CS447: Natural Language Processing
http://courses.engr.illinois.edu/cs447

Lecture 24: Semantic Role Labeling and Verb Semantics

Julia Hockenmaier
juliahmr@illinois.edu
3324 Siebel Center

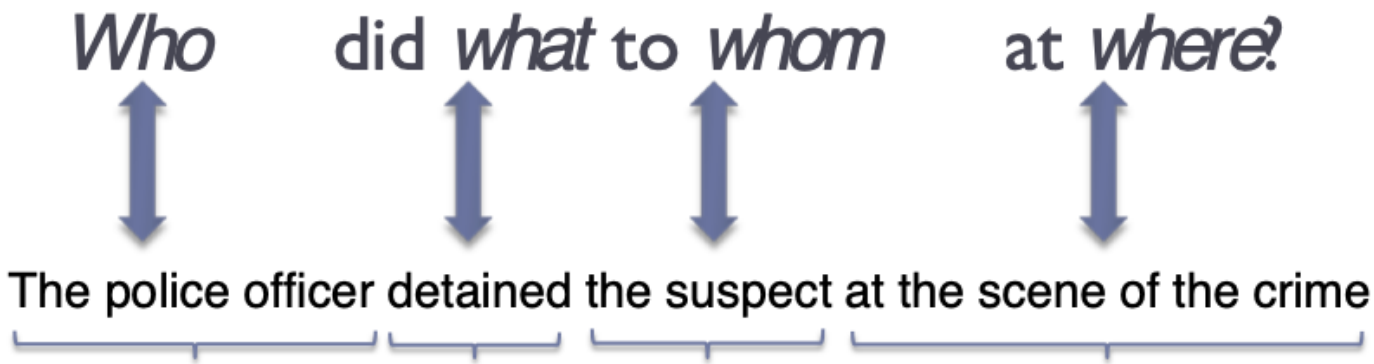
Where we're at

Last lecture: Lexical semantics, mostly for nouns
—Sense relations (e.g. hypernym/hyponym relations)
—Word Sense Disambiguation

Today: Verb semantics
— Argument structure
— Verb classes
— Semantic Role Labeling
(Chapter 20 in textbook)

The importance of predicate-argument structure

Predicate-argument structure

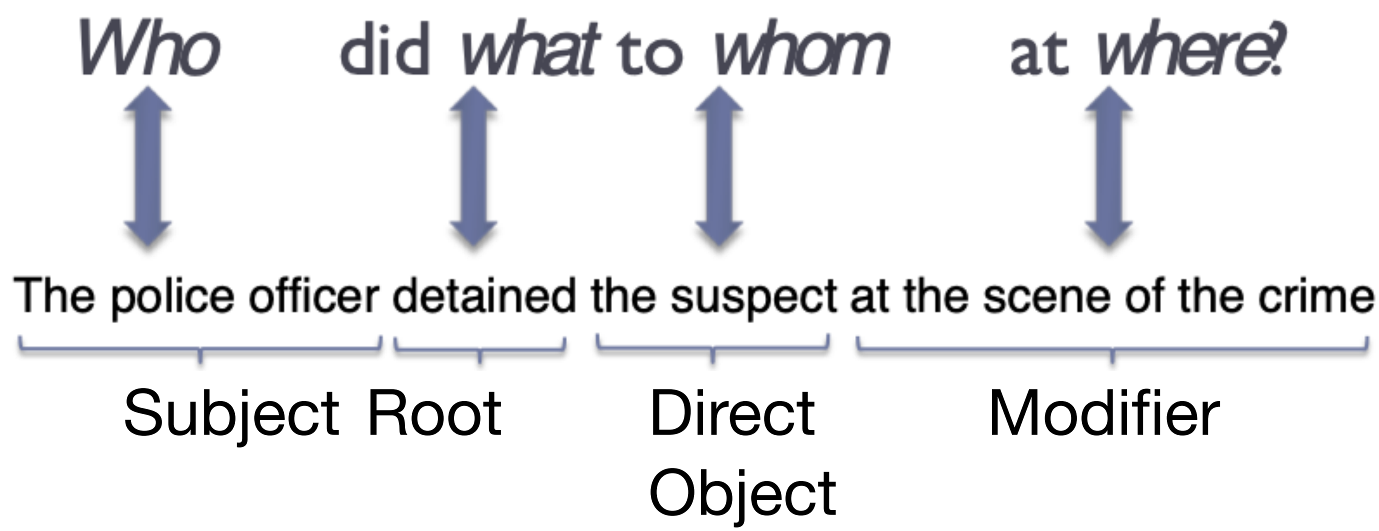


Understanding a sentence = knowing who did what (to whom, when, where, why…)

Verbs corresponds to predicates (what was done)

Their **arguments** (and modifiers) identify who did it, to whom, where, when, why, etc.)

Syntactic Parsing



Syntactic Parsing (e.g. dependency parsing) identifies grammatical roles (subject, (direct) object, etc.)

What do verbs mean?

Verbs describe **events** or **states** (‘eventualities’):

Tom broke the window with a rock.

The window broke.

The window was broken by Tom/by a rock.

We could translate verbs to (logical) predicates.

But: a naive translation

(e.g. subject = first argument, object = second argument, etc.)

does not capture that the similarities in meaning

```
break(Tom, window, rock)
break(window)
break(window, Tom)
break(window, rock)
```

There are many different ways to describe the same event

Grammatical roles ≠ Semantic roles

Tom **broke the window** with a rock.

The window broke.

The window was broken by Tom/by a rock.

Related verbs/nouns can describe the same event:

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...

Can we map these sentences to the same representation?

How do we represent verb semantics?

Neo-Davidsonian Event Representations

Predicate logic with explicit event variables *e*, and explicit predicates for each role:

Sasha broke the window

$$\exists e \exists y \text{Breaking}(e) \wedge \text{Broken}(e, y) \wedge \text{Breaker}(e, \text{Sasha}) \wedge \text{Window}(y)$$

Pat opened the door

$$\exists e \exists y \text{Opening}(e) \wedge \text{OpenedThing}(e, y) \wedge \text{Opener}(e, \text{Pat}) \wedge \text{Door}(y)$$

Explicit event variables make it easy to add adjuncts (Time(e, t)), and to express relations between events.

Here, break and open have verb-specific “deep” roles (**Breaker** and **Opener)**

Hard to reason about/with these roles, generalize

Towards Thematic roles

Breaker and **Opener** have something in common!

— Volitional actors

— Often animate

— Direct causal responsibility for their events

Thematic roles are a way to capture this semantic commonality between *Breakers* and *Eaters*.
They are both AGENTS.
The *BrokenThing* and *OpenedThing*, are THEMES.

prototypically inanimate objects affected in some way by the action

Semantic/Thematic roles

Verbs describe events or states ('eventualities'):

Tom broke the window with a rock.

The window broke.

The window was broken by Tom/by a rock.

Thematic roles refer to participants of these events:

Agent (who performed the action): Tom

Patient (who was the action performed on): window

Tool/Instrument (what was used to perform the action): rock

Semantic/thematic roles (agent, patient) are different from grammatical roles (subject or object).

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*

The inventory of thematic roles

To create systems that can identify thematic roles automatically, we need to create labeled training data.

This means we need to define an inventory of thematic roles

It is difficult to give a formal definition of thematic roles that generalizes across all verbs.

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*.

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 of this frame/grid:

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

A frame/grid identifies the set of roles associated with a particular event type.
These roles can be expressed ('realized') by different grammatical roles

Diathesis Alternations

Active/passive alternation:

Tom **broke** the window with a rock. (active voice)

The window **was broken** by Tom/by a rock. (passive voice)

Causative alternation:

Tom **broke** the window. ('causative'; active voice)

The window **broke**. ('anticausative'/'inchoative'; active voice)

Dative alternation

Tom **gave** the gift to Mary.

Tom **gave** Mary the gift.

Locative alternation:

Jessica **loaded** boxes into the wagon.

Jessica **loaded** the wagon with boxes.

Verb classes ("Levin classes")

Verbs with similar meanings undergo the same syntactic alternations, and have the same set of thematic roles (Beth Levin, 1993)

VerbNet (verbs.colorado.edu; Kipper et al., 2008)
A large database of verbs, their thematic roles and their alternations

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

Fewer roles: generalized semantic roles, defined as prototypes (Dowty 1991)

PROTO-AGENT
PROTO-PATIENT

More roles:
Define roles specific to a group of predicates

PropBank: generic roles with frame-specific interpretation
FrameNet: frame-specific roles

Datasets for Semantic Role Labeling

PropBank and FrameNet

Proposition Bank (PropBank):
Very coarse argument roles (arg0, arg1,…),
used for all verbs (but interpretation depends on the specific verb)

Arg0 = proto-agent
Arg1 = proto-patient
Arg2...: specific to each verb
ArgM-TMP/LOC/...: temporal/locative/... modifiers

FrameNet:

Verbs fall into classes that define different kinds of frames
(`change-position-on-a-scale` frame: rise, increase,...).
Each frame has its own set of “frame elements” (thematic roles)

PropBank

agree.01 Arg0: Agreer Arg1: Proposition
Arg2: Other entity agreeing

[$_{\text{Arg0}}$ The group] agreed [$_{\text{Arg1}}$ it wouldn't make an offer]

[$_{\text{Arg0}}$ John] agrees with [$_{\text{Arg2}}$ Mary]

fall.01 Arg1: patient/thing falling Arg2: extent/amount fallen
Arg3: start point Arg4: end point

[$_{\text{Arg1}}$ Sales] fell [$_{\text{Arg4}}$ to $251 million]

[$_{\text{Arg1}}$ Junk bonds] fell [$_{\text{Arg2}}$ by 5%]

Semantic role labeling: Recover the semantic roles of verbs (nowadays typically PropBank-style)

Machine learning; trained on PropBank

Syntactic parses provide useful information

PropBank

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106

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

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

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

PropBank Frame Files

agree.01
Arg0: Agr er
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].
Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ 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: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to \$25 million] [$_{\text{Arg3}}$ from \$27 million].
Ex2: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ 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 allows 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%].

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

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%].

1) [ITEM It] has *increased* [FINAL_STATE to having them 1 day a month].

2) [ITEM Microsoft shares] *fell* [FINAL_VALUE to 7 5/8].

3) [ITEM Colon cancer incidence] *fell* [DIFFERENCE by 50%] [GROUP among men].

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

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

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	

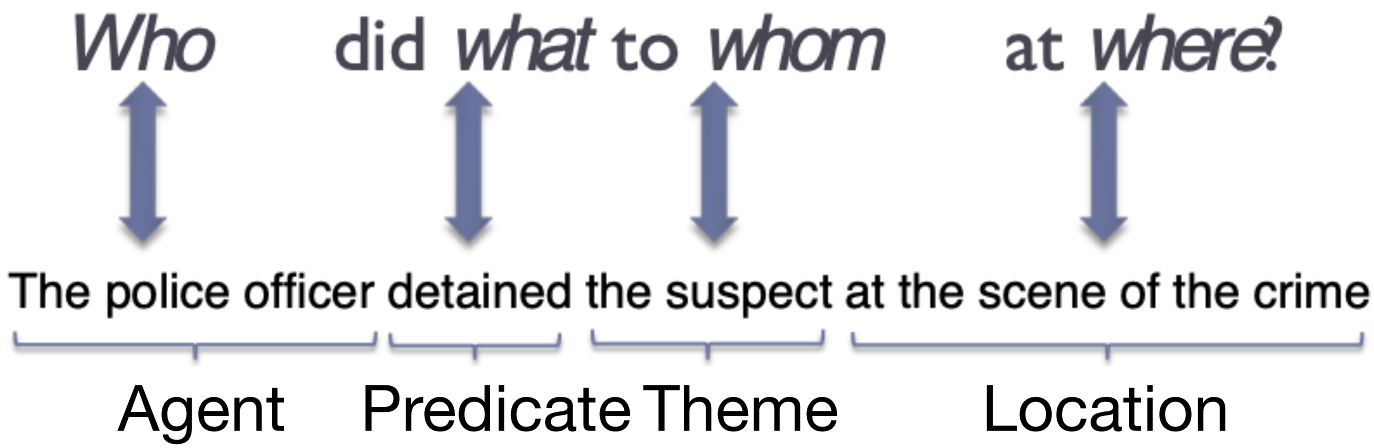
Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM’s state after the change in the ATTRIBUTE’s value as an independent predication.
FINAL_VALUE	The 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.

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:

Semantic Role Labeling algorithms

Semantic Role Labeling



Identify
— all predicates in a sentence
— the arguments of each predicate
and their semantic roles

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

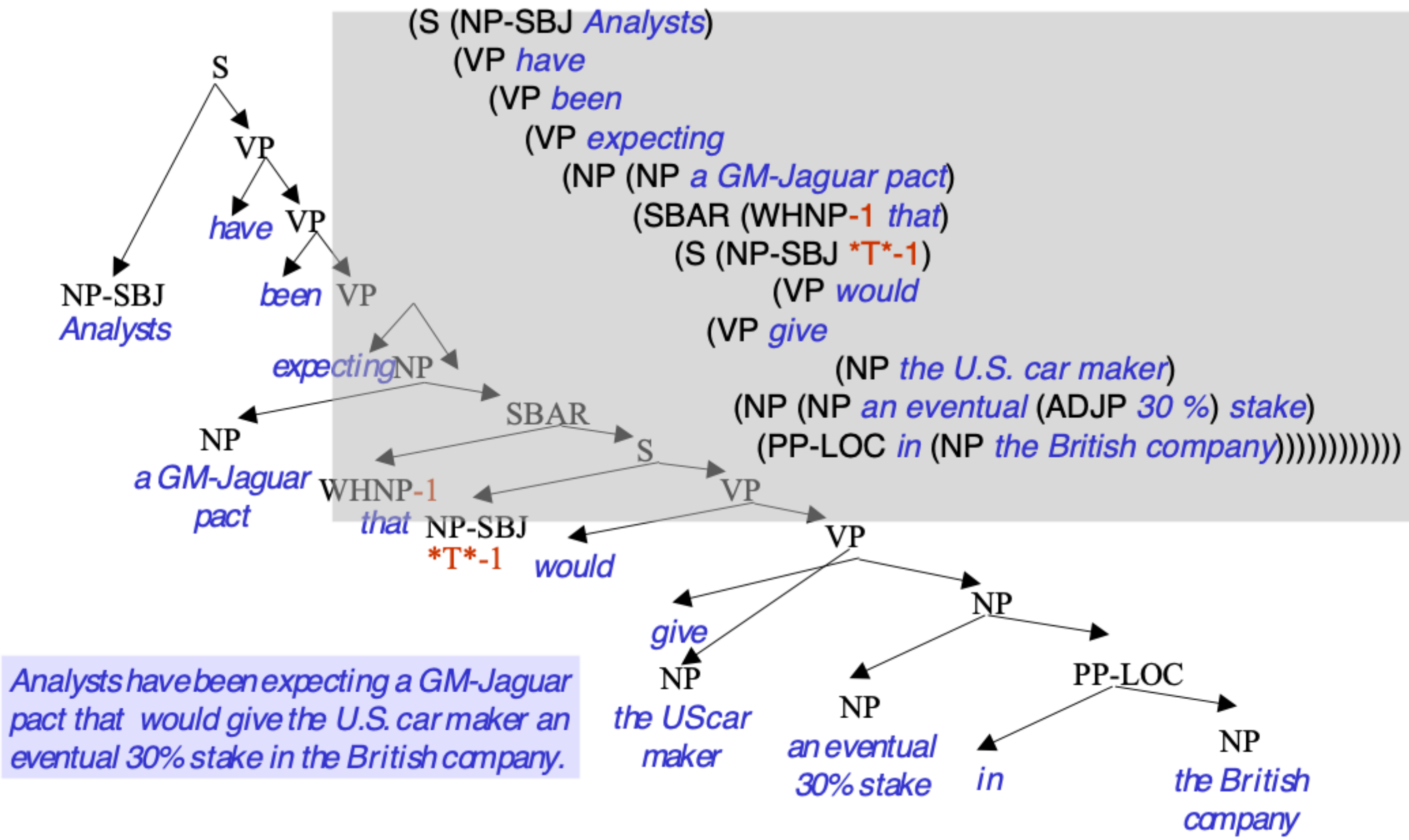
Simmons 1973, Marcus 1980:

- parser followed by hand-written rules for each verb
- dictionaries with verb-specific case frames (Levin 1977)

PropBanking a Sentence

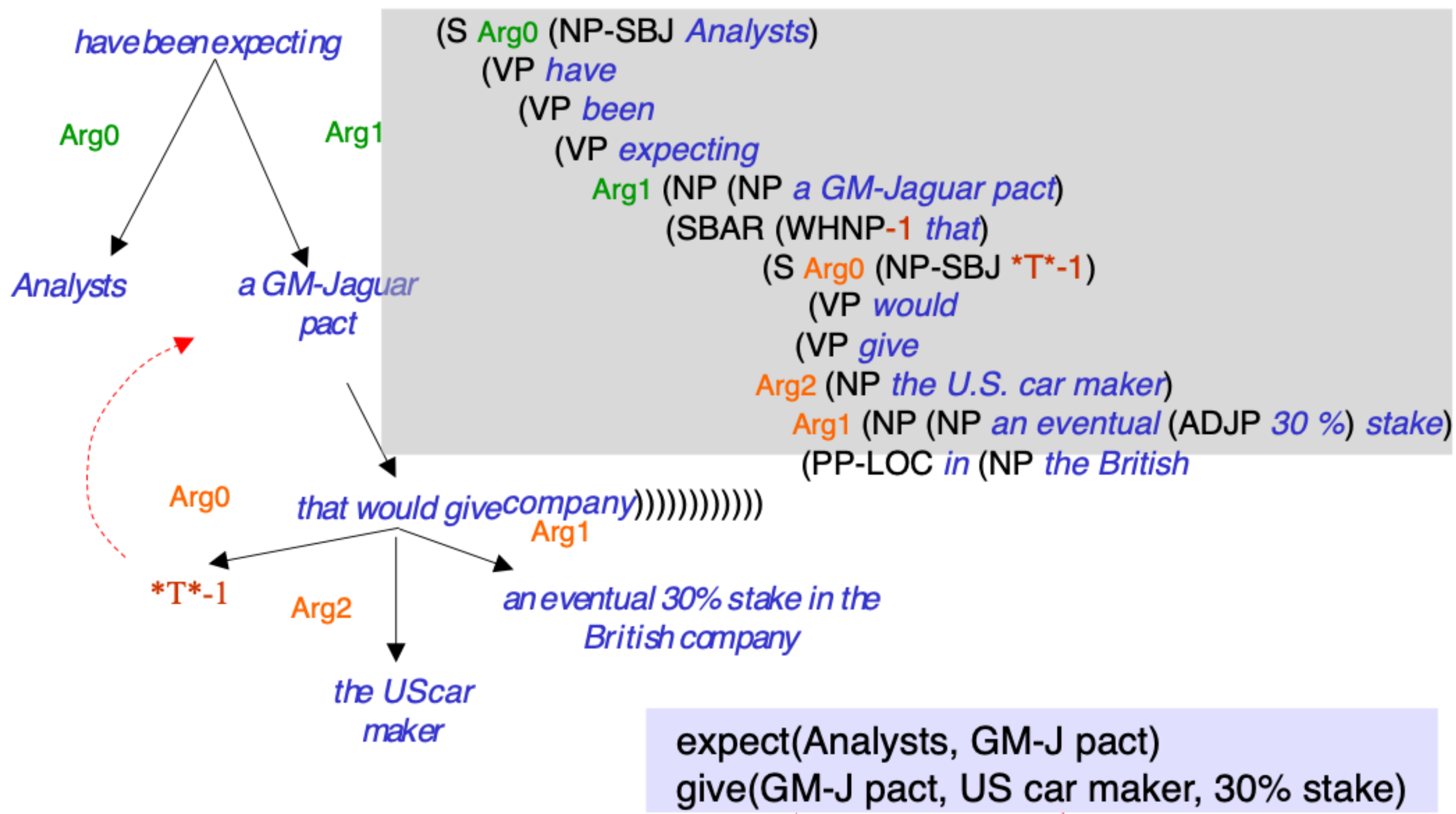
A sample parse tree

Martha Palmer 2013



The same parse tree PropBanked

Martha Palmer 2013



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

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$_{ARG1}$]"

Slide from Palmer 2013

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)
```

How do we decide what is a predicate

If we're just doing PropBank verbs

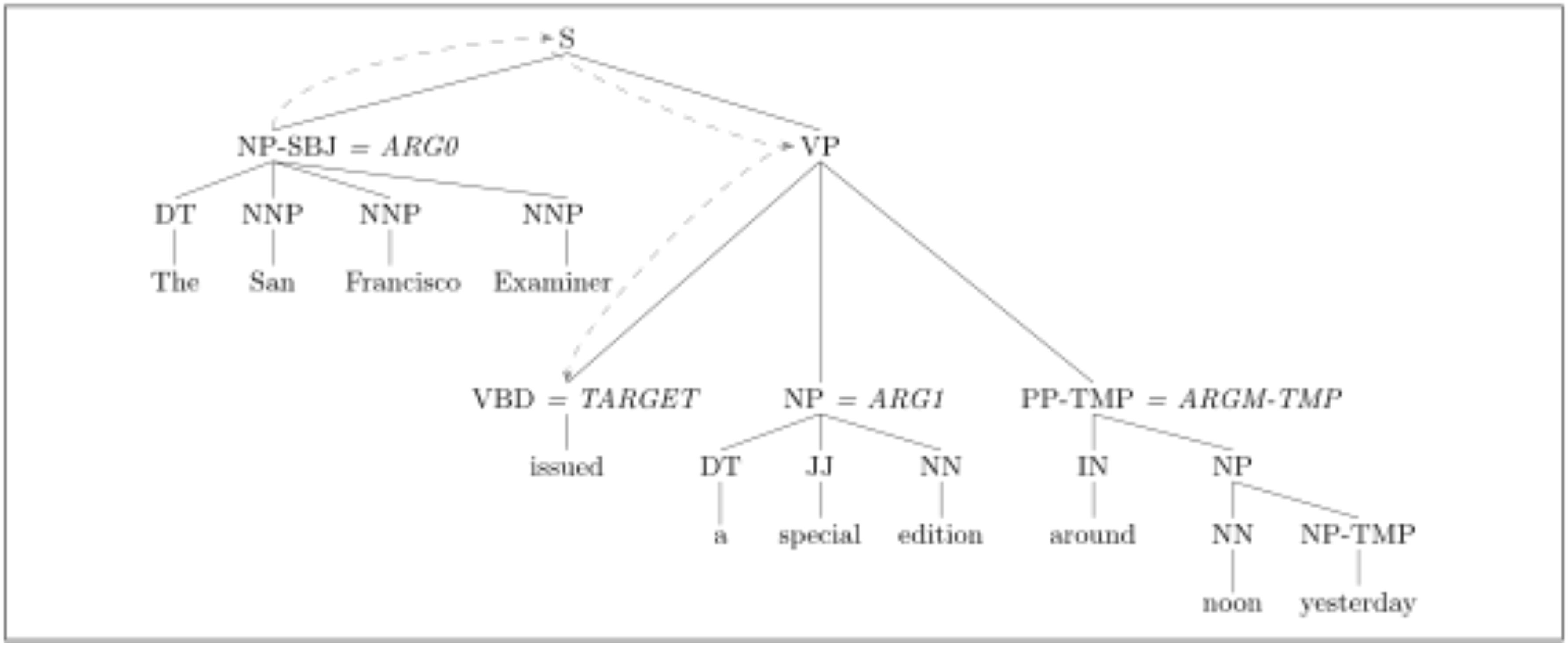
Choose all verbs

Possibly removing light verbs (from a list)

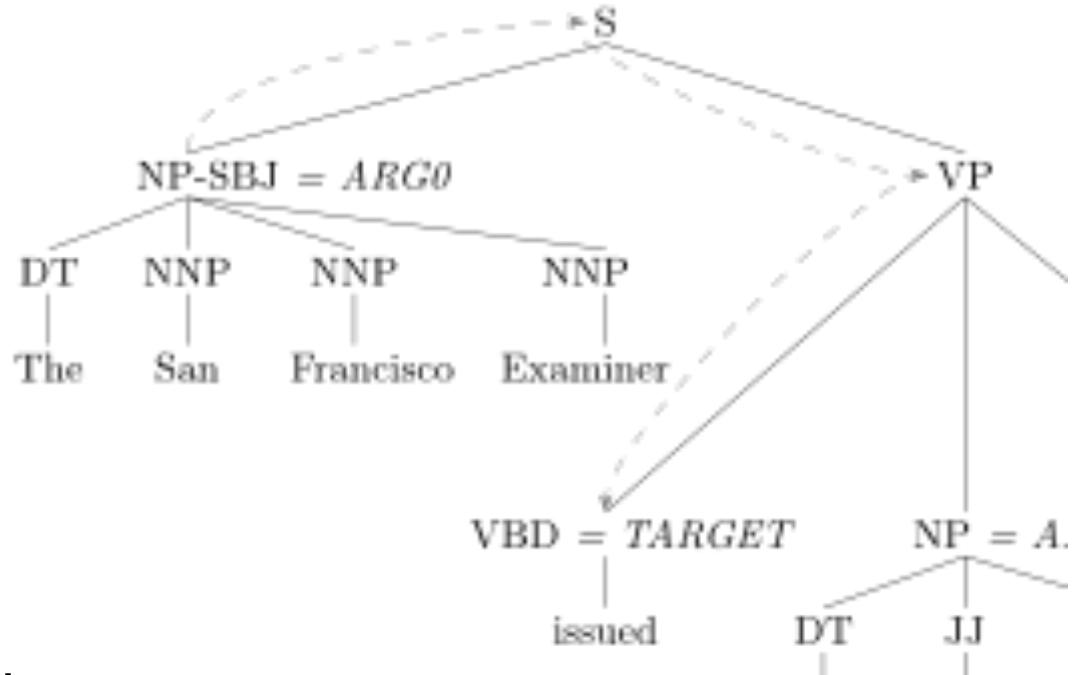
If we're doing FrameNet (verbs, nouns, adjectives)

Choose every word that was labeled as a target in training data

Semantic Role Labeling



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: issued: VBD->VP->S<-NP<-NNP examiner

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	

From Palmer, Gildea, Xue 2010

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

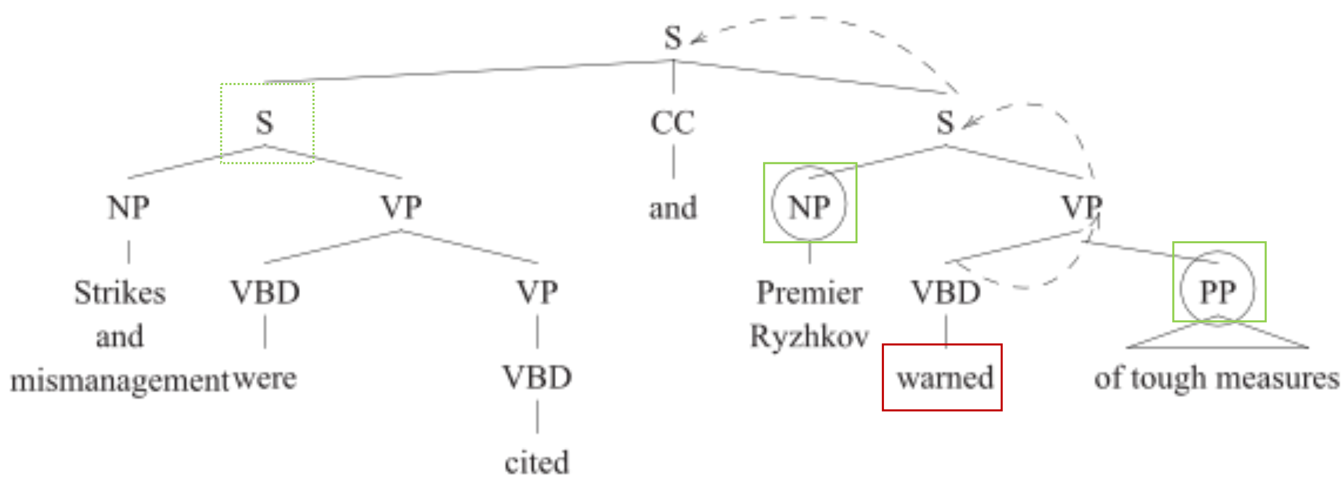
- positive samples (constituents that are arguments of predicate)
- negative samples (constituents that are not arguments of predicate)

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.

Pruning heuristics – Xue and Palmer

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

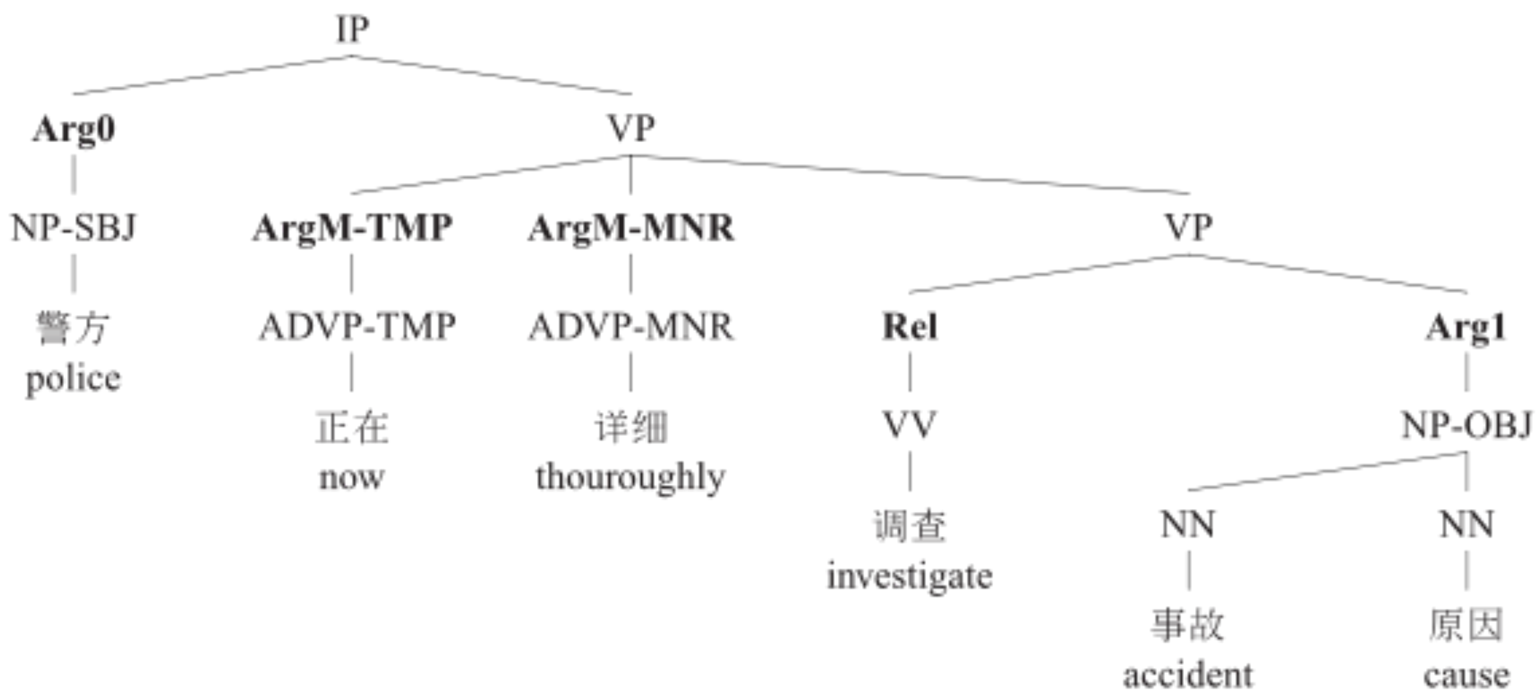
  Thus should increase the probability of another being ARG1

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)

Not just English



"The police are thoroughly investigating the cause of the accident."

Not just verbs: NomBank

Meyers et al. 2004

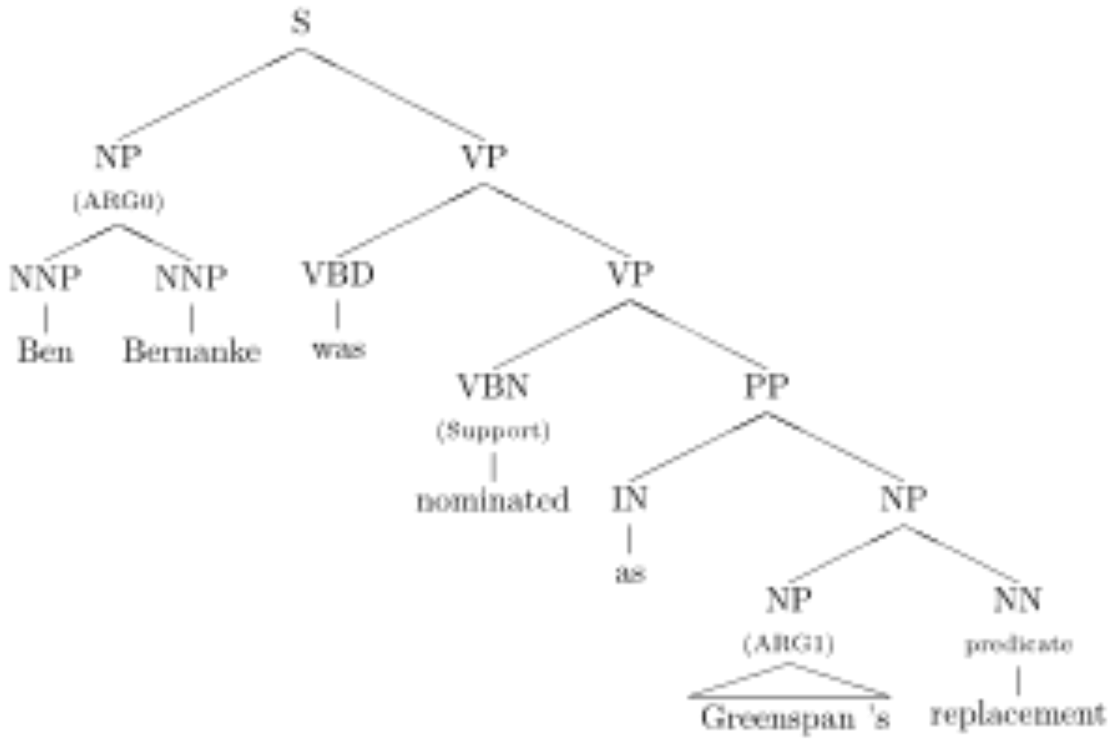


Figure from Jiang and Ng 2006

Additional Issues for nouns

Features:

- Nominalization lexicon (employment => employ)
- Morphological stem
  - Healthcare, Medicare => 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

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