARY</td>
<td>The beneficiary of an event</td>
</tr>
<tr>
<td>SOURCE</td>
<td>The origin of the object of a transfer event</td>
</tr>
<tr>
<td>GOAL</td>
<td>The destination of an object of a transfer event</td>
</tr>
</tbody>
</table>

Verb alternation (multiple argument structure realizations)

Doris gave the book to Cary.

AGENT THEME GOAL

Break: AGENT, INSTRUMENT, or THEME as subject

Doris gave Cary the book.

AGENT GOAL THEME

Give: THEME and GOAL in either order
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 (with resources)

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 Lemmas](#) (check out “break”)

22
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)
**PropBank Frame Files**

**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:

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

ArgM-TMP | when? | yesterday evening, now
LOC | where? | at the museum, in San Francisco
DIR | where to/from? | down, to Bangkok
MNR | how? | clearly, with much enthusiasm
PRP/CAU | why? | because ... , in response to the ruling
REC | | themselves, each other
ADV | | miscellaneous
PRD | secondary predication | ...ate the meat raw

PropBanking a Sentence

Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.
The same parse tree PropBanked

Expect(Analysts, GM-J pact) give(GM-J pact, US car maker, 30% stake)

Annotated PropBank Data

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

<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_ARG0] [decision_REL]
[to say look I don’t want to go through this anymore_ARG1]”

Example within an LVC: Make a decision
← “...[the President_ARG0] [made_REL-LVB]
the [fundamentally correct_ARGM-ADJ]
[decision_REL] [to get on offense_ARG1]”

Semantic Role Labeling

FrameNet
Capturing descriptions of the same event by different nouns/verbs

\[ \text{Arg}_1 \text{ The price of bananas] increased [Arg}_2 5\%.} \\
\text{Arg}_1 \text{ The price of bananas] rose [Arg}_2 5\%.} \\
\text{There has been a [Arg}_2 5\%] rise [Arg}_1 \text{ in the price of bananas].}

FrameNet

- Roles in PropBank are specific to a verb
- Role in FrameNet are specific to a **frame**: a background knowledge structure that defines a set of frame-specific semantic roles, called **frame elements**,  
  - includes a set of predicates that use these roles  
  - each word evokes a frame and profiles some aspect of the frame
- **Framenet Search** (e.g., “increase”)
The “Change position on a scale” Frame

This frame consists of words that indicate the change of an ITEM’s position on a scale (the ATTRIBUTE) from a starting point (INITIAL VALUE) to an end point (FINAL VALUE)

- [ITEM Oil] rose [ATTRIBUTE in price] [DIFFERENCE by 2%].
- [ITEM It] has increased [FINAL_STATE to having them 1 day a month].
- [ITEM Microsoft shares] fell [FINAL_VALUE to 7 5/8].
- [ITEM Colon cancer incidence] fell [DIFFERENCE by 50%] [GROUP among men].

  a steady increase [INITIAL_VALUE from 9.5] [FINAL_VALUE to 14.3] [ITEM in dividends]

  a [DIFFERENCE 5%] [ITEM dividend] increase...

The “Change position on a scale” Frame

<table>
<thead>
<tr>
<th>VERBS</th>
<th>dwindle</th>
<th>move</th>
<th>soar</th>
<th>escalation</th>
<th>shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>advance</td>
<td>edge</td>
<td>mushroom</td>
<td>swell</td>
<td>explosion</td>
<td>tumble</td>
</tr>
<tr>
<td>climb</td>
<td>explode</td>
<td>plummet</td>
<td>swing</td>
<td>fall</td>
<td></td>
</tr>
<tr>
<td>decline</td>
<td>fall</td>
<td>reach</td>
<td>triple</td>
<td>fluctuation</td>
<td>gain</td>
</tr>
<tr>
<td>decrease</td>
<td>fluctuate</td>
<td>rise</td>
<td>tumble</td>
<td></td>
<td>growth</td>
</tr>
<tr>
<td>diminish</td>
<td>gain</td>
<td>rocket</td>
<td></td>
<td></td>
<td>increasing</td>
</tr>
<tr>
<td>dip</td>
<td>grow</td>
<td>shift</td>
<td></td>
<td></td>
<td>hike</td>
</tr>
<tr>
<td>double</td>
<td>increase</td>
<td>skyrocket</td>
<td>decline</td>
<td>increase</td>
<td></td>
</tr>
<tr>
<td>drop</td>
<td>jump</td>
<td>slide</td>
<td>decrease</td>
<td>rise</td>
<td></td>
</tr>
</tbody>
</table>
The “Change position on a scale” Frame

<table>
<thead>
<tr>
<th>Core Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTRIBUTE</td>
</tr>
<tr>
<td>DIFFERENCE</td>
</tr>
<tr>
<td>FINAL_STATE</td>
</tr>
<tr>
<td>FINAL_VALUE</td>
</tr>
<tr>
<td>INITIAL_STATE</td>
</tr>
<tr>
<td>INITIAL_VALUE</td>
</tr>
<tr>
<td>ITEM</td>
</tr>
<tr>
<td>VALUE_RANGE</td>
</tr>
</tbody>
</table>

Some Non-Core Roles

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DURATION</td>
</tr>
<tr>
<td>SPEED</td>
</tr>
<tr>
<td>GROUP</td>
</tr>
</tbody>
</table>

Relation between frames

Inherits from:
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:
Relation between frames

“cause change position on a scale”
Is Causative of: Change_position_on_a_scale
Adds an agent Role

[AGENT They] raised [ITEM the price of their soda] [DIFFERENCE by 2%].

- add.v, crank.v, curtail.v, cut.n, cut.v, decrease.v, development.n, diminish.v, double.v, drop.v, enhance.v, growth.n, increase.v, knock down.v, lower.v, move.v, promote.v, push.n, push.v, raise.v, reduce.v, reduction.n, slash.v, step up.v, swell.v

Relations between frames

![Diagram](image)

Figure from Das et al 2010
Schematic of Frame Semantics

Austria, once expected to waltz smoothly into the European Union, is elbowing its partners treading on toes and pogo-dancing in a most un-Viennese manner.

Figure from Das et al (2014)

FrameNet Complexity

But there still aren’t enough ringers to ring more than six of the eight bells.

From Das et al. 2010
FrameNet and PropBank representations

In that time more than 1.2 million jobs have been created and the official jobless rate has been pushed below 17% from 21%.

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 EVALUEE 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

A simple modern algorithm

```plaintext
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

```
DT  NNP  NNP  NNP  
The  San  Francisco  Examiner
NP-SBJ = ARG0
VBD = TARGET
 issued
VP

PP-TMP = ARGM-TMP
NP

DT  JJ  NN  IN  
a  special  edition  around
NP

NN  NN  
oon  noon
NP-TMP
```

```
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
NP↑S↓VP↓VBD
Frequent path features

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Path</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.2%</td>
<td>VB↑VP↓PP</td>
<td>PP argument/adjunct</td>
</tr>
<tr>
<td>11.8</td>
<td>VB↑VP↑S↓NP</td>
<td>subject</td>
</tr>
<tr>
<td>10.1</td>
<td>VB↑VP↓NP</td>
<td>object</td>
</tr>
<tr>
<td>7.9</td>
<td>VB↑VP↑VP↑S↓NP</td>
<td>subject (embedded VP)</td>
</tr>
<tr>
<td>4.1</td>
<td>VB↑VP↓ADVP</td>
<td>adverbial adjunct</td>
</tr>
<tr>
<td>3.0</td>
<td>NN↑NP↑NP↑PP</td>
<td>prepositional complement of noun</td>
</tr>
<tr>
<td>1.7</td>
<td>VB↑VP↓PRT</td>
<td>adverbial particle</td>
</tr>
<tr>
<td>1.6</td>
<td>VB↑VP↑VP↑VP↑S↓NP</td>
<td>subject (embedded VP)</td>
</tr>
<tr>
<td>14.2</td>
<td>Other</td>
<td>no matching parse constituent</td>
</tr>
<tr>
<td>31.4</td>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

From Palmer, Gildea, Xue 2010

Final feature vector

- For “The San Francisco Examiner”,
- Arg0, [issued, NP, Examiner, NNP, active, before, VP→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 possible 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

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)

More complications: FrameNet

We need an extra step to find the frame

```python
function SEMANTICROLELABEL(words) returns labeled tree
    function SEMANTICROLELABEL(words) returns labeled tree
        parse ← PARSE(words)
        for each predicate in parse do
            Predicatevector ← ExtractFrameFeatures(predicate, parse)
            Frame ← ClassifyFrame(predicate, predicatevector)
        for each node in parse do
            Frame ← EXTRACTFEATURES(node, predicate, parse)
            CLASSIFYNODE(node, featurevector, parse, Frame)
```

59

60
Features for Frame Identification

Das et al (2014)

the POS of the parent of the head word of \( t_i \)
the set of syntactic dependencies of the head word \( t_i \)
if the head word of \( t_i \) is a verb, then the set of dependency labels of its children
the dependency label on the edge connecting the head of \( t_i \) and its parent
the sequence of words in the prototype, \( w_t \)
the lemmatized sequence of words in the prototype
the lemmatized sequence of words in the prototype and their part-of-speech tags \( \pi_t \)
WordNet relation \( \hat{p} \) holds between \( \ell \) and \( t_i \)
WordNet relation \( \hat{p} \) holds between \( \ell \) and \( t_i \), and the prototype is \( \ell \)
WordNet relation \( \hat{p} \) holds between \( \ell \) and \( t_i \), the POS tag sequence of \( \ell \) is \( \pi_t \), and the POS tag sequence of \( t_i \) is \( \pi_t \)

Not just English

```
IP
  \( \text{Arg0} \)
    \( \text{NP-SBJ} \)
      警方
      police
  \( \text{ArgM-TMP} \)
    正在
    now
  \( \text{ArgM-MNR} \)
    详细
    thoroughly
  VP
    \( \text{ADV-P-TMP} \)
      调查
      investigate
    \( \text{ADV-P-MNR} \)
      事故
      accident
    Rel
      \( \text{VV} \)
        原因
        cause
  \( \text{Arg1} \)
    \( \text{NP-OBJ} \)
      NN
      NN
```

“The police are thoroughly investigating the cause of the accident.”
Not just verbs: NomBank

Meyers et al. 2004

Additional Issues for nouns

• Features:
  • Nominalization lexicon (employment → employ)
  • Morphological stem
    • Healthcare, Medicate → care
• Different positions
  • Most arguments of nominal predicates occur inside the NP
  • Others are introduced by support verbs
  • Especially light verbs “X made an argument”, “Y took a nap”
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
Selectional Restrictions

Introduction
Selectional Restrictions

Consider the two interpretations of:

I want to eat someplace nearby.

a) sensible:
   Eat is intransitive and “someplace nearby” is a location adjunct

b) Speaker is Godzilla
   Eat is transitive and “someplace nearby” is a direct object

How do we know speaker didn’t mean b) ?
   Because the THEME of eating tends to be something edible

Selectional restrictions are associated with senses

• The restaurant serves green-lipped mussels.
  • THEME is some kind of food
• Which airlines serve Denver?
  • THEME is an appropriate location
Selectional restrictions vary in specificity

I often ask the musicians to imagine a tennis game.
To diagonalize a matrix is to find its eigenvalues.
Radon is an odorless gas that can’t be detected by human senses.

Representing selectional restrictions

Instead of representing “eat” as:
\[ \exists e, x, y \, Eating(e) \land Agent(e, x) \land Theme(e, y) \]

Just add:
\[ \exists e, x, y \, Eating(e) \land Agent(e, x) \land Theme(e, y) \land EdibleThing(y) \]

And “eat a hamburger” becomes
\[ \exists e, x, y \, Eating(e) \land Eater(e, x) \land Theme(e, y) \land EdibleThing(y) \land Hamburger(y) \]

But this assumes we have a large knowledge base of facts about edible things and hamburgers and whatnot.
Let’s use WordNet synsets to specify selectional restrictions

- The **THEME** of eat must be WordNet synset \{food, nutrient\}
  “any substance that can be metabolized by an animal to give energy and build tissue”

- Similarly
  
  **THEME of imagine**: synset \{entity\}
  **THEME of lift**: synset \{physical entity\}
  **THEME of diagonalize**: synset \{matrix\}

- This allows
  
  *imagine a hamburger* and *lift a hamburger*,

- Correctly rules out
  
  *diagonalize a hamburger.*

---

**Selectional Restrictions**

**Selectional Preferences**
Selectional Preferences

• In early implementations, selectional restrictions were strict constraints (Katz and Fodor 1963)
  • Eat [+FOOD]
• But it was quickly realized selectional constraints are really preferences (Wilks 1975)
  • But it fell apart in 1931, perhaps because people realized you can’t eat gold for lunch if you’re hungry.
  • In his two championship trials, Mr. Kulkarni ate glass on an empty stomach, accompanied only by water and tea.

Selectional Association (Resnik 1993)

• Selectional preference strength: amount of information that a predicate tells us about the semantic class of its arguments.
  • eat tells us a lot about the semantic class of its direct objects
  • be doesn’t tell us much
• The selectional preference strength
  • difference in information between two distributions:
    P(c) the distribution of expected semantic classes for any direct object
    P(c|v) the distribution of expected semantic classes for this verb
  • The greater the difference, the more the verb is constraining its object
Selectional preference strength

- Relative entropy, or the Kullback-Leibler divergence is the difference between two distributions
  \[ D(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \]

- Selectional preference: How much information (in bits) the verb expresses about the semantic class of its argument
  \[ S_R(v) = D(P(c|v)||P(c)) = \sum_c P(c|v) \log \frac{P(c|v)}{P(c)} \]

- Selectional Association of a verb with a class: The relative contribution of the class to the general preference of the verb
  \[ A_R(v,c) = \frac{1}{S_R(v)} P(c|v) \log \frac{P(c|v)}{P(c)} \]

Computing Selectional Association

- A probabilistic measure of the strength of association between a predicate and a semantic class of its argument
  - Parse a corpus
  - Count all the times each predicate appears with each argument word
  - Assume each word is a partial observation of all the WordNet concepts associated with that word
  - Some high and low associations:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Direct Object Semantic Class</th>
<th>Direct Object Semantic Class</th>
<th>Assoc</th>
<th>Assoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>read</td>
<td>WRITING</td>
<td>ACTIVITY</td>
<td>6.80</td>
<td>-.20</td>
</tr>
<tr>
<td>write</td>
<td>WRITING</td>
<td>COMMERCE</td>
<td>7.26</td>
<td>0</td>
</tr>
<tr>
<td>see</td>
<td>ENTITY</td>
<td>METHOD</td>
<td>5.79</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
Results from similar models

Ó Séaghdha and Korhonen (2012)

Instead of using classes, a simpler model of selectional association

- Model just the association of predicate $v$ with a noun $n$
  
  (one noun, as opposed to the whole semantic class in WordNet)

  - Parse a huge corpus
  - Count how often a noun $n$ occurs in relation $r$ with verb $v$:
    
    $\log \text{count}(n, v, r)$

  - Or the probability:
    
    $$P(n|v, r) = \begin{cases} 
    \frac{C(n, v, r)}{C(v, r)} & \text{if } C(n, v, r) > 0 \\
    0 & \text{otherwise}
    \end{cases}$$
Evaluation from Bergsma, Lin, Goebel

<table>
<thead>
<tr>
<th>Verb</th>
<th>Plaus./Implaus.</th>
</tr>
</thead>
<tbody>
<tr>
<td>see</td>
<td>friend/method</td>
</tr>
<tr>
<td>read</td>
<td>article/fashion</td>
</tr>
<tr>
<td>find</td>
<td>label/fever</td>
</tr>
<tr>
<td>hear</td>
<td>story/issue</td>
</tr>
<tr>
<td>write</td>
<td>letter/market</td>
</tr>
<tr>
<td>urge</td>
<td>daughter/contrast</td>
</tr>
<tr>
<td>warn</td>
<td>driver/engine</td>
</tr>
<tr>
<td>judge</td>
<td>contest/climate</td>
</tr>
<tr>
<td>teach</td>
<td>language/distance</td>
</tr>
<tr>
<td>show</td>
<td>sample/travel</td>
</tr>
<tr>
<td>expect</td>
<td>visit/mouth</td>
</tr>
<tr>
<td>answer</td>
<td>request/tragedy</td>
</tr>
<tr>
<td>recognize</td>
<td>author/pocket</td>
</tr>
<tr>
<td>repeat</td>
<td>comment/journal</td>
</tr>
<tr>
<td>understand</td>
<td>concept/session</td>
</tr>
<tr>
<td>remember</td>
<td>reply/smoke</td>
</tr>
</tbody>
</table>

Selectional Restrictions

Conclusion
Summary: Selectional Restrictions

- Two classes of models of the semantic type constraint that a predicate places on its argument:
  - Represent the constraint between predicate and WordNet class
  - Represent the constraint between predicate and a word
- One fun recent use case: detecting metonomy (type coercion)
  - Coherent with selectional restrictions: \(^{83}\) Pustejovsky et al (2010)
    - The spokesman denied the statement (PROPOSITION).
    - The child threw the stone (PHYSICAL OBJECT)
  - Coercion:
    - The president denied the attack (EVENT → PROPOSITION).
    - The White House (LOCATION → HUMAN) denied the statement.