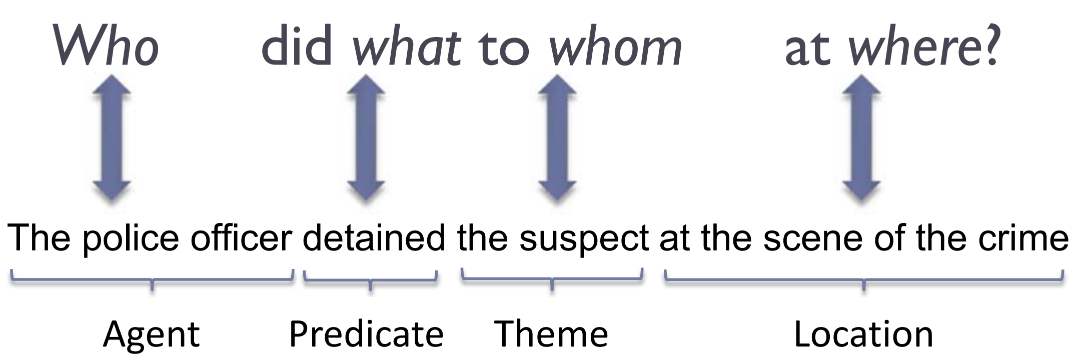
# Semantic Role Labeling

Chapter 20

## Semantic Role Labeling



## 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...

3

## A Shallow Semantic Representation: Semantic Roles

Predicates (bought, sold, purchase) represent an **event** and **semantic roles** express the abstract role that arguments of a predicate can take in the event

More specific ⟷ More general

**buyer** **agent** **proto-agent**

4

## Getting to semantic roles

What roles are involved in a breaking event?

First order logic event representation for Sasha broke the window:

5

## Getting to semantic roles

First order logic event representation:

Sasha broke the window

Pat opened the door

$$\exists e,x,y\ Breaking(e) \wedge Breaker(e,Sasha) \wedge BrokenThing(e,y) \wedge Window(y)$$

$$\exists e,x,y\ Opening(e) \wedge Opener(e,Pat) \wedge OpenedThing(e,y) \wedge Door(y)$$

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

6

## 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

7

## 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*

8

## Thematic roles

- A typical set:

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

9

## Thematic grid, case frame

Example usages of "break"

- 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

10

## Thematic grid, case frame

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

11

## Thematic grid, case frame

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**
Break:
AGENT, THEME, INSTRUMENT.

Some realizations:

AGENT/Subject, THEME/Object
AGENT/Subject, THEME/Object, INSTRUMENT/$PP_{with}$
INSTRUMENT/Subject, THEME/Object
THEME/Subject

What type of parsing?

12

## Diathesis alternations (or verb alternation)

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

*Doris gave Cary the book.*
AGENT GOAL THEME

*Break:* AGENT, INSTRUMENT, or THEME as subject

*Give:* THEME and GOAL in either order

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

13

## 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.

14

## 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**

15

## PropBank

- Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106
- http://verbs.colorado.edu/~mpalmer/projects/ace.html

16

## 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

17

## 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

18

*(Arg2-Arg5 are not really that consistent, causes a problem for labeling)*

## PropBank Frame Files

**http://verbs.colorado.edu/propbank/framesets-english-aliases/agree.html**

**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%].

19

## 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:

Big Fruit Co. increased the price of bananas.
The price of bananas was increased again by Big Fruit Co.
The price of bananas increased 5%

20

## 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. ] increased [$_{\text{Arg1}}$ the price of bananas].
[$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co. ]
[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

21

## Modifiers or adjuncts of the predicate: Arg-M

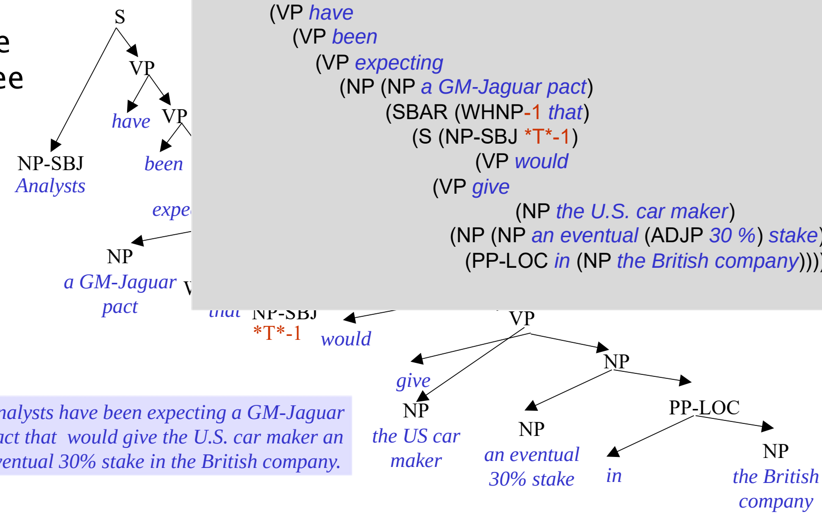
| | | |
|---|---|---|
| **ArgM-TMP** | when? | yesterday evening, now |
| **LOC** | where? | at the museum, in San Francisco |
| **DIR** | where to/from? | down, to Bangkok |
| **MNR** | how? | clearly, with much enthusiasm |
| **PRP/CAU** | why? | because ... , in response to the ruling |
| **REC**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

22

# PropBanking a Sentence

Martha Palmer 2013

A sample parse tree



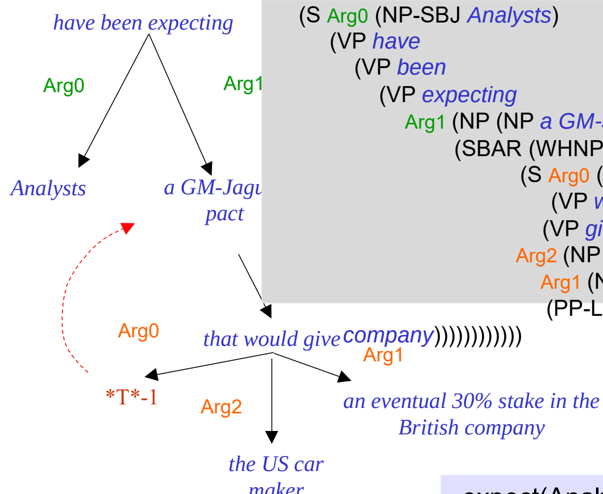

(S (NP-SBJ *Analysts*)
(VP *have*
(VP *been*
(VP *expecting*
(NP (NP *a GM-Jaguar pact*)
(SBAR (WHNP-1 *that*)
(S (NP-SBJ *T*-1)
(VP *would*
(VP *give*
(NP *the U.S. car maker*)
(NP (NP *an eventual* (ADJP *30 %*) *stake*)
(PP-LOC *in* (NP *the British company*))))))))))))

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

23

# The same parse tree PropBanked

Martha Palmer 2013




(S Arg0 (NP-SBJ *Analysts*)
(VP *have*
(VP *been*
(VP *expecting*
Arg1 (NP (NP *a GM-Jaguar pact*)
(SBAR (WHNP-1 *that*)
(S Arg0 (NP-SBJ *T*-1)
(VP *would*
(VP *give*
Arg2 (NP *the U.S. car maker*)
Arg1 (NP (NP *an eventual* (ADJP *30 %*) *stake*)
(PP-LOC *in* (NP *the British company*))))))))))))

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

24

## 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)

| *Language* | *Final Count* |
|---|---|
| English | 10,615* |
| Chinese | 24, 642 |
| Arabic | 7,015 |

25

From Martha Palmer 2013 Tutorial

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

[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].
[$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].
There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

26

## FrameNet

- Baker et al. 1998, Fillmore et al. 2003, Fillmore and Baker 2009, Ruppenhofer et al. 2006
- 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
- **https://framenet.icsi.berkeley.edu/fndrupal/**

27

## 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)

[$_{\text{ITEM}}$ Oil] *rose* [$_{\text{ATTRIBUTE}}$ in price] [$_{\text{DIFFERENCE}}$ by 2%].

[$_{\text{ITEM}}$ It] has *increased* [$_{\text{FINAL\_STATE}}$ to having them 1 day a month].

[$_{\text{ITEM}}$ Microsoft shares] *fell* [$_{\text{FINAL\_VALUE}}$ to 7 5/8].

[$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

a steady *increase* [$_{\text{INITIAL\_VALUE}}$ from 9.5] [$_{\text{FINAL\_VALUE}}$ to 14.3] [$_{\text{ITEM}}$ in dividends]

a [$_{\text{DIFFERENCE}}$ 5%] [$_{\text{ITEM}}$ dividend] *increase*...

28

## The "Change position on a scale" Frame

| | | | | | |
|---|---|---|---|---|---|
| **VERBS:** | dwindle | move | soar | escalation | shift |
| advance | edge | mushroom | swell | explosion | tumble |
| climb | explode | plummet | swing | fall | |
| decline | fall | reach | triple | fluctuation | **ADVERBS:** |
| decrease | fluctuate | rise | tumble | gain | increasingly |
| diminish | gain | rocket | | growth | |
| dip | grow | shift | **NOUNS:** | hike | |
| double | increase | skyrocket | decline | increase | |
| drop | jump | slide | decrease | rise | |

29

## The "Change position on a scale" Frame

| | |
|---|---|
| **Core Roles** | |
| ATTRIBUTE | The ATTRIBUTE is a scalar property that the ITEM possesses. |
| DIFFERENCE | The distance by which an ITEM changes its position on the scale. |
| FINAL_STATE | A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication. |
| FINAL_VALUE | The position on the scale where the ITEM ends up. |
| INITIAL_STATE | A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication. |
| INITIAL_VALUE | The initial position on the scale from which the ITEM moves away. |
| ITEM | The entity that has a position on the scale. |
| VALUE_RANGE | A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate. |
| **Some Non-Core Roles** | |
| DURATION | The length of time over which the change takes place. |
| SPEED | The rate of change of the VALUE. |
| GROUP | The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way. |

30

## 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:

31

## Relation between frames

"cause change position on a scale"

Is Causative of: Change_position_on_a_scale

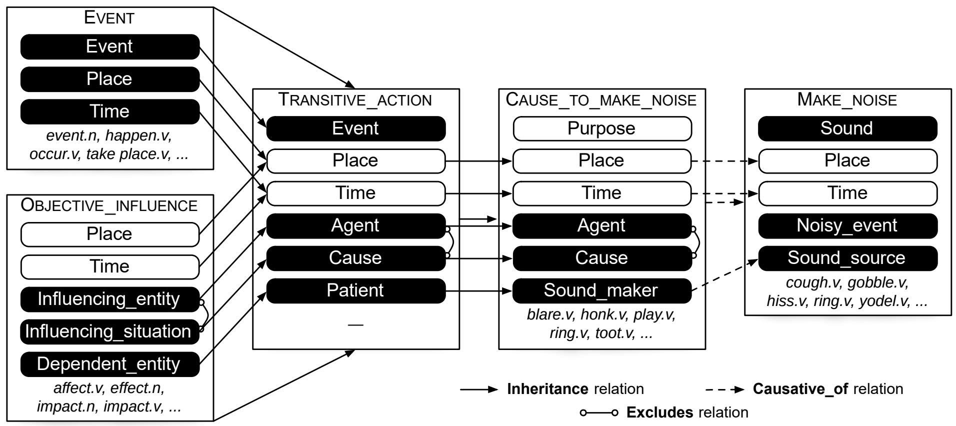
Adds an agent Role

[$_{\text{AGENT}}$ They] *raised* [$_{\text{ITEM}}$ the price of their soda] [$_{\text{DIFFERENCE}}$ 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*

32

# Relations between frames




33

Figure from Das et al 2010

# Schematic of Frame Semantics

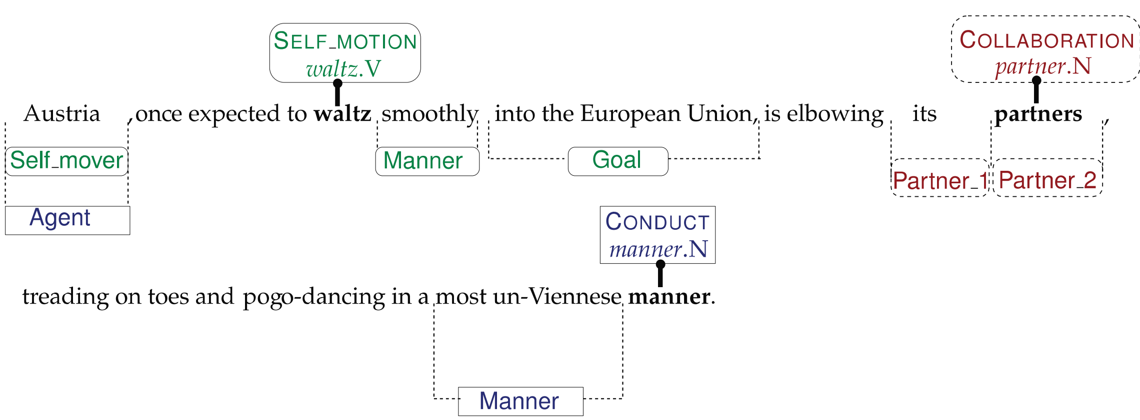



Figure from Das et al (2014)

34

## Homework 3

| | |
|---|---|
| Minimum | 0.00 |
| Maximum | 100.00 |
| | |
| Average | 86.175 |
| Median | 93.00 |

35

## Observations from Rav

- Using other algorithms such as NB didn't improve the baseline with statistical significance.
- Addressing the data imbalance directly, e.g. regrouping the labels or oversampling, did find an improvement ... sometimes.
- Pre-processing and manipulating how many words to consider: there is some number of features (between 2000 and 8000) that maximizes the accuracy, and that normalizing the text too much hurts the performance.
  - No one handled unknown words though
- General summary: the best performance is achieved through proper and thoughtful feature extraction and management.

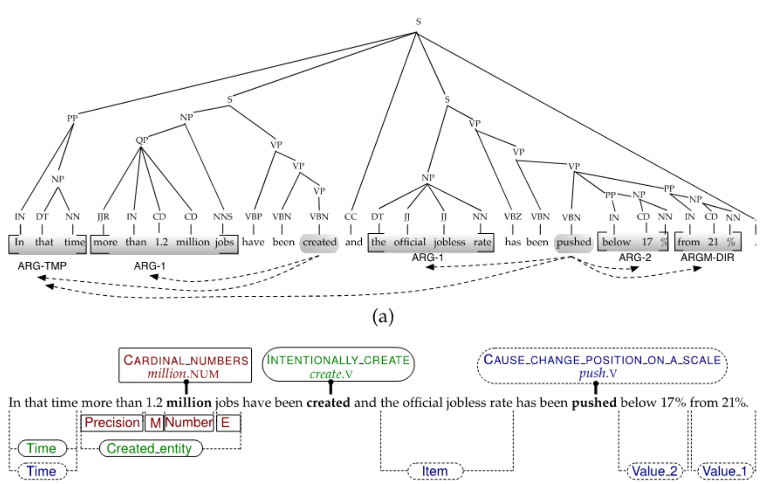
36

## Review

- Semantic roles
- Human-created resources
  - PropBank
  - FrameNet

37

## FrameNet and PropBank representations




38

## Semantic role labeling (SRL) algorithms

- 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 |

39

## 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)

40

## Why Semantic Role Labeling

- A useful shallow semantic representation
- Improves downstream NLP tasks like
  - question answering
  - machine translation

41

## A simple modern algorithm

**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*)

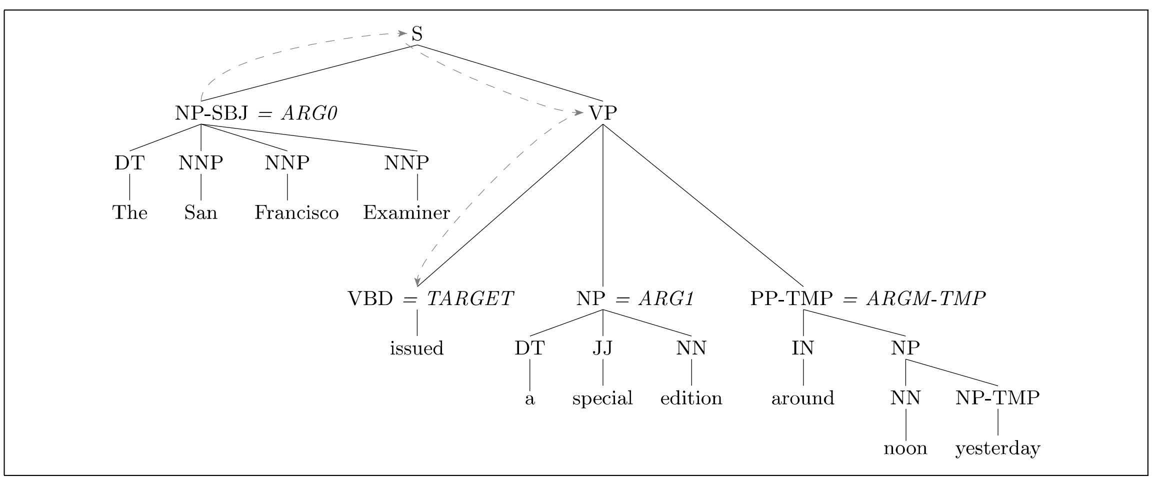
42

## How do we decide what is a predicate

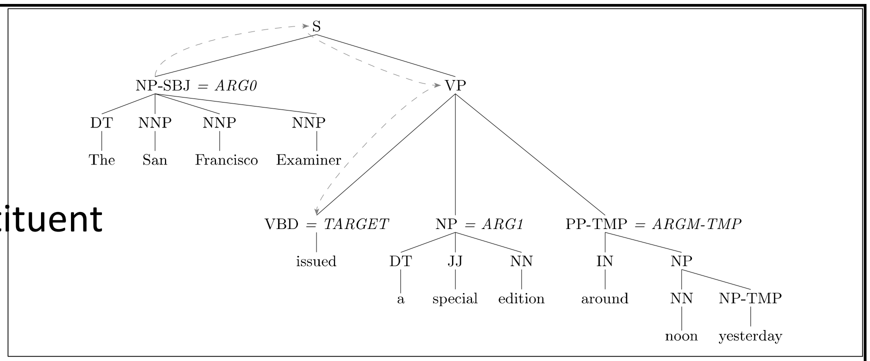
- If we're just doing PropBank verbs
  - Choose all verbs
- If we're doing FrameNet (verbs, nouns, adjectives)
  - Choose every word that was labeled as a target in training data



## Semantic Role Labeling



## Features: 1st constituent




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

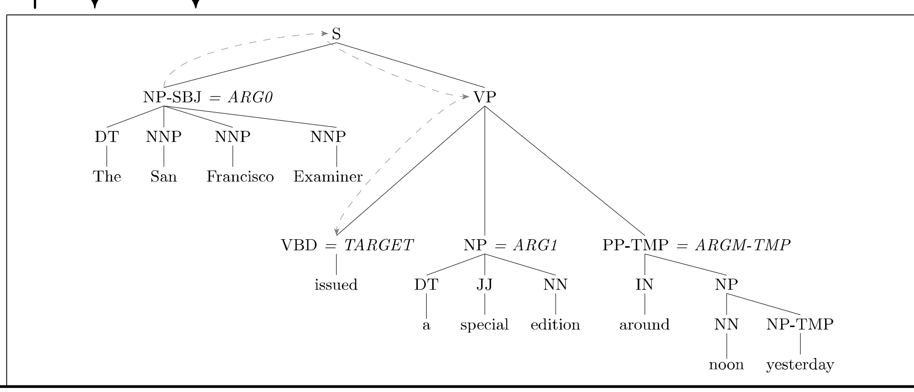
Linear position,clause re: predicate
- before

45

## Path Features

**Path** in the parse tree from the constituent to the predicate

$$NP\uparrow S\downarrow VP\downarrow VBD$$




46

## Frequent path features

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

47

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

48

## 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

49

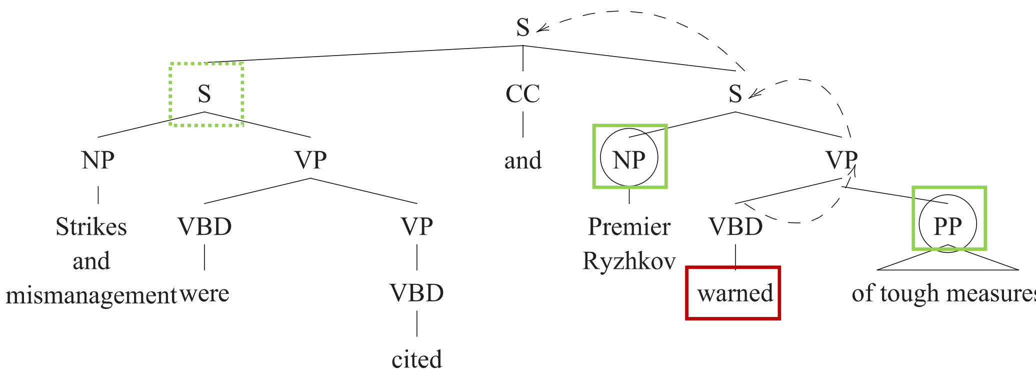
## Why add Pruning and Identification steps?

- Algorithm is looking at one predicate at a time
- Very few of the nodes in the tree could be possible 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.

50

## Pruning heuristics – Xue and Palmer (2004)

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




51

## 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

52

## 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)

53

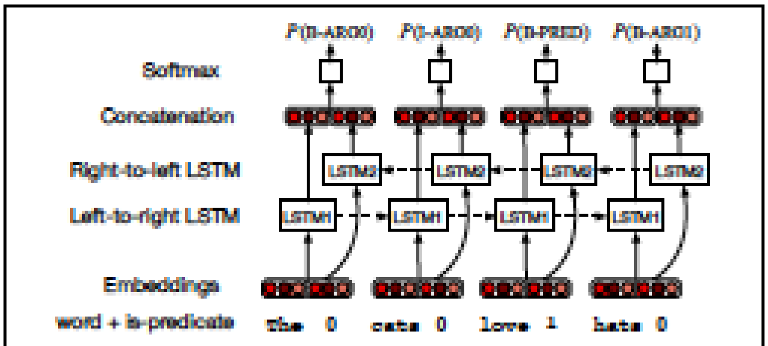
## Neural Approaches too

- Typically same models as used for other "tagging" tasks (e.g., POS, NER)
- Instead of parsing first, uses and end-to-end (map straight from words) approach

54

## Neural Approaches too

- Typically model used for other "tagging" tasks (e.g., POS, NER)




[redacted] A bi-LSTM approach to semantic role labeling. Most actual networks are much deeper than shown in this figure; 3 to 4 bi-LSTM layers (6 to 8 total LSTMs) are common. The input is a concatenation of an embedding for the input word and an embedding of a binary variable which is 1 for the predicate to 0 for all other words. After He et al. (2017).

55

## More complications: FrameNet

We need an extra step to find the frame

**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**
    *featurevector* ← EXTRACTFEATURES(*node*, *predicate*, *parse*)
    CLASSIFYNODE(*node*, *featurevector*, *parse* , Frame)

56

## 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[21] of $t_i$
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, $w_\ell$
the lemmatized sequence of words in the prototype
the lemmatized sequence of words in the prototype and their part-of-speech tags $\pi_\ell$
WordNet relation[22] $\rho$ holds between $\ell$ and $t_i$
WordNet relation[22] $\rho$ holds between $\ell$ and $t_i$, and the prototype is $\ell$
WordNet relation[22] $\rho$ holds between $\ell$ and $t_i$, the POS tag sequence of $\ell$ is $\pi_\ell$, and the POS tag sequence of $t_i$ is $\pi_t$

57

## Evaluation

- Each argument label must be assigned to the exactly correct word sequence or parse constituent
- Recall/Precision/F
- Common to use shared task datasets from CoNLL (Computational Natural Language Learning)

58

## SRL Summary

- 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

59

## Selectional Restrictions

Consider :

I want to eat someplace nearby.

60

## 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*

61

## 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

62

## 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.

63

## Representing selectional restrictions

Instead of representing "eat" as:

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y)$$

Just add:

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y) \wedge EdibleThing(y)$$

And "eat a hamburger" becomes

$$\exists e,x,y\ Eating(e) \wedge Eater(e,x) \wedge Theme(e,y) \wedge EdibleThing(y) \wedge Hamburger(y)$$

But this assumes we have a large knowledge base of facts about edible things and hamburgers and whatnot.

64

## 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*.

65

## Selectional Preferences

- In early implementations, selectional restrictions were strict constraints
  - Eat [+FOOD]
- But it was quickly realized selectional constraints are really **preferences**
  - 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.

66

## 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

67

## 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

$$\begin{aligned} S_R(v) &= D(P(c|v)||P(c)) \\ &= \sum_c P(c|v) \log \frac{P(c|v)}{P(c)} \end{aligned}$$

- Selectional Association of a verb with a class: The relative contribution of the class to the general preference of the verb

68

$$A_R(v,c) = \frac{1}{S_R(v)} P(c|v) \log \frac{P(c|v)}{P(c)}$$

## 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:

| Verb | Direct Object Semantic Class | Assoc | Direct Object Semantic Class | Assoc |
|---|---|---|---|---|
| read | WRITING | 6.80 | ACTIVITY | -.20 |
| write | WRITING | 7.26 | COMMERCE | 0 |
| see | ENTITY | 5.79 | METHOD | -0.01 |

69

## 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:

$$P(n|v,r) = \begin{cases} \frac{C(n,v,r)}{C(v,r)} & \text{if } C(n,v,r) > 0 \\ 0 & \text{otherwise} \end{cases}$$

70

## Evaluation from Bergsma, Lin, Goebel

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

71

## Evaluation

- Pseudowords
  - Choose between real argument and created confounders
- Compare to human preferences

72

## Primitive Decomposition of Predicates

- Semantic roles define the roles that arguments play for a predicate in a decompositional way based on finite lists
- Can do something similar to define predicate meaning itself!

73

## 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

74