Semantic Role Labeling

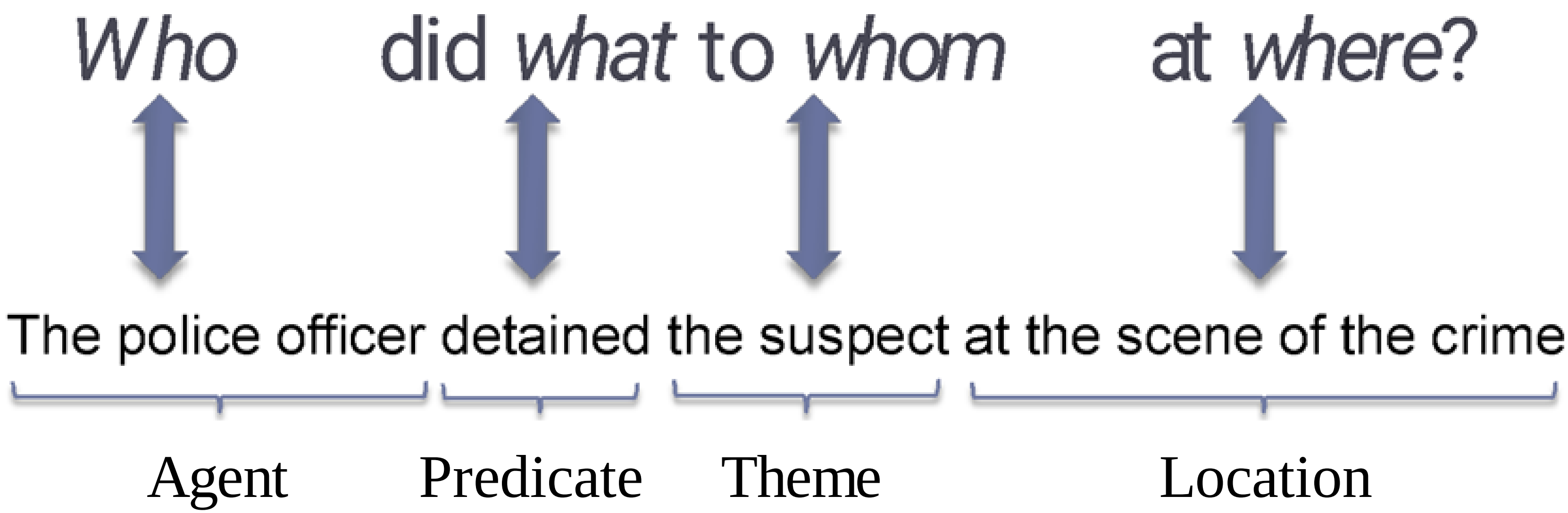
CS-585

Natural Language Processing

Sonjia Waxmonsky

Based on slides from Kai Shu and Derrick Higgins

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...

Predicate-Argument Structure and Semantic Roles

- **Predicates** bear the central meaning of a situation expressed by a text.
  - Usually verbs or verbal nouns (bought, sold, purchase)
- **Arguments:** phrases that fill meaning slots of a situation expressed by a predicate, define its essential details
  - Answer questions like “who?”, “did what?”, “to whom?”
  - Usually nouns or noun phrases (stock, XYZ Corp.)
- **Semantic roles** express the abstract role that **arguments** of a predicate can take in the event
  - More specific: "Buyer"
  - Less specific: "Agent"

SEMANTIC ROLES

Semantic roles: First order logic

First order logic: Quantifies variables (Some, All)

"Neo-Davidsonian" **event semantics**:

- Event represented as variable ***e***
- Event arguments linked via relations

Sasha broke the window

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

Pat opened the door

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

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

Deep roles specific to each event (breaking, opening)

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 both have the **role** of: **AGENT**.
- *BrokenThing* and *OpenedThing* have role of **THEME**.
  - Prototypically inanimate objects affected in some way by the action

Thematic roles

A typical set of thematic roles:

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*.

- Many sets exists
- No set is universally agreed-upon

Beyond Thematic Roles

Problems with Thematic roles for NLP:

- Hard to create and define a **standard** set of roles
- Often roles need to be fragmented to be defined.

Alternatives for NLP:

1. **Fewer roles**: generalized semantic roles, defined as prototypes (Dowty 1991)
    - PROTO-AGENT
    - PROTO-PATIENT
2. **More roles**: Define roles specific to a <u>group of predicates</u>

THE PROPOSITION BANK (PROPBANK)

PropBank

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

- System of role representation
- Defines arguments for individual **verb senses** (i.e. individual word sense for individual verbs)
- Corpus of sentences annotated by Role
- English PropBank: Annotated on **Penn Treebank**

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: 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

ArgM: Adjunct meanings (e.g. time); not necessary to complete meaning of predicate

(Arg2-Arg4 are not really that consistent, causes a problem for labeling)

PropBank Frame Files

agree.01

Arg0: Agreer
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%].

← No Arg0?
No proto-agent

Advantage of a PropBank 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%].

FRAMENET

FrameNet

Goal: Capturing descriptions of the **same event** by **different nouns and verbs**

The price of bananas increased 5%.
The price of bananas rose 5%.
There has been a 5% increase in the price of bananas.

Which word or phrase in each sentence indicates prices are higher? What is its part of speech in each sentence?

FrameNet

Goal: Capturing descriptions of the same event by **different nouns and 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].

FrameNet and Frames

- **Frame**:
  - Descriptions of situations or events
  - Evoked by lexical units – usually but not always verbs
  - Include **frame elements** (similar concept to roles)

- FrameNet
  - Lexicon of ~1000 frames
  - Corpus of ~200k annotated sentences that use those frames
  - Groups verbs into frames
  - Links semantically related roles across frames

Example Frame:
"Change position on a scale"

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*...

Frame Elements: Example

Frame **elements** for frame “**Change position on a scale**":

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.

https://web.stanford.edu/~jurafsky/slp3/24.pdf

Frame Target Words: Example

Target words for frame "**Change position on a scale**":

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	

https://web.stanford.edu/~jurafsky/slp3/24.pdf

Relationships between Frames

FrameNet encodes relationships between frames to allow sharing of frame elements (i.e. frame roles)

Example frame relationship: **Causation**

- Frame "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%].

Relationships between Frames

FrameNet encodes relationships between frames to allow sharing of frame elements (i.e. frame roles)

Example frame relationship: **Inheritance**

- Child Frames "**Commerce-Sell**" and "**Lending**" both inherit from parent frame "**Giving**"

FrameNet vs ProbBank

- Role in FrameNet are specific to a **frame**
  - Multiple target words can evoke same frame, usually but not always verbs
  - Roles/elements are shared across frames
- Roles in PropBank are specific to a **verb sense**
  - ARG0 and ARG1 are consistent across verbs
  - ARG2, ARG3, … are (somewhat) verb-specific

SEMANTIC ROLE LABELING

Semantic role labeling (SRL)

- Task: Recover the predicate-argument structure of a sentence
- Find the semantic roles of each argument of each **predicate** in a sentence
- Can apply FrameNet representation...

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

- ... or PropBank representation:

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

Shallow Semantics

- Semantic Role Labeling: Roles are filled with **tokens from the text:**
  – Identify **predicates**
  – Specify **span** of text (tokens or sequences of tokens) that fills each role
- Why "shallow"?
  – Roles are <u>not</u> filled with symbolic expressions
  – <u>Not</u> applying first-order logic concepts
  – Intermediate between parses and full semantics

Why Semantic Role Labeling

- Provides a useful shallow semantic representation
- Improves NLP tasks like:
  - Question answering
  - Information extraction
  - Textual entailment
    - "Disney's purchase of ABC..."→Disney bought ABC
  - Machine translation

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**

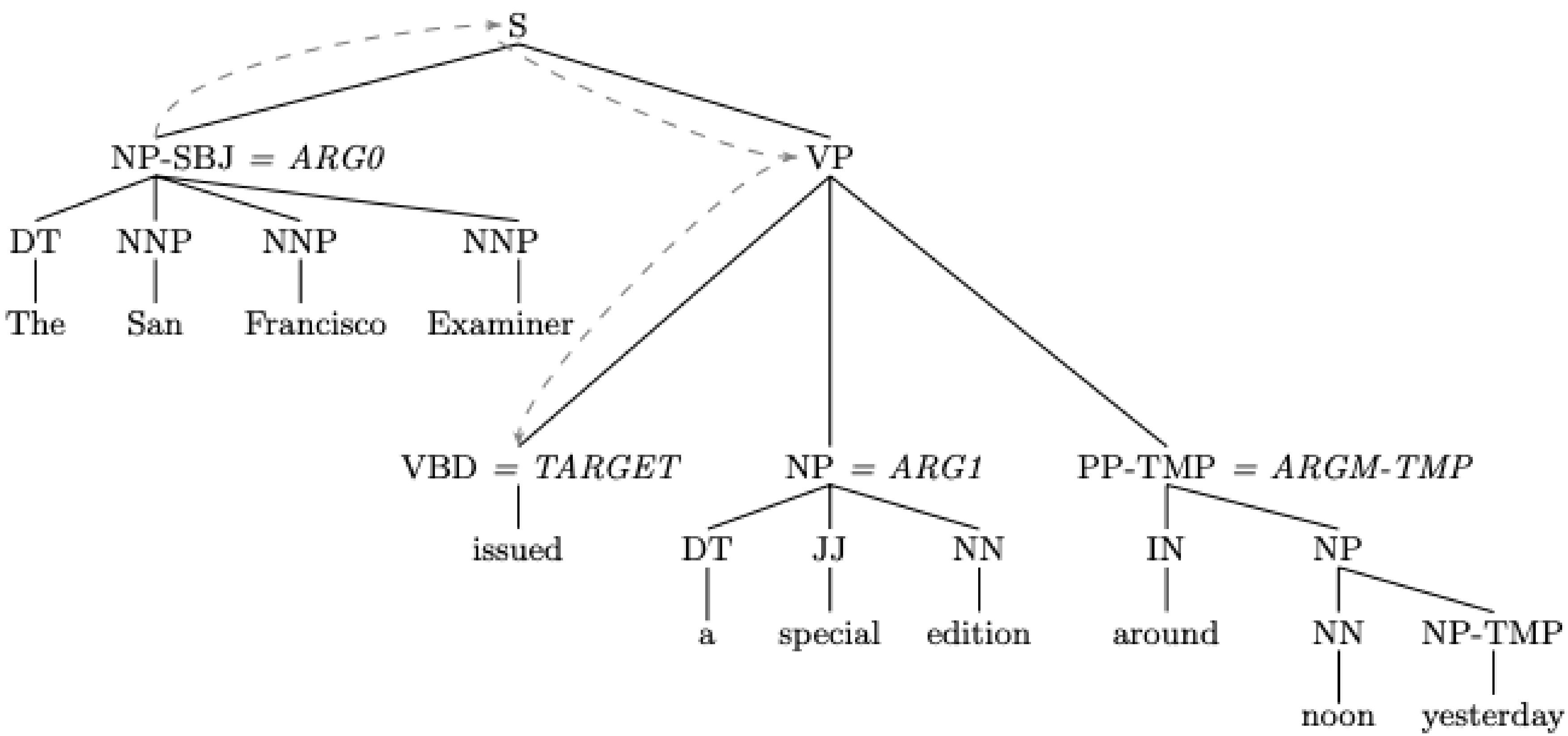
SRL as Classification

- SRL can be modeled as a classification of phrasal constituents
- Steps:
  - Parse sentence
  - Finding predicates in the sentence
  - For each predicate, classify each parse tree constituent

function SEMANTICROLELABEL(*words*) **returns** labeled tree

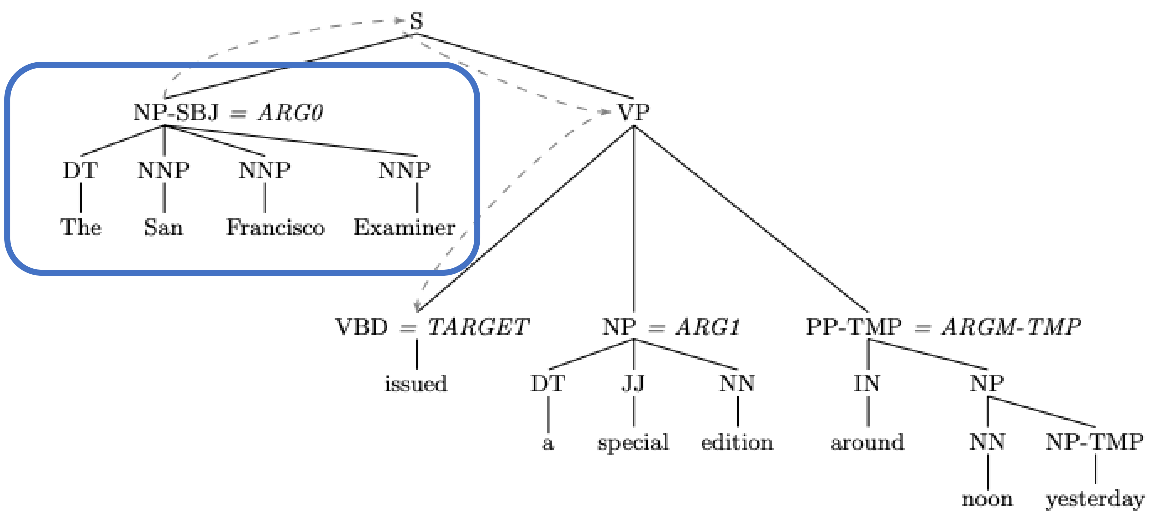
parse ← PARSE(*words*)
for each *predicate* **in** *parse* **do**
    for each *node* **in** *parse* **do**
        featurevector ← EXTRACTFEATURES(*node*, *predicate*, *parse*)
        CLASSIFYNODE(*node*, *featurevector*, *parse*)

SRL as Classification



https://web.stanford.edu/~jurafsky/slp3/24.pdf

Features for SRL



Features for:
"The San Francisco Examiner"

Headword of constituent

Examiner

Headword POS

NNP

Voice of the clause

Active

Subcategorization of pred

VP -> VBD NP PP

Named Entity type of constituent

ORGANIZATION

First and last words of constituent

The, Examiner

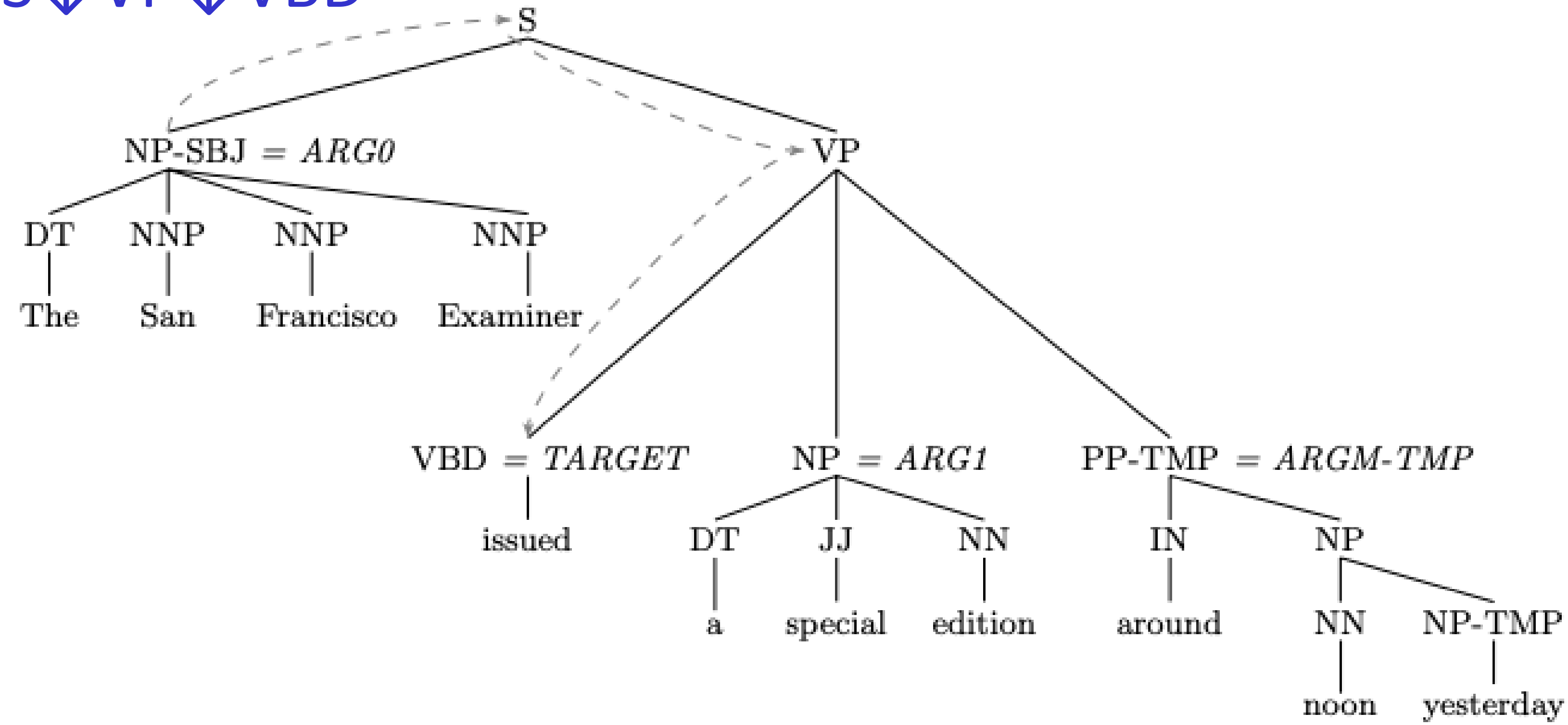
Linear position,clause re: predicate

before

Path Features

Path in the parse tree from the constituent to the predicate

NP↑S↓VP↓VBD



Frequent path features

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

Features for SRL

For “The San Francisco Examiner”,

- Label: Arg0
- Features: [issued, NP, Examiner, NNP, active, before, VP->VBD NP PP, ORG, The, Examiner, NP↑S↓VP↓VBD]

Other features could be used as well

- sets of n-grams inside the constituent
- other path features
  - the upward or downward halves
  - whether particular nodes occur in the path

Constrained Optimization for SRL

- SRL has sentence level constraints, e.g.:
  - For each verb, there can only be one argument of each type
  - Arguments cannot be from overlapping spans
- Can be imposed with integer linear programming
- Example: Each non-null role type r can occur only once over all constituents in a sentence:

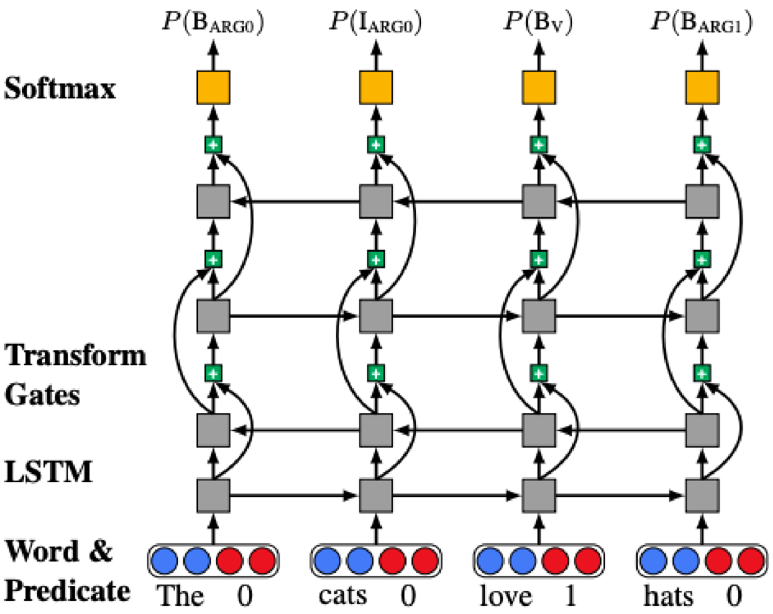
$$\forall r \neq \varnothing, \quad \sum_{(i,j) \in \tau} z_{i,j,r} \leq 1.$$

[E=NLP Formula 13.7]

Deep Learning for SRL

Deep Semantic Role Labeling: What Works and What's Next [He et al, 2017]

- Treats SRL as B-I-O sequential tagging task, similar to previous NN-based approaches
- Applies Bi-LSTM model based on word and binary predicate embedding
- Finds minimal benefit from decoding constraints on SRL structure



https://aclanthology.org/P17-1044.pdf

Deep Learning for SRL

Simple BERT Models for Relation Extraction and Semantic Role Labeling [Shi and Lin, ~~2003~~ 2019]

Typo in year corrected after lecture

- Applies pre-trained transformer-based models (BERT)
- Also applies B-I-O SRL tagging (e.g B-ARG0)
- Shows that SRL is possible without lexical or syntactic features
- Promising for low-resource languages where parsers not available
- Also does not apply decoding constraints

https://arxiv.org/pdf/1904.05255.pdf