Semantics

Semantics

荃者所以在鱼，得鱼而忘荃

蹄者所以在兔，得兔而忘蹄

言者所以在意，得意而忘言

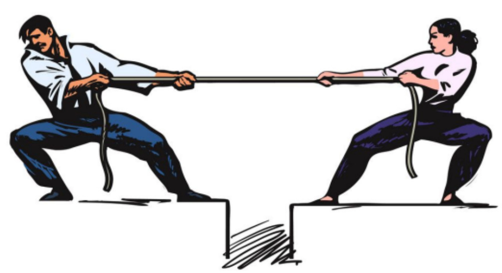
——庄子

Semantics

- Semantics studies meaning, connecting language to the real world
  - Lexical semantics: the meanings of words
  - Sentence semantics
- Implicit vs. explicit meaning representation

Implicit: vector semantics (e.g., word embedding)

Explicit: formal, symbolic representation



Vector Representation of Sentences

- See Ch.3 & Ch.8
- Pros:
  - Automatic learning from data
  - Seamless integration with downstream neural models
  - Impressive performance on many NLP tasks
- Cons:
  - Blackbox: not interpretable and manipulable

Symbolic Representation of Sentences

- The focus of this chapter
- Pros:
  - Interpretable and manipulable
  - Seamless integration with symbolic knowledge bases and inference engines
- Cons:
  - Many forms of representations, unclear which one is “best”
  - Difficult to build an accurate semantic parser

Lexical Semantics

SLP3 Ch 18; INLP Ch 4.2

Word Sense

- A lemma is the dictionary headword form of one or more words (Ref: Ch.2)
  - mouse, mice $\rightarrow$ mouse
  - sing, sang, sung $\rightarrow$ sing
- A lemma can have multiple meanings (polysemous)
  - mouse (N)
    1. any of numerous small rodents...
    2. a hand-operated device that controls a cursor...
  - Each of these is call a word sense

Word Sense

- How to decide different uses of a word should be treated as different senses
  - independent truth conditions
  - Ex:
    - They rarely serve red meat, preferring to prepare seafood.
      - “Help to some food; help with food or drink”
    - He might have served his time, come out and led an upstanding life.
      - “Spend time in prison or in a labor camp”
    - Quite different situations in which the sentences would be true

Word Sense

- How to decide different uses of a word should be treated as different senses
  - independent truth conditions
  - different syntactic behaviors
  - Ex:
    - He might have <u>served his time</u>, come out and led an upstanding life.
      - serve + noun phrase
    - He <u>served as U.S. ambassador</u> to Norway in 1976 and 1977.
      - serve as + noun phrase

Word Sense

- How to decide different uses of a word should be treated as different senses
  - independent truth conditions
  - different syntactic behaviors
  - exhibit antagonistic meanings
  - Ex:
    - Which of those flights serve breakfast?
    - Does Air France serve Philadelphia?
    - Does Air France serve breakfast and Philadelphia?
    - The last sentence is ill-formed.

Word Sense

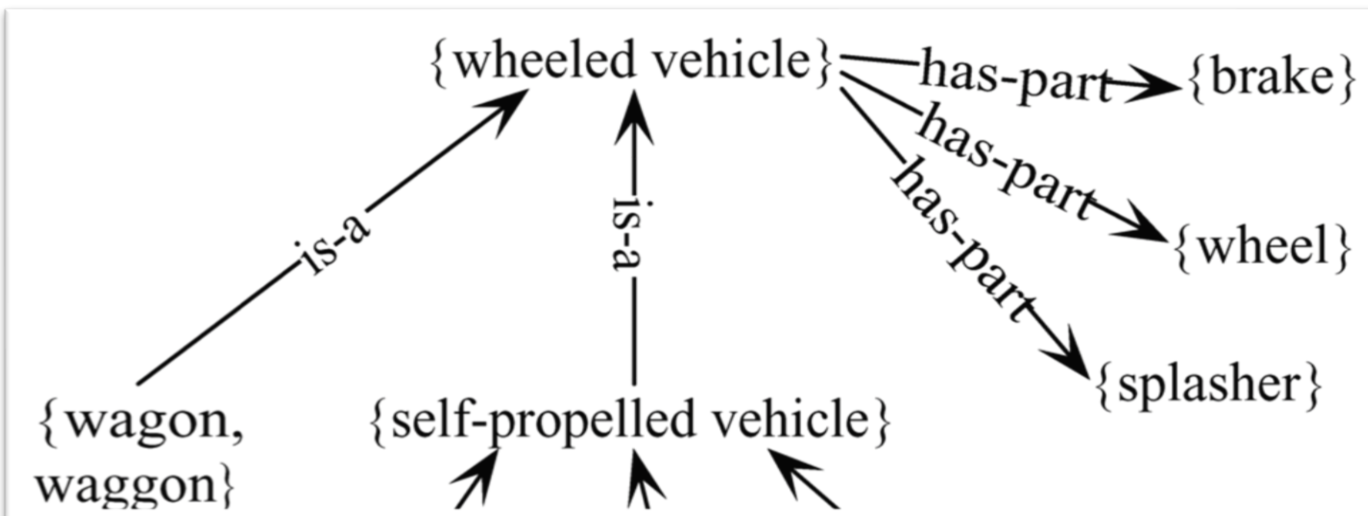
- How to decide different uses of a word should be treated as different senses
  - independent truth conditions
  - different syntactic behaviors
  - exhibit antagonistic meanings
  - independent sense relations
  - Ex:
    - They rarely serve red meat, preferring to prepare seafood.
      - Hypernym: provide, supply
    - He might have served his time, come out and led an upstanding life.
      - Hypernym: spend

Some Semantic Relations

- **Synonymy** – equivalence
  - <small, little>
- **Antonymy** – opposition
  - <small, large>
- **Hyponymy** – subset; is-a relation
  - < dog, mammal>
- **Hypernymy** – superset
  - <mammal, dog>
- **Meronymy** – part-of relation
  - <liver, body>
- **Holonymy** – has-a relation
  - <body, liver>

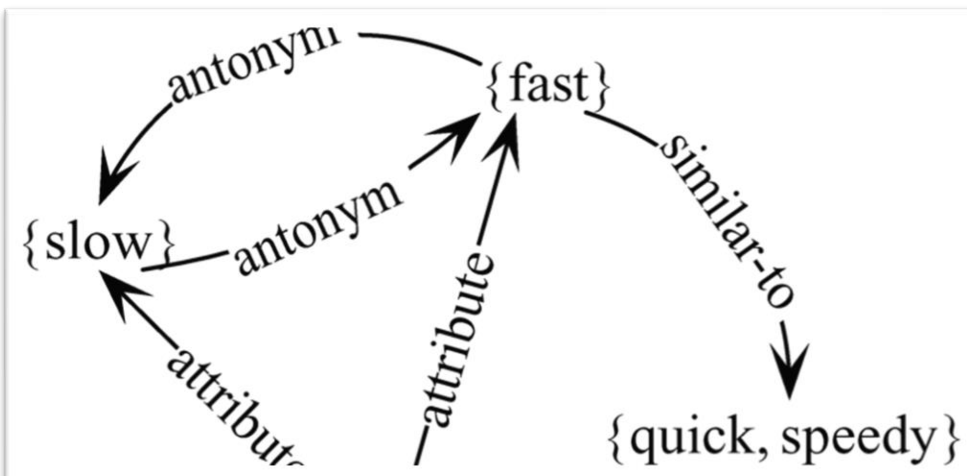
WordNet http://wordnetweb.princeton.edu/perl/webwn

- WordNet is a lexical resource that organizes word senses according to their semantic relations.
- Synset: group of word senses that are synonymous
- Synsets are associated to one another by semantic relations



WordNet http://wordnetweb.princeton.edu/perl/webwn

- WordNet is a lexical resource that organizes word senses according to their semantic relations.
- Synset: group of word senses that are synonymous
- Synsets are associated to one another by semantic relations



Synsets for *dog (n)*

- (n) **dog, domestic dog, Canis familiaris** (a member of the genus Canis that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night" ⊳ dog.n.1
- (n) **frump, dog** (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- (n) **dog** (informal term for a man) "you lucky dog"
- (n) **cad, bounder, blackguard, dog, hound, heel** (someone who is morally reprehensible) "you dirty dog"
- (n) **frank, frankfurter, hotdog, hot dog, dog, wiener, wienerwurst, weenie** (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- …

Direct Hyponyms (subset) for *dog.n.1*

- (n) **puppy** (a young dog)
- (n) **pooch, doggie, doggy, barker, bow-wow** (informal terms for dogs)
- (n) **cur, mongrel, mutt** (an inferior dog or one of mixed breed)
- (n) **lapdog** (a dog small and tame enough to be held in the lap)
- (n) **toy dog, toy** (any of several breeds of very small dogs kept purely as pets)
- (n) **hunting dog** (a dog used in hunting game)
- …

Direct Hypernym (superset) for *dog.n.1*

- (n) **canine, canid** (any of various fissiped mammals with nonretractile claws and typically long muzzles)
- (n) **domestic animal, domesticated animal** (any of various animals that have been tamed and made fit for a human environment)

Part Meronym (part-of) for *dog.n.1*

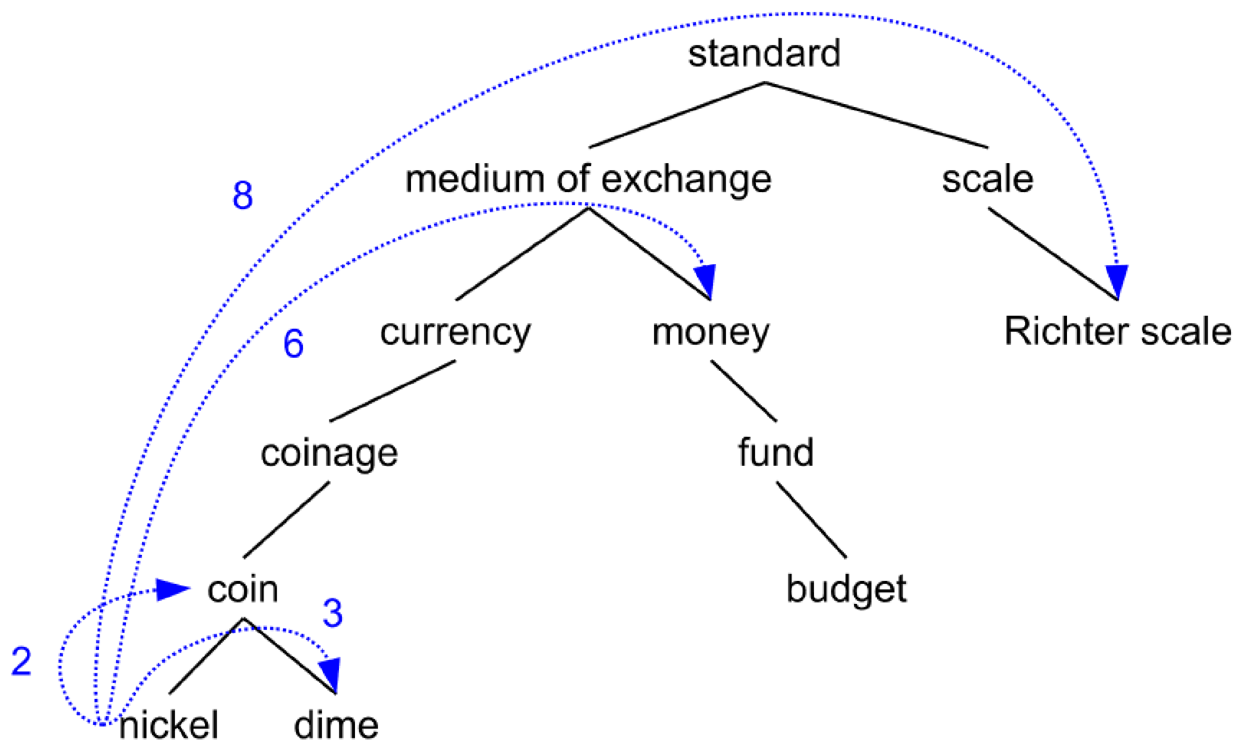
- (n) **flag** (a conspicuously marked or shaped tail)

Member Holonym (has-a) for *dog.n.1*

- (n) **Canis, genus Canis** (type genus of the Canidae: domestic and wild dogs; wolves; jackals)
- (n) **pack** (a group of hunting animals)

Semantic distance based on WordNet

- Given two words, we can compute their semantic distance by the length of the shortest path between their synsets in the hypernymy/hyponymy graph.



Word Sense Disambiguation (WSD)

- Selecting the correct sense for a word in context
  - The set of senses of each word is given (e.g., WordNet)
- Ex: He cashed a check at the bank.

Word Sense Disambiguation (WSD)

- Selecting the correct sense for a word in context
  - The set of senses of each word is given (e.g., WordNet)
- Ex: He **cashed** a check at the bank.
  - (v) cash, cash in (exchange for cash)
  - (adj) cashed (for which money has been paid)

Word Sense Disambiguation (WSD)

- Selecting the correct sense for a word in context
  - The set of senses of each word is given (e.g., WordNet)
- Ex: He cashed a **check** at the bank.
  - ✓ (n) check, bank check, cheque (a written order directing a bank to pay money)
  - (n) check mark, check, tick (a mark indicating that something has been noted or completed etc.)
  - (v) check, check up on, look into, check out, suss out, check over, go over, check into (examine so as to determine accuracy, quality, or condition)
  - (v) check (make an examination or investigation)
  - …

Word Sense Disambiguation (WSD)

- Selecting the correct sense for a word in context
  - The set of senses of each word is given (e.g., WordNet)
- Ex: He cashed a check at the **bank**.
  - (n) bank (sloping land (especially the slope beside a body of water))
  - (n) depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities)
  - (n) bank (a long ridge or pile)
  - (v) bank (tip laterally)
  - …

Word Sense Disambiguation (WSD)

- Selecting the correct sense for a word in context
  - The set of senses of each word is given (e.g., WordNet)
- Methods
  - It's a sequence labeling problem!
  - Can use any of the methods discussed earlier.

Sentence Semantics

SLP3 Ch 15, 16, 19; INLP Ch 12, 13

Sentence Semantics

- Based on lexical semantics
  - The meaning representation of a sentence often contains representations of words in the sentence.
- Can also be applied to a paragraph containing multiple sentences.

Formal Meaning Representation

Meaning Representations

- **Unambiguity:** one representation should have exactly one meaning
- **Canonical form:** one meaning should have exactly one representation

- **Expressiveness:** should be able to handle a wide variety of subject matter
- **Inference ability:** should be able to draw conclusions
- There is a tradeoff between the two…

Meaning Representations

- Special-purpose representations
  - Database query
  - Robot control commands
  - …
- General-purpose representations
  - Formal logic
  - Semantic graphs

Database queries

- To facilitate data exploration and analysis, you might want to parse natural language into database queries (SQL)

which country had the highest carbon emissions last year

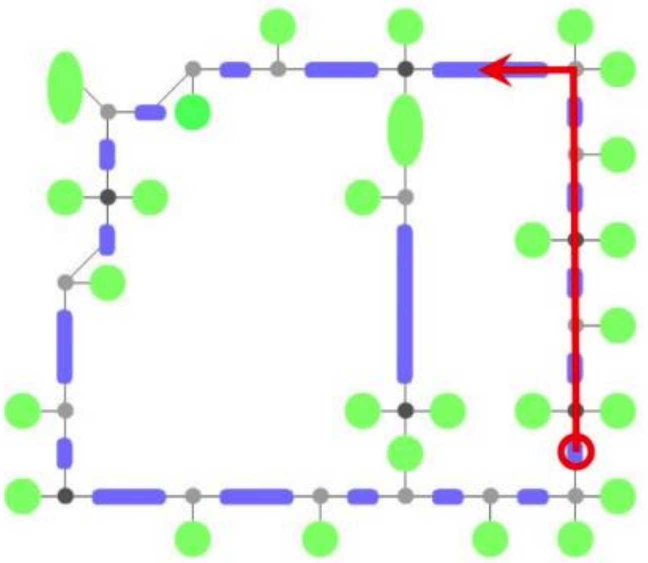
```
SELECT   country.name
FROM     country, co2_emissions
WHERE    country.id = co2_emissions.country_id
AND      co2_emissions.year = 2014
ORDER BY co2_emissions.volume DESC
LIMIT    1;
```

Robot control

- For a robot control application, you might want a custom-designed procedural language:

Go to the third junction and take a left.

```
(do-sequentially
  (do-n-times 3
    (do-sequentially
      (move-to forward-loc)
      (do-until
        (junction current-loc)
        (move-to forward-loc))))
  (turn-left))
```



Intents and arguments

- For smartphone voice commands, you might want relatively simple meaning representations, with intents and arguments:

directions to SF by train

```
(TravelQuery
  (Destination /m/0d6lp)
  (Mode TRANSIT))
```

angelina jolie net worth

```
(FactoidQuery
  (Entity /m/0f4vbz)
  (Attribute /person/net_worth))
```

text my wife on my way

```
(SendMessage
  (Recipient 0x31cbf492)
  (MessageType SMS)
  (Subject "on my way"))
```

play sunny by boney m

```
(PlayMedia
  (MediaType MUSIC)
  (SongTitle "sunny")
  (MusicArtist /m/017mh))
```

First-Order Logic (FOL)

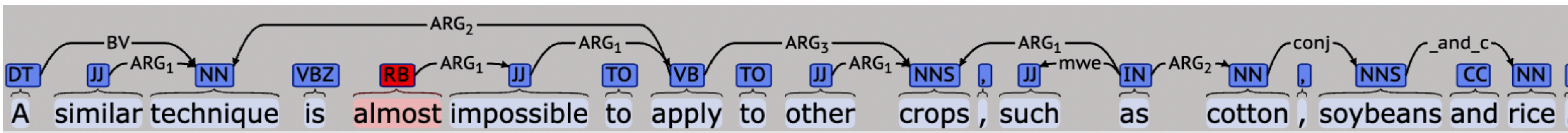
- **Term:** a constant or a variable
- **Formula:** defined recursively
  - If R is an n-ary relation and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a formula.
  - If ϕ is a formula, then its negation, $\neg\phi$, is a formula.
  - If ϕ and ψ are formulas, then binary logical connectives can be used to create formulas:
    - $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \Rightarrow \psi$, …
  - If ϕ is a formula and v is a variable, then quantifiers can be used to create formulas:
    - Universal quantifier: $\forall v, \phi$
    - Existential quantifier: $\exists v, \phi$

Translating Between FOL and Natural Language

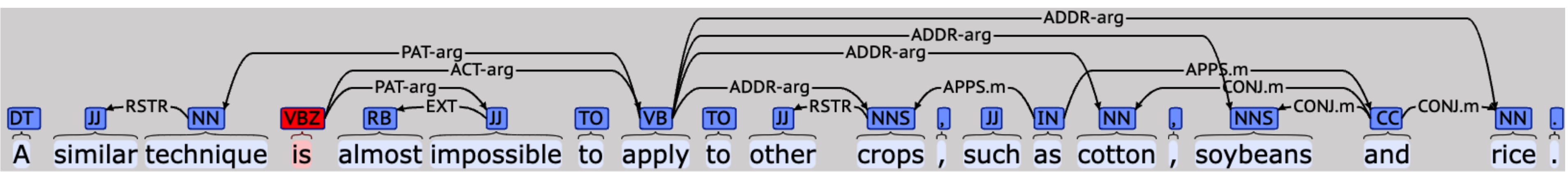
- Alice is not tall
  - $\neg Tall(a)$
- Some people like Broccoli
  - $\exists x, Human(x) \wedge Likes(x, br)$
- If a person likes Thai, then he isn't a friend with Donald
  - $\forall x, Human(x) \wedge Likes(x, th) \Rightarrow \neg Friends(x, d)$
- $\forall x, Restaurant(x) \Rightarrow \left(Longwait(x) \vee \neg Likes(a, x)\right)$
  - Every restaurant has a long wait or is disliked by Adrian
- $\forall x, \exists y, \neg Likes(x, y)$
  - Everybody has something he doesn't like
- $\exists y, \forall x, \neg Likes(x, y)$
  - There exists something that nobody likes

Semantic Graphs

- Flavor 0
  - Node: word
  - Edge: relation
  - Ex: DELPH-IN Minimal Recursion Semantics (DM)

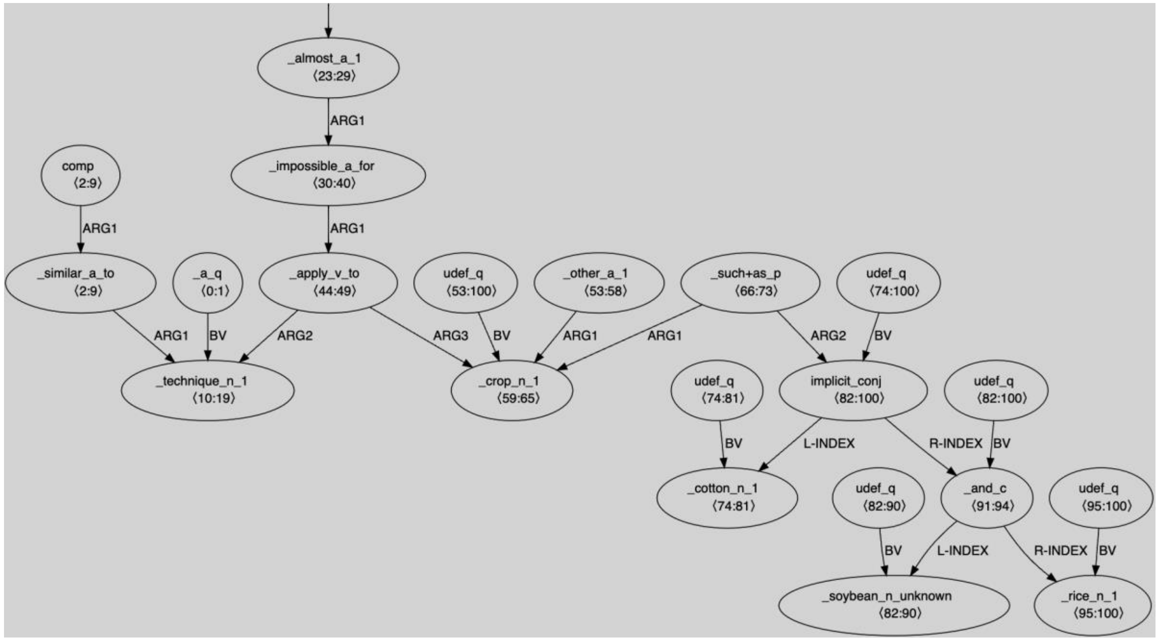


  - Ex: Prague Semantic Dependencies (PSD)



Semantic Graphs

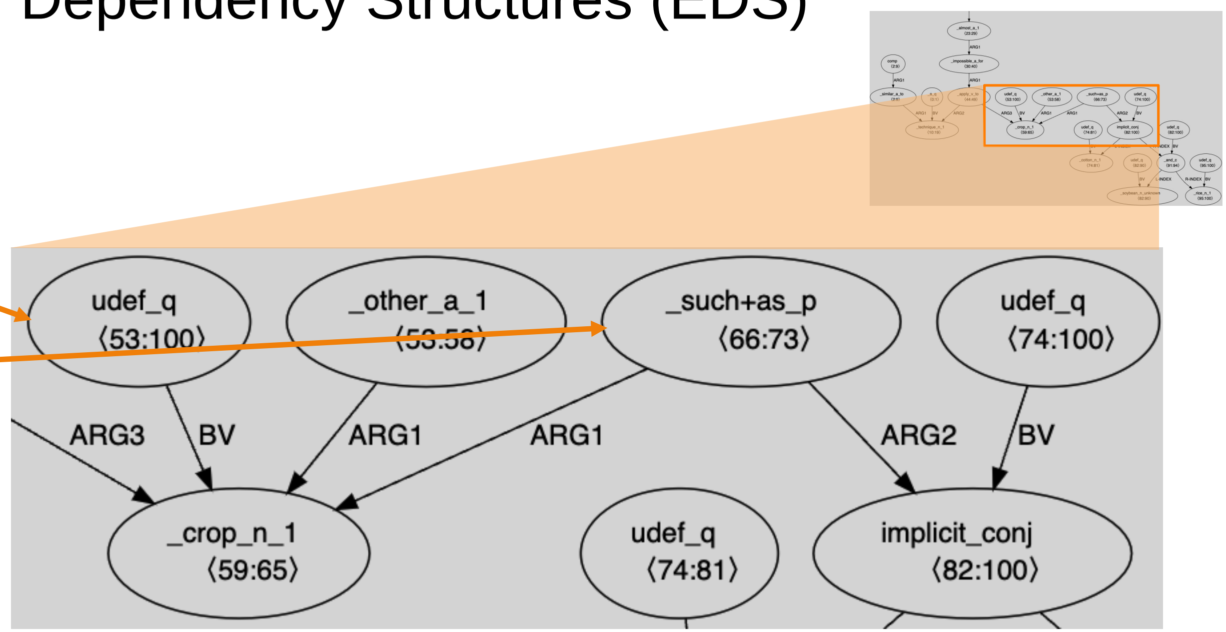
- Flavor 1
  - Node: arbitrary part of the sentence (sub-word, multiple words, no word)
  - Edge: relation
  - Ex: Elementary Dependency Structures (EDS)



Semantic Graphs

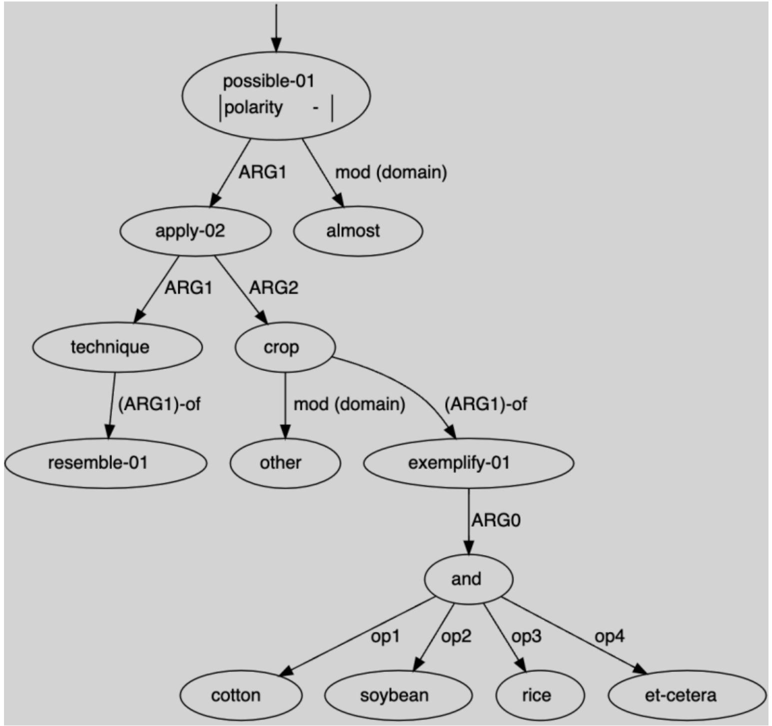
- Flavor 1
  - Node: arbitrary part of the sentence (sub-word, multiple words, no word)
  - Edge: relation
  - Ex: Elementary Dependency Structures (EDS)

…
53 *other*
59 *cops,*
66 *such*
71 *as*
74 *cotton,*
82 *soybeans*
…



Semantic Graphs

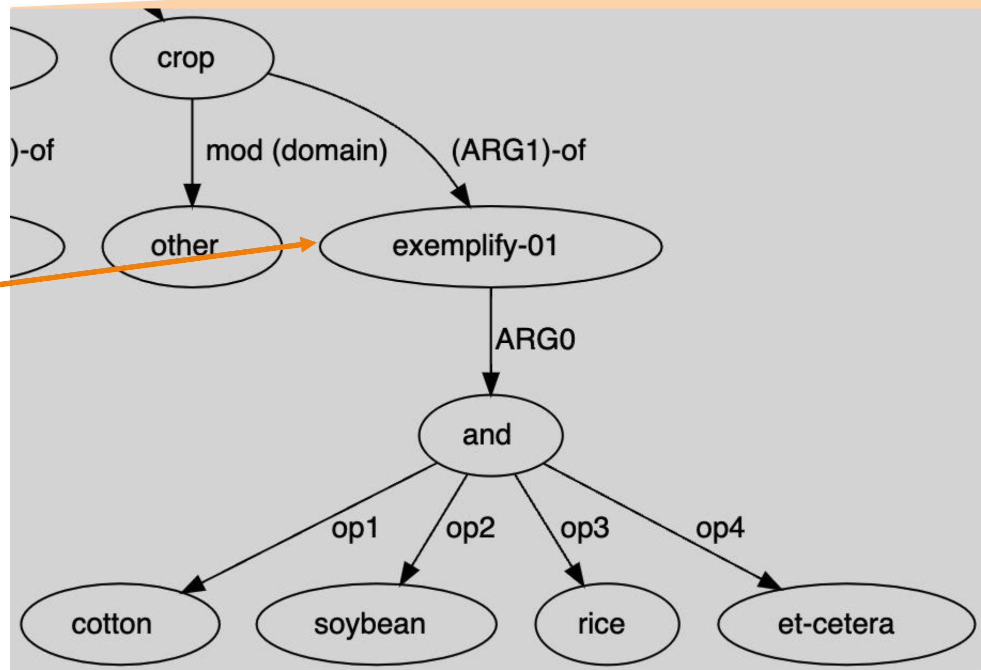
- Flavor 2
  - Node: not necessarily corresponding to words (unanchored)
  - Edge: relation
  - Ex: Abstract Meaning Representation

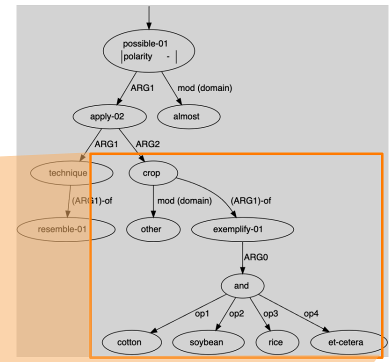


Semantic Graphs

- Flavor 2
  - Node: not necessarily corresponding to words (unanchored)
  - Edge: relation
  - Ex: Abstract Meaning Representation

…
other
cops,
such
as
cotton,
soybeans
…





Semantic Graphs

- Flavor 0
  - Node: word
  - Edge: relation
- Flavor 1
  - Node: arbitrary part of the sentence (sub-word, multiple words, no word)
  - Edge: relation
- Flavor 2
  - Node: not necessarily corresponding to words (unanchored)
  - Edge: relation

Semantic Parsing

- Translating a sentence to its semantic representation
  - Syntax-driven approach
  - Neural approach

Syntax-Driven Semantic Parsing

The Principle of Compositionality

- The meaning of a NL phrase is determined by the meanings of its sub-phrases.
- Phrase ⇨ sub-phrases: this is syntax (constituency parse)!
- Syntax-driven semantic parsing
  - follow a constituency syntactic tree from bottom up
  - repeatedly compose semantics of sub-phrases together

- First of all, we need a way to express semantics of phrases
  - We've already talked about sentence meaning representations, e.g., FOL
  - But phrases are incomplete pieces of meanings

λ-Calculus

- Informally, two extensions over FOL
  - λ-abstraction
    - If ϕ is a FOL formula and v is a variable, then $\lambda v.\phi$ is a λ-term, meaning an unnamed function or map from values (of v) to formulas (usually involving v)
    - Notational conventions:
      $$\lambda x.\left(\lambda y.f(x,y)\right) = \lambda x.\lambda y.f(x,y) = \lambda xy.f(x,y)$$
  - Application (or λ-reduction)
    - If we have $\lambda v.\phi$ and ψ, then $[\lambda v.\phi](\psi)$ is a formula.
    - It can be reduced by substituting every instance of v in ϕ with ψ

λ-Calculus Examples

- $\lambda x. Likes(x, NLP)$
  - A map of someone to a statement that he likes NLP
  - $[\lambda x. Likes(x, NLP)](a)$ reduces to $Likes(a, NLP)$
- $\lambda x. \lambda y. Friends(x, y)$
  - A map of thing x to a map of thing y to a statement that x and y are friends
  - $[\lambda x. \lambda y. Friends(x, y)](a)$ reduces to $\lambda y. Friends(a, y)$
  - $\big[[\lambda x. \lambda y. Friends(x, y)](a)\big](b)$ reduces to $[\lambda y. Friends(a, y)](b)$, which reduces to $Friends(a, b)$
- $\lambda f. f(a, b)$
  - A map of relation f to a statement that a and b have relation f
  - $[\lambda f. f(a, b)](\lambda x. \lambda y. Friends(x, y))$ reduces to $[\lambda x. \lambda y. Friends(x, y)](a, b)$, which reduces $Friends(a, b)$

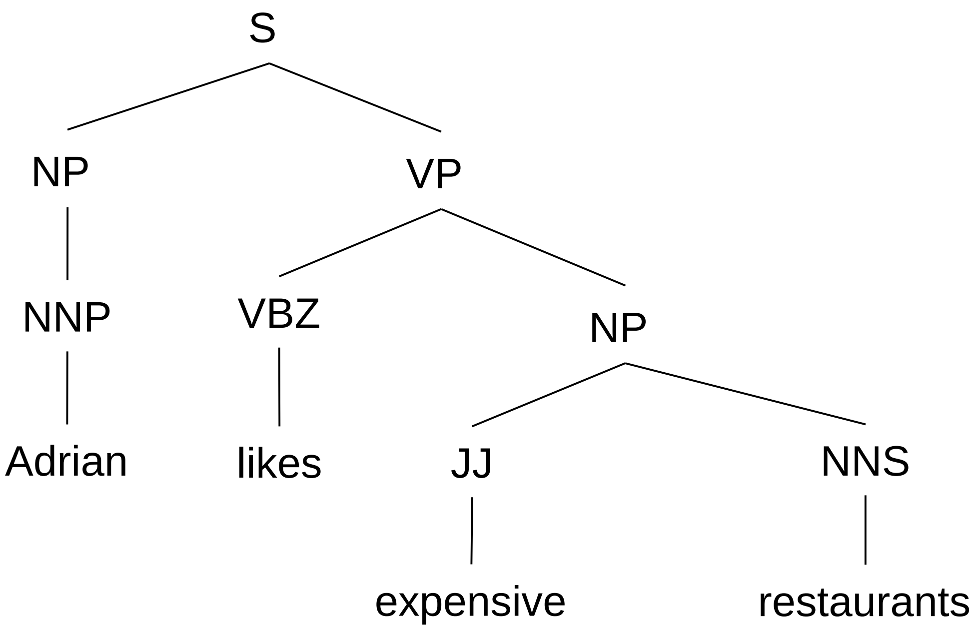
Example CFG

- NNP → Adrian
- VBZ → likes
- JJ → expensive
- NNS → restaurants
- NP → NNP
- NP → JJ NNS
- VP → VBZ NP
- S → NP VP

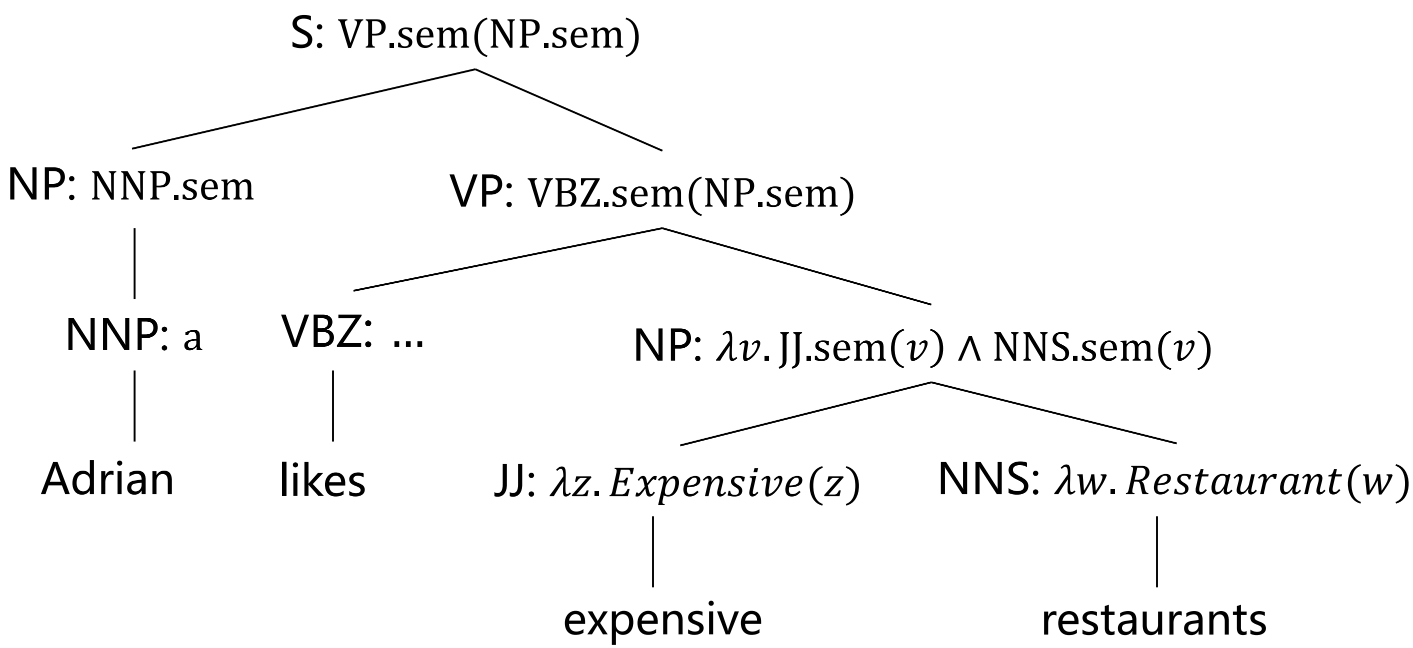
Semantic Attachments to CFG

- NNP $\rightarrow$ Adrian $\{a\}$
- VBZ $\rightarrow$ likes $\{\lambda f y. \forall x f(x) \Rightarrow Likes(y, x)\}$
- JJ $\rightarrow$ expensive $\{\lambda x. Expensive(x)\}$
- NNS $\rightarrow$ restaurants $\{\lambda x. Restaurant(x)\}$
- NP $\rightarrow$ NNP $\{\boxed{\text{NNP.sem}}\}$ an undetermined formula of NNP
- NP $\rightarrow$ JJ NNS $\{\lambda x. \text{JJ.sem}(x) \wedge \text{NNS.sem}(x)\}$
- VP $\rightarrow$ VBZ NP $\{\text{VBZ.sem}(\text{NP.sem})\}$
- S $\rightarrow$ NP VP $\{\text{VP.sem}(\text{NP.sem})\}$

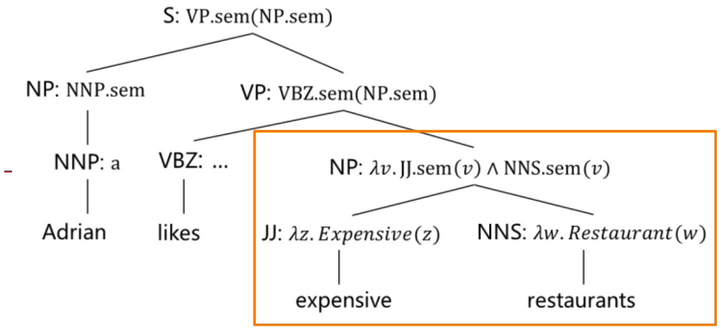
Example

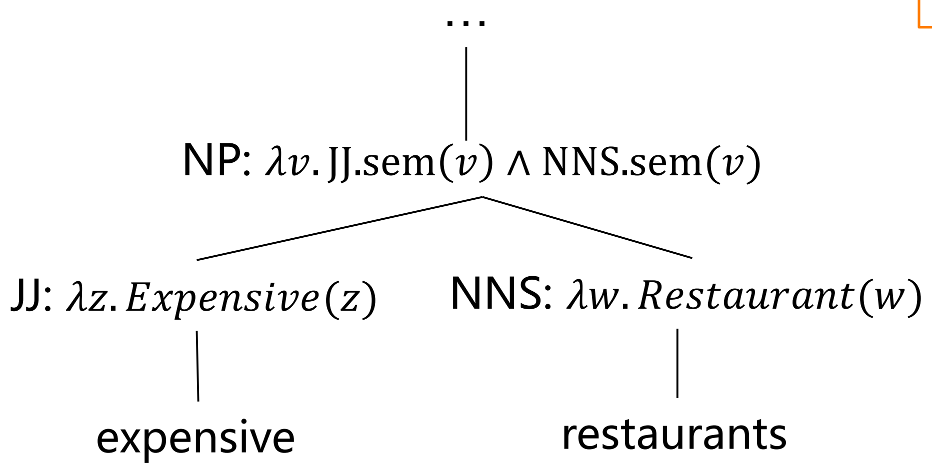


Example


S: VP.sem(NP.sem)
NP: NNP.sem
VP: VBZ.sem(NP.sem)
NNP: a
VBZ: …
NP: $\lambda v.\,$JJ.sem$(v) \wedge$ NNS.sem(v)
Adrian
likes
JJ: $\lambda z.\,Expensive(z)$
NNS: $\lambda w.\,Restaurant(w)$
expensive
restaurants

Example

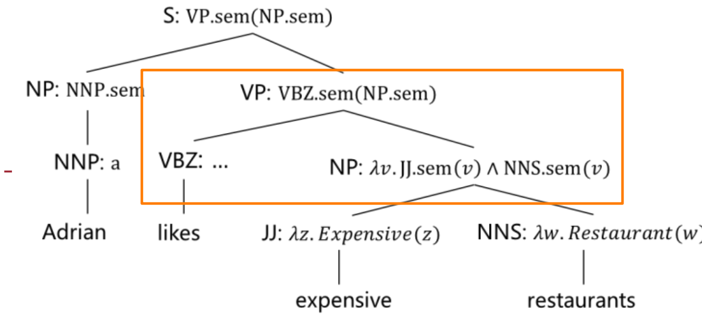


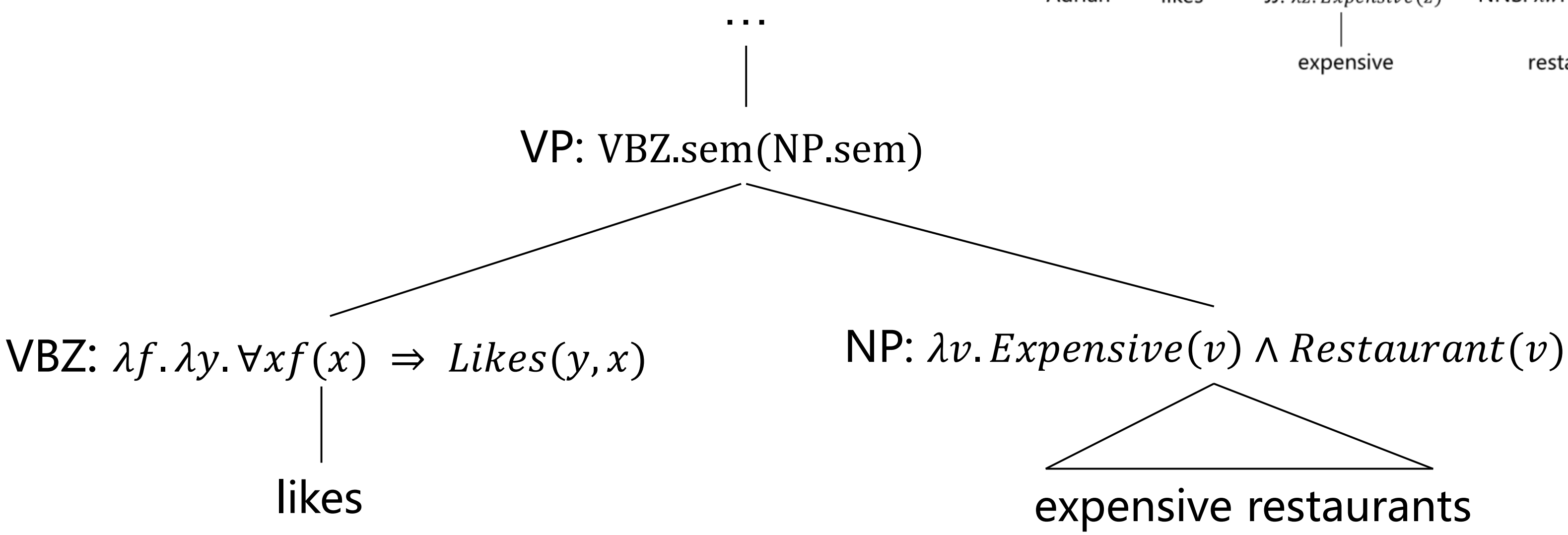


$$\lambda v.\, [\underbrace{\lambda z.\, Expensive(z)}_{\text{JJ.sem}}](v) \wedge [\underbrace{\lambda w.\, Restaurant(w)}_{\text{NNS.sem}}](v)$$

$$\lambda v.\, Expensive(v) \wedge Restaurant(v)$$

Example



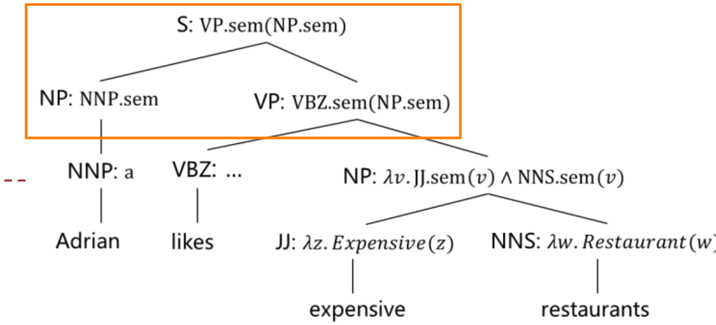


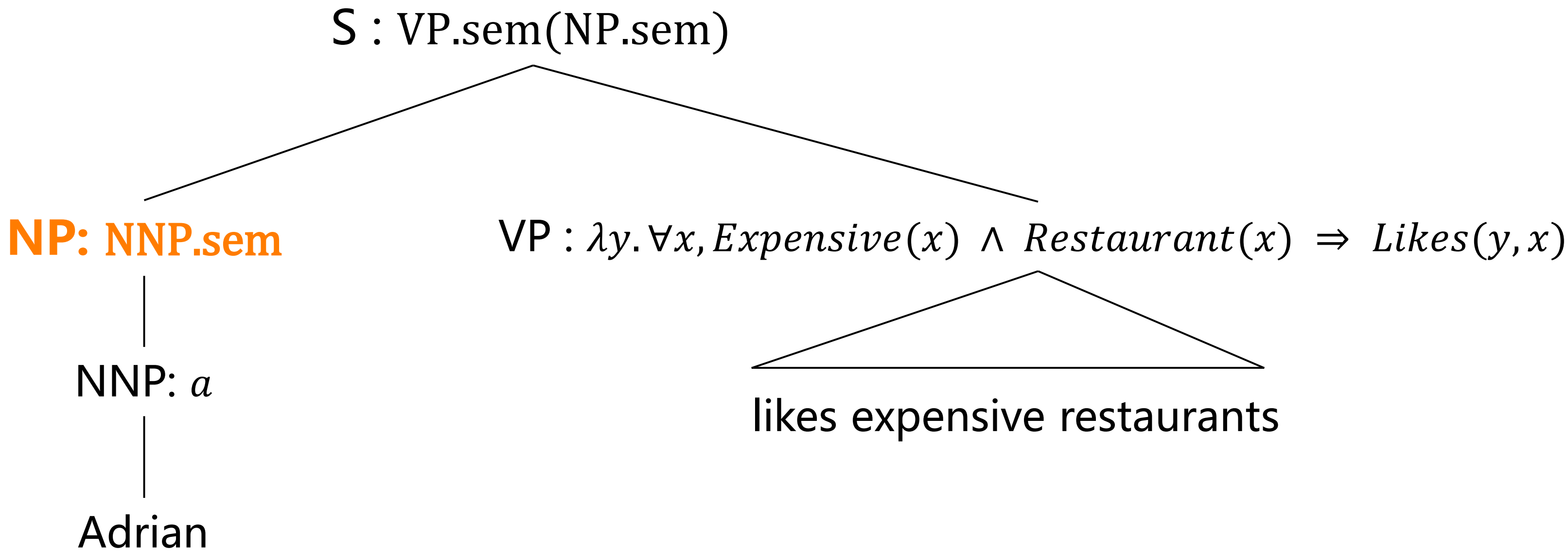
$$\underbrace{[\lambda f.\lambda y.\forall x f(x) \Rightarrow Likes(y,x)]}_{\text{VBZ.sem}}\underbrace{(\lambda v.Expensive(v) \wedge Restaurant(v))}_{\text{NP.sem}}$$

$$\lambda y.\forall x\,[\lambda v.Expensive(v) \wedge Restaurant(v)]\,(x) \Rightarrow Likes(y,x)$$

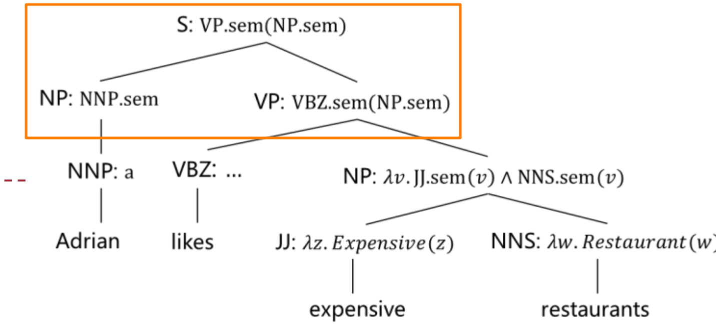
$$\lambda y.\forall x, Expensive(x) \wedge Restaurant(x) \Rightarrow Likes(y,x)$$

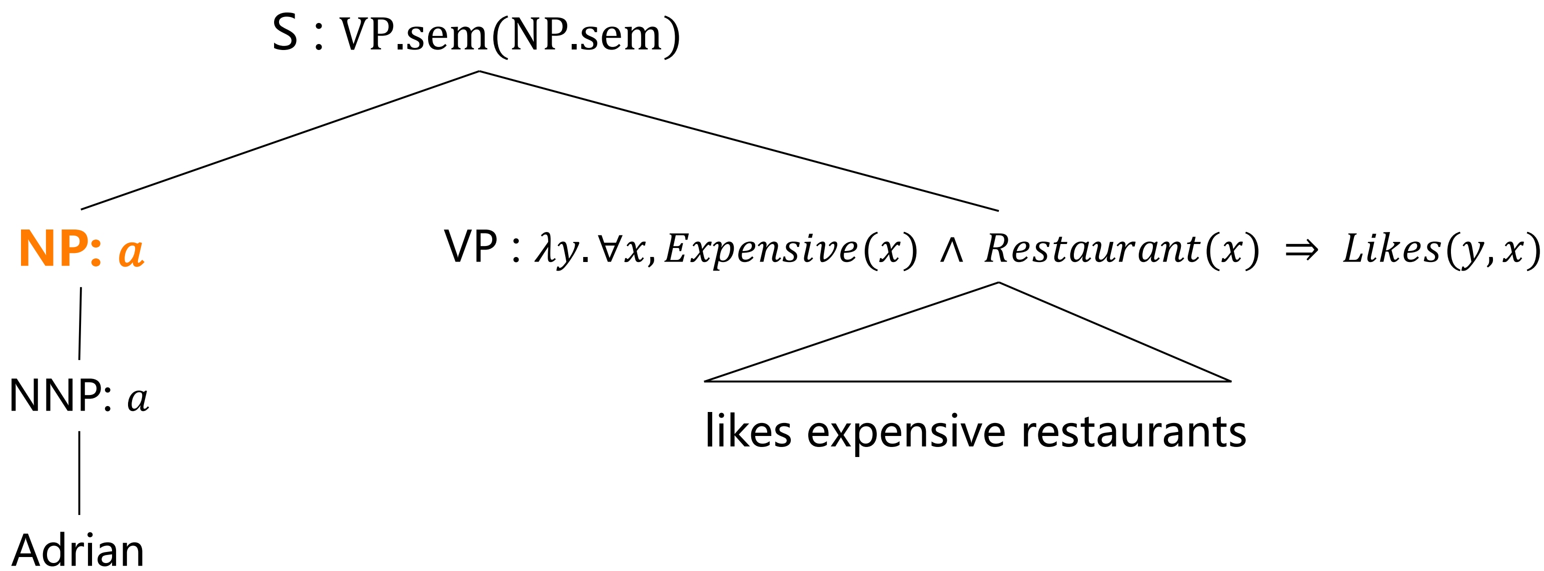
Example


S: VP.sem(NP.sem)
NP: NNP.sem
VP: VBZ.sem(NP.sem)
NNP: a
VBZ: ...
NP: $\lambda v.$ JJ.sem$(v) \wedge$ NNS.sem(v)
Adrian
likes
JJ: $\lambda z. Expensive(z)$
NNS: $\lambda w. Restaurant(w)$
expensive
restaurants

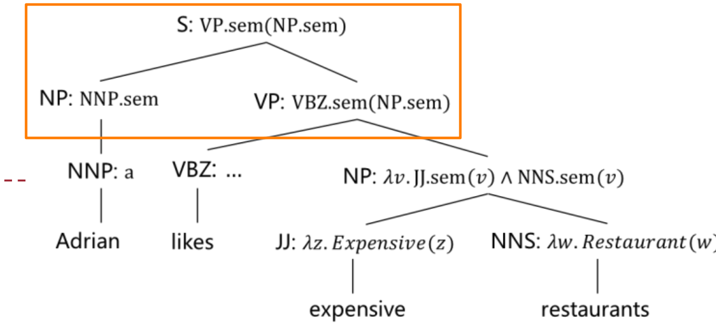

S : VP.sem(NP.sem)
NP: NNP.sem
VP : $\lambda y. \forall x, Expensive(x) \wedge Restaurant(x) \Rightarrow Likes(y, x)$
NNP: a
likes expensive restaurants
Adrian

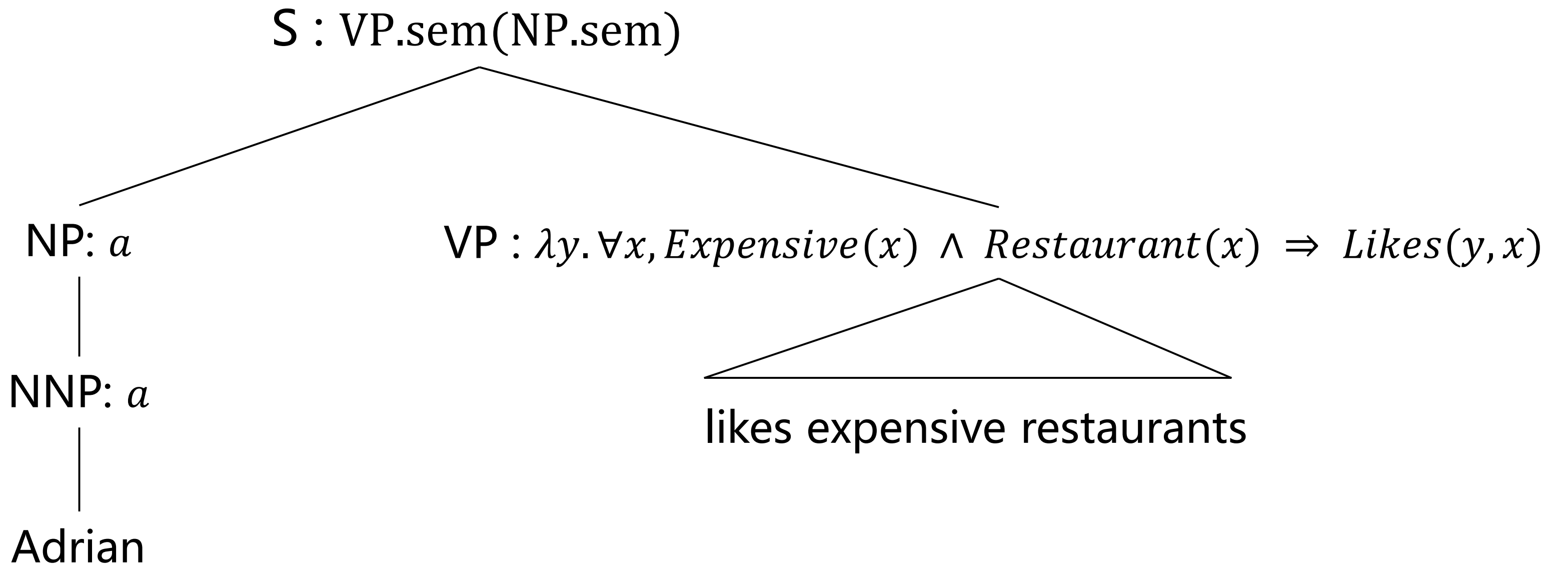
Example


S: VP.sem(NP.sem)
NP: NNP.sem
VP: VBZ.sem(NP.sem)
NNP: a
VBZ: ...
NP: $\lambda v.$ JJ.sem$(v) \wedge$ NNS.sem(v)
Adrian
likes
JJ: $\lambda z.\, Expensive(z)$
NNS: $\lambda w.\, Restaurant(w)$
expensive
restaurants


S : VP.sem(NP.sem)
NP: a
VP : $\lambda y. \forall x, Expensive(x) \wedge Restaurant(x) \Rightarrow Likes(y, x)$
NNP: a
Adrian
likes expensive restaurants

Example





$$\underbrace{[\lambda y. \forall x, Expensive(x) \land Restaurant(x) \Rightarrow Likes(y, x)]}_{\text{VP.sem}}\underbrace{(\ a\ )}_{\text{NP.sem}}$$

$$\forall x, Expensive(x) \land Restaurant(x) \Rightarrow Likes(a, x)$$

Neural Semantic Parsing

Neural Models

- Sequence-to-sequence
  - Input: sentence
  - Output:
    - Logic formula
    - Linearized semantic graph (e.g., depth-first traversal)

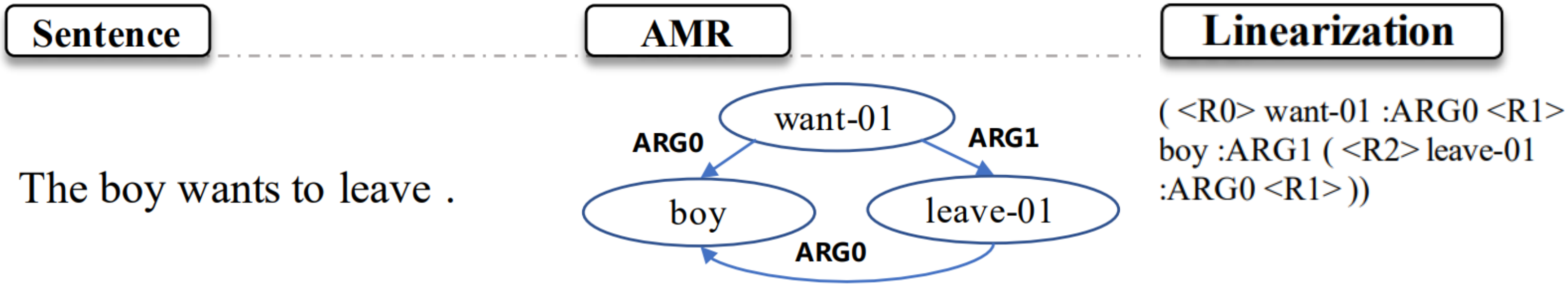


Image from Chen et al., AMRize, then Parse! Enhancing AMR Parsing with PseudoAMRs. 2022

Neural Models

- Parsing to semantic graph
  - Transition-based method
    - Similar to transition-based parsing, but with actions that build a graph instead of a tree

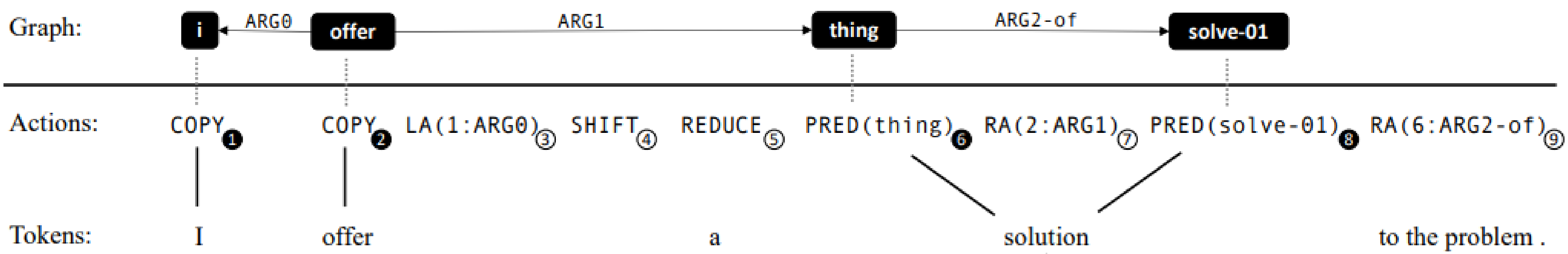


Image from Jiawei et al. AMR Parsing with Action-Pointer Transformer. 2021.

  - Graph-based method
    - First generate a set of nodes (using seq2seq or seq2set) and then predict edges between them (like dependency parsing)

Learning

- Supervised learning
  - Challenge: manual annotation of semantic representations is difficult and costly
- Weakly supervised learning
  - Correct semantic representation not available
  - But in many scenarios, semantic representation is executable and we know the correct outcomes
    - Ex. The correct SQL for a NL question is not known, but the correct answer is known
      - "What is the capital of France?" ⇨ Paris
  - Supervised learning with latent variables, reinforcement learning

Semantic Role Labeling

Semantic Role Labeling (SRL)

- Semantic parsing produces the complete meaning representation of a sentence
- SRL only identifies predicate-argument structures in a sentence
  - A shallow semantic representation
  - No fine-grained meaning representation inside each argument

Who | **did what** | **to whom** | **at where?**

$\underbrace{\text{The police officer}}_{\text{Agent}}\ \underbrace{\text{detained}}_{\text{Predicate}}\ \underbrace{\text{the suspect}}_{\text{Theme}}\ \underbrace{\text{at the scene of the crime}}_{\text{Location}}$

Examples: Who did What to Who(m)?

- **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**...

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

Two widely used semantic role specifications

- **FrameNet**
  - more roles
  - define roles specific to a group of predicates
- **PropBank**
  - fewer roles
  - define generalized semantic roles (prototypes)

<u>XYZ corporation</u> bought the stock

More specific ⟷ More general

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

The volitional causer of an event

PropBank

- Data resource: annotated on top of the Penn Treebank (so arguments are always constituents).
- Each verb sense has a specific set of roles.
- These roles are given numbers rather than names (e.g., Arg0, Arg1).

PropBank Roles

- Arg0: 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)
- Arg1: PROTO-PATIENT
  - Undergoes change of state
  - Causally affected by another participant
  - Stationary relative to movement of another participant

PropBank Roles

- Arg0: PROTO-AGENT
- Arg1: PROTO-PATIENT
- *Arg2-5 are not really that consistent*
  - Arg2: usually: benefactive, instrument, attribute, or end state
  - Arg3: usually: start point, benefactive, instrument, or attribute
  - Arg4: usually: the end point

PropBank Roles

- Arg-M: modifiers or adjuncts of the predicate

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

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.01 (move downward)

- Arg1: logical subject, patient, thing falling
- Arg2: extent, amount fallen
- Arg3: starting point
- Arg4: ending point
- ArgM-LOC: medium

Examples:

- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Palo Alto.

fall.08 (fall back, rely on in emergency)

- Arg0: thing falling back
- Arg1: thing fallen back on

Example:

- World Bank president Paul Wolfowitz has fallen back on his last resort.

fall.08 (fall back, rely on in emergency)

- Arg0: thing falling back
- Arg1: thing fallen back on

Example:

- World Bank president Paul Wolfowitz has fallen back on his last resort.

fall.08 (fall back, rely on in emergency)

- Arg0: thing falling back
- Arg1: thing fallen back on

Example:

- World Bank president Paul Wolfowitz has fallen back on his last resort.

fall.10 (fall for a trick; be fooled by)

- Arg1: the fool
- Arg2: the trick

Example:

- Many people keep falling for the idea that lowering taxes on the rich benefits everyone.

fall.10 (fall for a trick; be fooled by)

- Arg1: the fool
- Arg2: the trick

Example:

- Many people keep falling for the idea that lowering taxes on the rich benefits everyone.

fall.10 (fall for a trick; be fooled by)

- Arg1: the fool
- Arg2: the trick

Example:

- Many people keep falling for the idea that lowering taxes on the rich benefits everyone.

FrameNet https://framenet.icsi.berkeley.edu

- Roles are specific to a frame.
- Frames can be any content word (verb, noun, adjective, adverb)
- About 1,000 frames, each with its own roles
- Both frames and roles are hierarchically organized
- Annotated without syntax, so that arguments can be anything

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)
- It consists of the following words:
  - **Verbs**: advance, climb, decline, decrease, diminish, dip, double, drop, dwindle, edge, explode, fall, fluctuate, gain, grow, increase, jump, move, mushroom, plummet, reach, rise, rocket, shift, skyrocket, slide, soar, swell, swing, triple, tumble
  - **Nouns**: decline, decrease, escalation, explosion, fall, fluctuation, gain, growth, hike, increase, rise, shift, tumble
  - **Adverb**: increasingly

The “Change position on a scale” Frame

- **Item**: entity that has a position on the scale
- **Attribute**: scalar property that the Item possesses
- **Difference**: distance by which an Item changes its position
- **Final state**: Item’s state after the change
- **Final value**: position on the scale where Item ends up
- **Initial state**: Item’s state before the change
- **Initial value**: position on the scale from which the Item moves
- **Value range**: portion of the scale along which values of Attribute fluctuate
- **Duration**: length of time over which the change occurs
- **Speed**: rate of change of the value
- **Group**: the group in which an Item changes the value of an Attribute

The “Change position on a scale” Frame

- Examples

[$_{\text{ITEM}}$ 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].

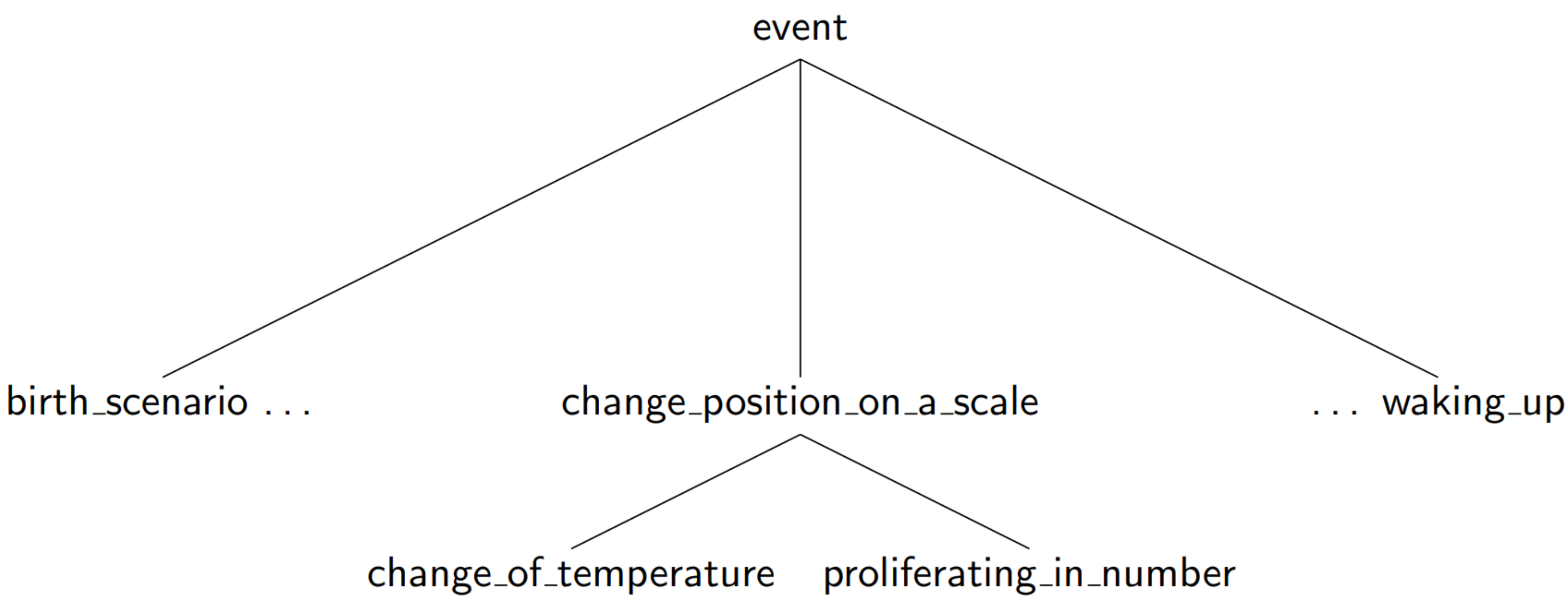
[$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

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*...

The “Change position on a scale” Frame

- Hierarchical organization



Semantic Role Labeling

- Input: a sentence x
- Output:
  - A collection of **predicates**, each consisting of:
    - A span (typically one word)
    - A label, sometimes called the **frame**
    - A set of arguments, each consisting of:
      - A span
      - A label, usually called the **role**

In some settings, predicates are given.

Example:

$$\overbrace{\text{You}}^{\text{ARG0}}\ \text{can't}\ \overbrace{\text{blame}}^{\text{TARGET}}\ \overbrace{\text{the program}}^{\text{ARG1}}\ \overbrace{\text{for being unable to identify it}}^{\text{ARG2}}$$

$$\text{You can't blame}\ \overbrace{\text{the program}}^{\text{ARG0}}\ \text{for being unable to}\ \overbrace{\text{identify}}^{\text{TARGET}}\ \overbrace{\text{it}}^{\text{ARG1}}$$

Methods

- As sequence labeling
  - Use the BIO scheme to represent predicate and argument spans and labels
  - One tag sequence for each predicate

Example

$$\overbrace{\text{You}}^{\text{ARG0}}\ \text{can't}\ \overbrace{\text{blame}}^{\text{TARGET}}\ \overbrace{\text{the program}}^{\text{ARG1}}\ \overbrace{\text{for being unable to identify it}}^{\text{ARG2}}$$

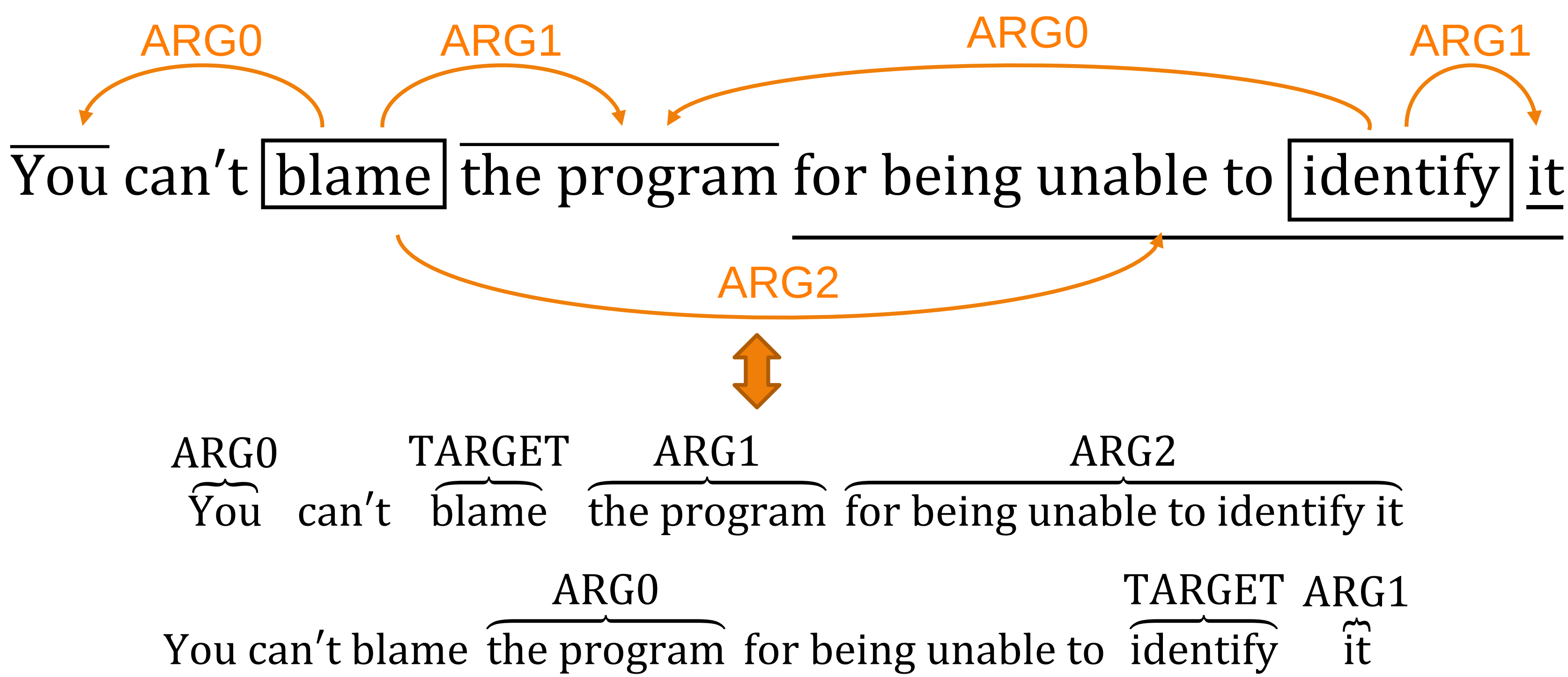
$$\text{You can't blame}\ \overbrace{\text{the program}}^{\text{ARG0}}\ \text{for being unable to}\ \overbrace{\text{identify}}^{\text{TARGET}}\ \overbrace{\text{it}}^{\text{ARG1}}$$

You	can't	blame	the	program	for	being	unable	to	identify	it
B-ARG0	O	B-V	B-ARG1	I-ARG1	B-ARG2	I-ARG2	I-ARG2	I-ARG2	I-ARG2	I-ARG2
O	O	O	B-ARG0	I-ARG0	O	O	O	O	B-V	B-ARG1

Methods

- Graph-based methods
  - First predict predicate and argument spans
  - Then predicate roles as dependency arcs

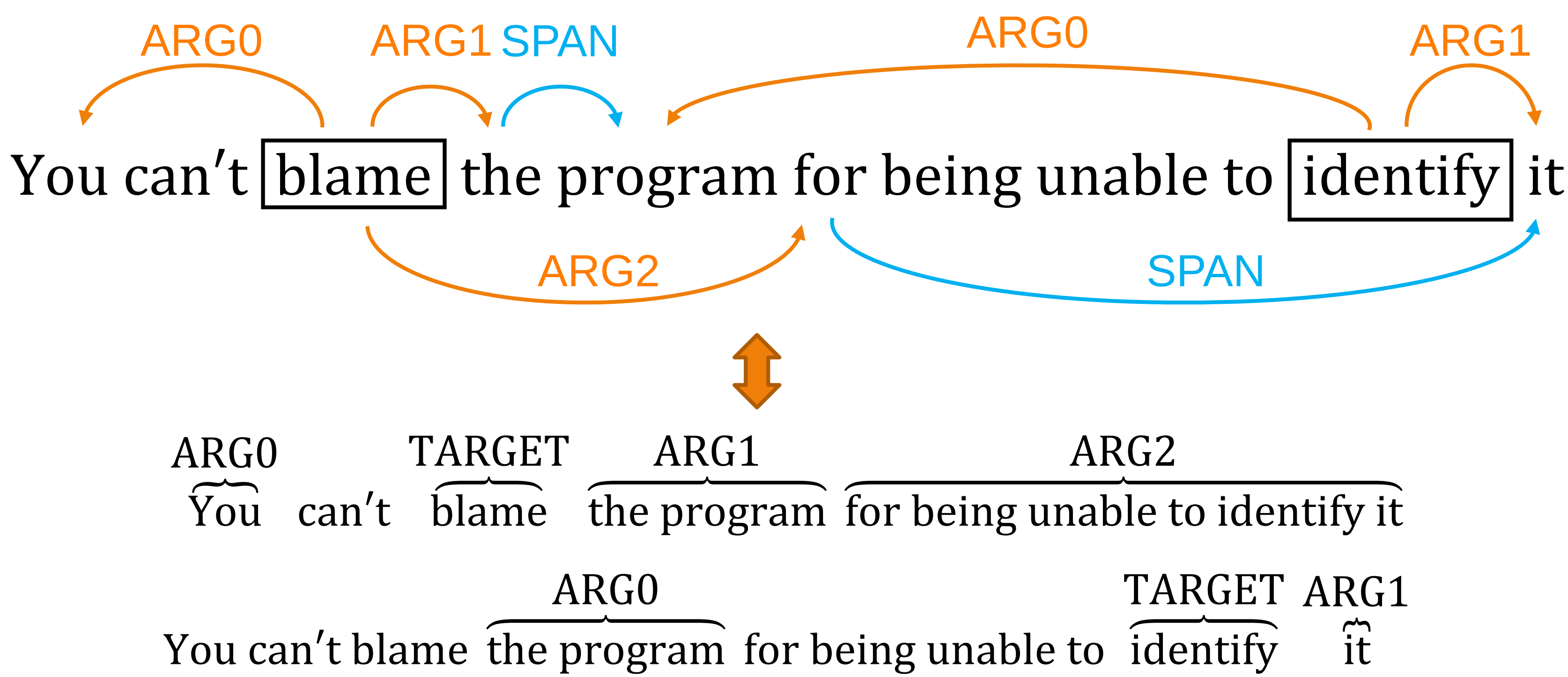
Example



Methods

- Graph-based methods
  - Predict both spans and roles as dependency arcs
  - Can utilize high-order dependency parsing

Example



Methods

- Sequence to sequence
  - Input:
    - Tolkien's epic novel The Lord of the Rings was published in 1954-1955, years after the book was completed.
  - Output:
    - Tolkien's epic novel [The Lord of the Rings | *Arg1*] [was published | *predicate*] [in 1954-1955 | *ArgM-TMP*], years after the book was completed.

Paolini, et al. "Structured prediction as translation between augmented natural languages." ICLR 2021

Information Extraction

SLP Ch 17; INLP Ch 17

Information Extraction (IE)

- Semantic parsing produces the complete meaning representation of a sentence
- SRL only identifies predicate-argument structures in a sentence
- IE only identifies specific information in a sentence, esp. entities and their relations

Named Entity Recognition

Elizabeth Warren, the liberal firebrand who emerged as a top Democratic contender for the **White House** on the strength of an anti-corruption platform backed by a dizzying array of policy proposals, ended her campaign on Thursday. A former bankruptcy law professor who forged a national reputation as a scourge of **Wall Street** even before entering politics, **Warren** had banked on a strong showing on Super Tuesday after a string of disappointing finishes in the early states. But she trailed far behind front-runners **Bernie Sanders** and **Joe Biden**, placing third in her home state of **Massachusetts**, which she continues to represent in the **U.S.** Senate.

- Label certain kinds of proper nouns:
  - Personal names
  - Organizations
  - Geopolitical entities
  - Locations
  - Etc.

Nested Named Entity Recognition

Elizabeth Warren, the liberal firebrand who emerged as a top Democratic contender for the **White House** on the strength of an anti-corruption platform backed by a dizzying array of policy proposals, ended her campaign on Thursday...

… as a top Democratic contender for the White House on …

But she trailed far behind front-runners **Bernie Sanders** and **Joe Biden**, placing third in her home state of **Massachusetts**, which she continues to represent in the **U.S.** Senate.

… in the U.S. Senate.

Entity Linking

Elizabeth Warren, the liberal firebrand who emerged as a top Democratic contender for the White House on the strength of an anti-corruption platform backed by a dizzying array of policy proposals, ended her campaign on Thursday. A former bankruptcy law professor who forged a national reputation as a scourge of Wall Street even before entering politics, **Warren** had banked on a strong showing on Super Tuesday after a string of disappointing finishes in the early states. But she trailed far behind front-runners **Bernie Sanders** and **Joe Biden**, placing third in her home state of Massachusetts, which she continues to represent in the U.S. Senate.

https://www.wikidata.org/wiki/Q434706

Relation Extraction

member_of

WASHINGTON/SELMA, Ala. (Reuters) - **Democratic** U.S. presidential front-runner **Bernie Sanders** raised $46.5 million in February, his campaign said on Sunday, and will launch new television ad buys in nine states with primaries later this month after this week’s Super Tuesday contests. **Joe Biden**’s campaign reported raising $5 million the day of the South Carolina primary. His February haul was $18 million, spokesman Michael Gwin said. Meanwhile, rival **Elizabeth Warren**, who struggled to a fifth-place finish in South Carolina, raised more than $29 million in February, her campaign manager Roger Lau said in a memo to supporters on Sunday.

Event Extraction

WASHINGTON/SELMA, Ala. (Reuters) - **Democratic U.S. presidential front-runner Bernie Sanders** **raised** **$46.5 million** in **February**, his campaign said on Sunday, and will launch new television ad buys in nine states with primaries later this month after this week's Super Tuesday contests. Joe Biden's campaign reported raising $5 million the day of the South Carolina primary. His February haul was $18 million, spokesman Michael Gwin said. Meanwhile, rival Elizabeth Warren, who struggled to a fifth-place finish in South Carolina, raised more than $29 million in February, her campaign manager Roger Lau said in a memo to supporters on Sunday.

Event	TRANSFER-MONEY
Trigger	raised
Recipient-Arg	… Bernie Sanders
Money-Arg	$46.5 million
Time-Arg	February

Similar to SRL

Methods: NER

- As sequence labeling
  - Use the BIO scheme to represent entity spans and types

…	as	a	top	Democratic	contender	for	the	White	House	on	…
…	O	O	O	O	O	O	O	B-ORG	I-ORG	O	…

Methods: NER

- As sequence labeling
  - Use the BIO scheme to represent entity spans and types
  - Problematic when handling nested entities
    - Ambiguous labeling

…	as	a	top	Democratic	contender	for	the	White	House	on	…
…	O	O	O	O	O	O	O	B-ORG	I-ORG	O	…
…	O	O	O	B-MISC	I-MISC	I-MISC	I-MISC	I-MISC	I-MISC	O	…

Methods: NER

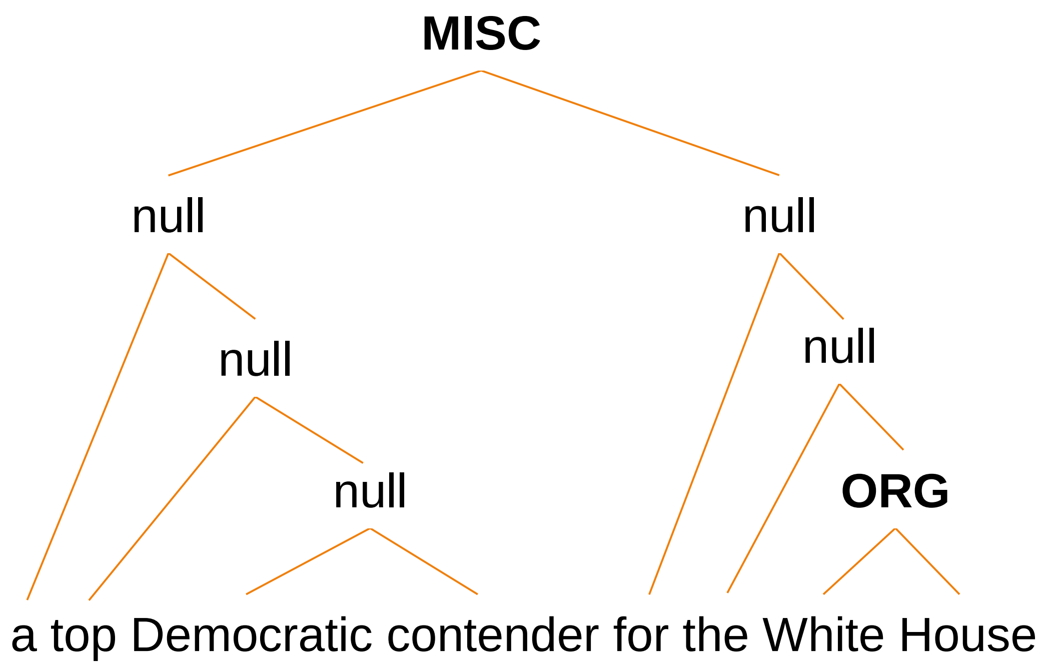
- As span classification
  - For each span, predict its entity type (including NONE)

0 1 2 3 4 5 6 7 8 9 10 …
… as a top **Democratic** **contender** **for** **the** **White** **House** on …

	…	6	7	8	9	10	…
…	…	…	…	…	…	…	…
3	…	NONE	NONE	NONE	**MISC**	NONE	…
4	…	NONE	NONE	NONE	NONE	NONE	…
5		NONE	NONE	NONE	NONE	NONE	…
6			NONE	NONE	NONE	NONE	…
7				NONE	**ORG**	NONE	…
8					NONE	NONE	…
…							…

Methods: NER

- As constituency parsing
  - Entities are constituents in a partially-observed constituency parse tree



Methods: Relation Extraction

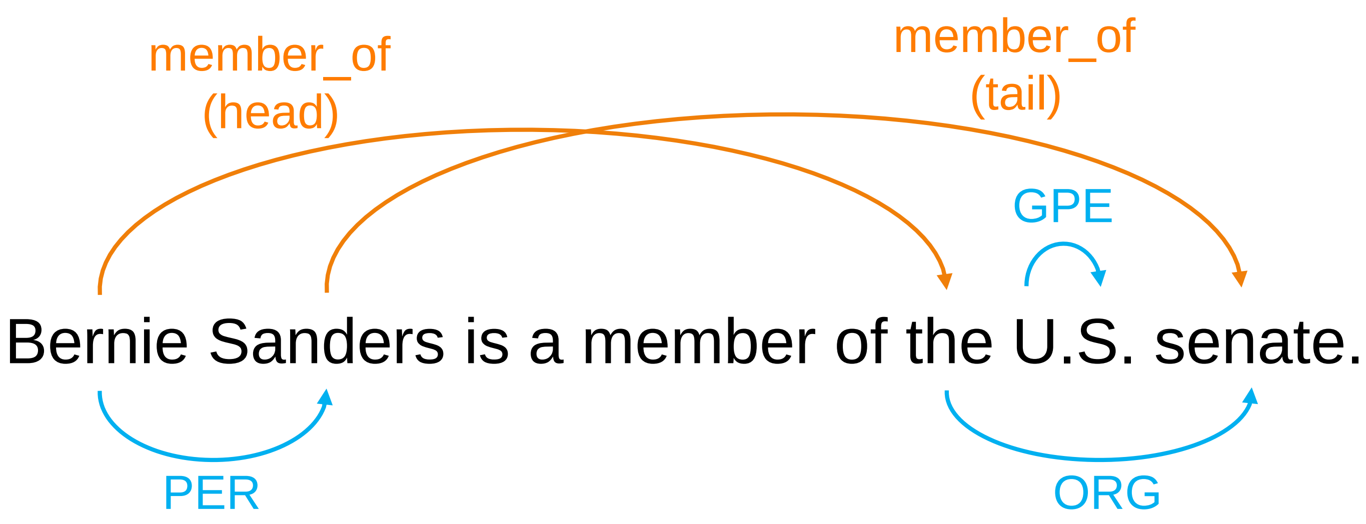
- Given entity spans, predicting relations between them is just like predicting dependency arcs between words
  - Input features
    - Labels of the two entities
    - Text spans of the two entities
    - Text between the two entities
    - Syntactic dependency path between the two entities
  - Output
    - Relation type (including NONE)

Joint Extraction

- Predicting entities and relations jointly, instead of the "entity → relation" pipeline
  - Avoid error propagation
  - Relations may place constraints on entity types
    - Ex: the LIVEIN relation should appear between a PERSON and a LOCATION entity.
- Event extraction is similar
  - Triggers and arguments are like entities
  - Roles are like relations
  - They can be predicted jointly

Joint Extraction

- A graph-based method
  - Each entity represented by one arc
  - Each relation represented by two arcs
  - Scoring and inference are similar to high-order graph-based dependency parsing



TPLinker: Single-stage Joint Extraction of Entities and Relations Through Token Pair Linking. Wang et al. 2020.

Decoding-based IE methods

- Seq2Seq

Steve became CEO of Apple in 1997.

```
(
  (person: Steve
    (work for: Apple)
  )
  (start-position: became
    (employee: Steve)
    (employer: Apple)
    (time: 1997)
  )
  (organization: Apple)
  (time: 1997)
)
```

Unified Structure Generation for Universal Information Extraction. Lu et al. 2022.

Decoding-based IE methods

- Conversation based (e.g., ChatGPT)
  - NER:
    - Q: The given sentence is “My Love Diary is a TV series released in Beijing in 1990”. Given the list of entity types: Person, Location, what entity types are included in this sentence?
    - A: Location
    - Q: Please identify the entities of type “Location” in the given sentence.
    - A: Beijing
  - The same method can be applied to relation and event extraction.

Zero-Shot Information Extraction via Chatting with ChatGPT. Wei et al. 2023.

Summary

Lexical Semantics

- Word Senses
- WordNet
  - Organizing word senses according to their semantic relations
- Word Sense Disambiguation

Sentence Semantics

- Vector vs. symbolic representation of sentences
- Formal Meaning Representation
  - Special-purpose representations
  - General-purpose representations: formal logic, semantic graphs
- Syntax-Driven Semantic Parsing
  - λ-Calculus, Semantic Attachments to CFG
- Neural Semantic Parsing
  - Seq2seq, parsing to graph
- Semantic Role Labeling
  - PropBank, FrameNet
  - Methods: sequence labeling, graph-based methods, seq2seq
- Information Extraction