UMSI

SI 630

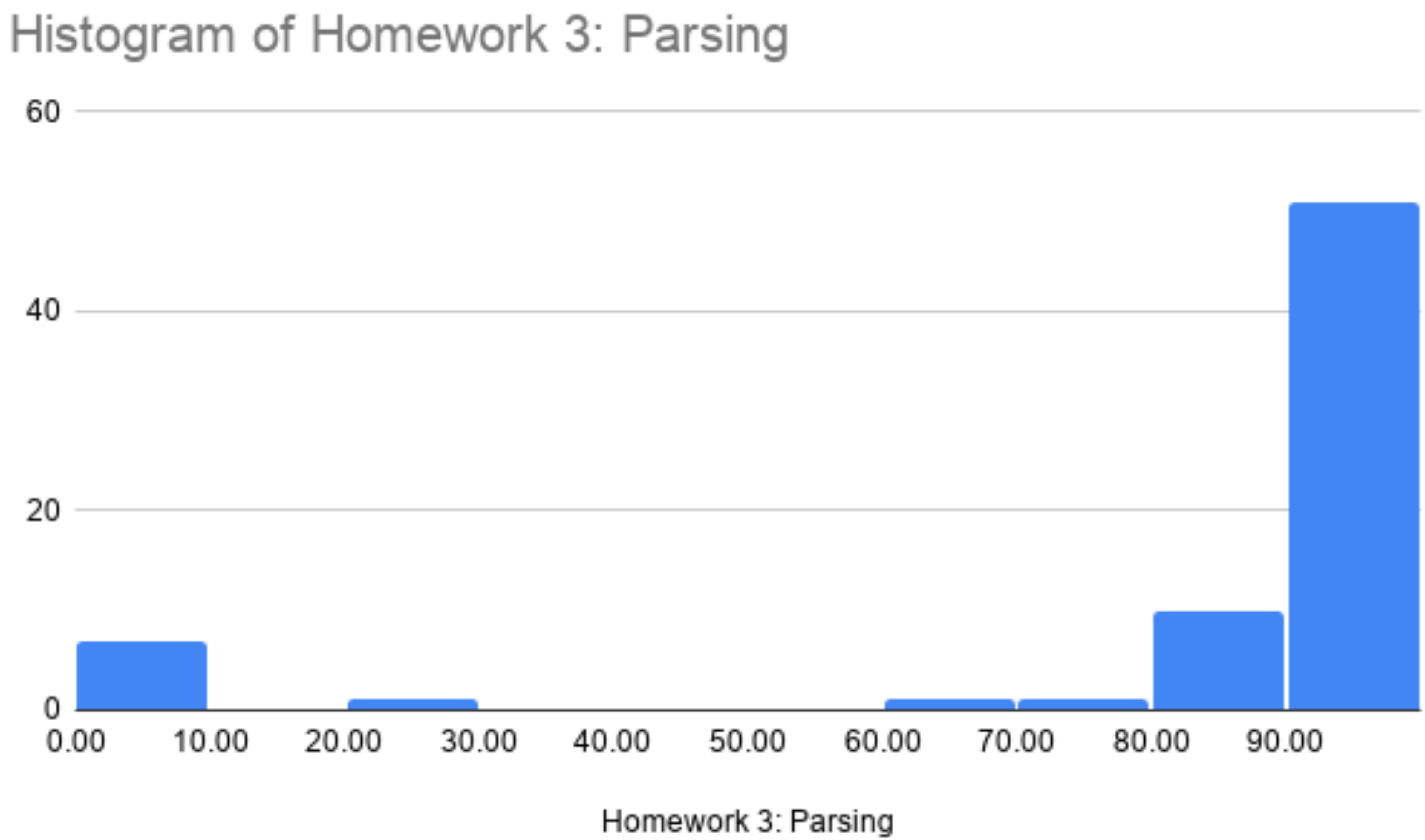
Natural Language Processing: Algorithms and People

Lecture 11: Semantics + Midterm Review

March 25, 2019

Some slides adapted from David Bamman, Dan Jurafsky, and Ray Mooney

Administration


Histogram of Homework 3: Parsing
60
40
20
0
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
Homework 3: Parsing

Homework 5 goes out today (the last one!!)

- Story Generation!

Homework 5
Due Wednesday, April 15

Homework 5
Due Wednesday, April 15

- Two person homework!

Homework 5
Due Wednesday, April 15

- Two person homework!

- We give you lots of stories as training data

Homework 5
Due Wednesday, April 15

- Two person homework!
- We give you lots of stories as training data
- Two-fold goal:
  - Train a deep learning system (GPT) to generate stories
  - Train a deep learning system (BERT) to detect computer generated stories

Goals of HW5

Goals of HW5

- Get familiar with using cutting edge deep learning libraries

Goals of HW5

- Get familiar with using cutting edge deep learning libraries
- Write very little code—think of this like following an sklearn script/blog post

Goals of HW5

- Get familiar with using cutting edge deep learning libraries
- Write very little code—think of this like following an sklearn script/blog post
- Be able to do generation/classification tasks for course projects, jobs, research, fun, etc.

Goals of HW5

- Get familiar with using cutting edge deep learning libraries
- Write very little code—think of this like following an sklearn script/blog post
- Be able to do generation/classification tasks for course projects, jobs, research, fun, etc.
- Make you super employable

Homework 5 probably needs a GPU

Homework 5 probably needs a GPU

- Training time for suggested models is ~10 minutes on a GPU (bare minimum)

Homework 5 probably needs a GPU

- Training time for suggested models is ~10 minutes on a GPU (bare minimum)
- We have free access to GPUs on Great Lakes through this class

Homework 5 probably needs a GPU

- Training time for suggested models is ~10 minutes on a GPU (bare minimum)
- We have free access to GPUs on Great Lakes through this class
  - However, we only have 5 GPUs at a time

Homework 5 probably needs a GPU

- Training time for suggested models is ~10 minutes on a GPU (bare minimum)
- We have free access to GPUs on Great Lakes through this class
  - However, we only have 5 GPUs at a time
  - Jobs are limited to two hours

Project ~~Presentations~~ Blog Posts

Project ~~Presentations~~ Blog Posts

- Instructions go out this week
  - with example blog posts!

Project ~~Presentations~~ Blog Posts

- Instructions go out this week
  - with example blog posts!
- Overarching goals:
  - Practice communicating technical results to a general audience
  - Show off all your hard work
  - Be useful to other people interested in your topic

Project ~~Presentations~~ Blog Posts

- Instructions go out this week
    - with example blog posts!
- Overarching goals:
    - Practice communicating technical results to a general audience
    - Show off all your hard work
    - Be useful to other people interested in your topic
- Ask yourself: what can I teach others about my topic and how to solve my research problem?

Project ~~Presentations~~ Blog Posts

- You can use any platform you want, though we recommend Medium
- Blog must be public, unless you seek approval first
- Blog post should not be:
    - A jupyter notebook with just a few comments
    - A mostly word-based blog (we want to see example, figures, tables, and even some code snippets!)
    - A copy of your final report
        - Important distinction:
            - report is expert-to-expert communication
            - blog post is expert-to-layperson (who is tech-savvy)

Midterm next Friday

Midterm next Friday

- Will be released at 12:01am Friday, due 11:59 Friday

Midterm next Friday

- Will be released at 12:01am Friday, due 11:59 Friday
- No late submissions—not even one minute (srsly)

Midterm next Friday

- Will be released at 12:01am Friday, due 11:59 Friday
- No late submissions—not even one minute (srsly)
- ~2 hours if you have been paying attention in class. Could be longer if not (but you have time)

Midterm next Friday

- Will be released at 12:01am Friday, due 11:59 Friday
- No late submissions—not even one minute (srsly)
- ~2 hours if you have been paying attention in class. Could be longer if not (but you have time)
    - Don’t start at 10:00pm without at least looking at it prior

Semantics

Word representation

…	…
the	0
a	0
an	0
for	0
in	0
on	0
dog	1
cat	0
…	…

dog

4.1
-0.9

dog is a point in V-dimensional space

dog is a point in 2-dimensional space

Polysemy

the movie is so bad , in fact , that it retains that ridiculous tarzan call that was so tirelessly mocked in last summer's comedy " george of the jungle .

Polysemy

the movie is so bad , in fact , that it retains that ridiculous tarzan call that was so tirelessly mocked in last summer's comedy " george of the jungle .

the performances are bad .

Polysemy

the movie is so bad , in fact , that it retains that ridiculous tarzan call that was so tirelessly mocked in last summer's comedy " george of the jungle .

the performances are bad .

Polysemy

Oxford English Dictionary

Polysemy

bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.

Oxford English Dictionary

Polysemy

bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.

Oxford English Dictionary

Polysemy

bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.
awesome	Inspiring awe; appalling, dreadful, weird.	Expressing enthusiastic approval: great, excellent, highly impressive; fantastic.

Oxford English Dictionary

Polysemy

bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.
awesome	Inspiring awe; appalling, dreadful, weird.	Expressing enthusiastic approval: great, excellent, highly impressive; fantastic.
wicked	Bad in moral character, disposition, or conduct	Excellent, splendid; remarkable. slang (orig. U.S.).

Oxford English Dictionary

Polysemy

Sense 1

bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.
awesome	Inspiring awe; appalling, dreadful, weird.	Expressing enthusiastic approval: great, excellent, highly impressive; fantastic.
wicked	Bad in moral character, disposition, or conduct	Excellent, splendid; remarkable. slang (orig. U.S.).

Oxford English Dictionary

Polysemy

	Sense 1	Sense 2
bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.
awesome	Inspiring awe; appalling, dreadful, weird.	Expressing enthusiastic approval: great, excellent, highly impressive; fantastic.
wicked	Bad in moral character, disposition, or conduct	Excellent, splendid; remarkable. slang (orig. U.S.).

Oxford English Dictionary

Word senses

I’m going to the bank

Word senses

I’m going to the bank

- bank_1 = “financial institution”

Word senses

I'm going to the bank

- bank_1 = "financial institution"
- bank_2 = "sloping mound"

Word senses

I'm going to the bank

- bank_1 = "financial institution"
- bank_2 = "sloping mound"
- bank_3 = "biological repository"

Word senses

I’m going to the bank

- $bank_1$ = “financial institution”
- $bank_2$ = “sloping mound”
- $bank_3$ = “biological repository”
- $bank_4$ = “building where a $bank_1$ does its business”

Word senses

- A word sense is a representation of one aspect of a word's meaning.

Breakout session time! (5 min)

- How many meanings are there for
  - “line” (noun)
  - “hot” (adj)
  - “serve” (verb)
- We’ll discuss, so write your definitions down notes somewhere

“Bank”

- S: (n) **bank** (sloping land (especially the slope beside a body of water)) *"they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the currents"*
- S: (n) depository financial institution, **bank**, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities) *"he cashed a check at the bank"; "that bank holds the mortgage on my home"*
- S: (n) **bank** (a long ridge or pile) *"a huge bank of earth"*
- S: (n) **bank** (an arrangement of similar objects in a row or in tiers) *"he operated a bank of switches"*
- S: (n) **bank** (a supply or stock held in reserve for future use (especially in emergencies))
- S: (n) **bank** (the funds held by a gambling house or the dealer in some gambling games) *"he tried to break the bank at Monte Carlo"*
- S: (n) **bank**, cant, camber (a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force)
- S: (n) savings bank, coin bank, money box, **bank** (a container (usually with a slot in the top) for keeping money at home) *"the coin bank was empty"*
- S: (n) **bank**, bank building (a building in which the business of banking transacted) *"the bank is on the corner of Nassau and Witherspoon"*
- S: (n) **bank** (a flight maneuver; aircraft tips laterally about its longitudinal axis (especially in turning)) *"the plane went into a steep bank"*

line

Noun

- S: (n) **line** (a formation of people or things one beside another) *"the li soldiers advanced with their bayonets fixed"; "they were arrayed in lin battle"; "the cast stood in line for the curtain call"*
- S: (n) **line** (a mark that is long relative to its width) *"He drew a line on chart"*
- S: (n) **line** (a formation of people or things one behind another) *"the l stretched clear around the corner"; "you must wait in a long line at th checkout counter"*
- S: (n) **line** (a length (straight or curved) without breadth or thickness; trace of a moving point)
- S: (n) **line** (text consisting of a row of words written across a page or computer screen) *"the letter consisted of three short lines"; "there are lines in every stanza"*
- S: (n) **line** (a single frequency (or very narrow band) of radiation in a spectrum)
- S: (n) **line** (a fortified position (especially one marking the most forwa position of troops)) *"they attacked the enemy's line"*
- S: (n) argumentation, logical argument, argument, line of reasoning, l course of reasoning aimed at demonstrating a truth or falsehood; the methodical process of logical reasoning) *"I can't follow your line of reasoning"*
- S: (n) cable, **line**, transmission line (a conductor for transmitting elect optical signals or electric power)
- S: (n) course, **line** (a connected series of events or actions or developr *"the government took a firm course"; "historians can only point out th lines for which evidence is available"*
- S: (n) **line** (a spatial location defined by a real or imaginary unidimens extent)
- S: (n) wrinkle, furrow, crease, crinkle, seam, **line** (a slight depression in the smoothness of a surface) *"his face has many lines"; "ironing ge most wrinkles"*
- S: (n) pipeline, **line** (a pipe used to transport liquids or gases) *"a pipe runs from the wells to the seaport"*
- S: (n) **line**, railway line, rail line (the road consisting of railroad track a roadbed)
- S: (n) telephone line, phone line, telephone circuit, subscriber line, **lin** telephone connection)

Word senses

- “Serve”

Word senses

- "Serve"

S: (v) **serve**, function (serve a purpose, role, or function) *"The tree stump serves as a table"; "The female students served as a control group"; "This table would serve very well"; "His freedom served him well"; "The table functions as a desk"*
S: (v) **serve** (do duty or hold offices; serve in a specific function) *"He served as head of the department for three years"; "She served in Congress for two terms"*
S: (v) **serve** (contribute or conduce to) *"The scandal served to increase his popularity"*
S: (v) service, **serve** (be used by; as of a utility) *"The sewage plant served the neighboring communities"; "The garage served to shelter his horses"*
S: (v) **serve**, help (help to some food; help with food or drink) *"I served him three times, and after that he helped himself"*
S: (v) **serve**, serve up, dish out, dish up, dish (provide (usually but not necessarily food)) *"We serve meals for the homeless"; "She dished out the soup at 8 P.M."; "The entertainers served up a lively show"*
S: (v) **serve** (devote (part of) one's life or efforts to, as of countries, institutions, or ideas) *"She served the art of music"; "He served the church"; "serve the country"*
S: (v) **serve**, serve well (promote, benefit, or be useful or beneficial to) *"Art serves commerce"; "Their interests are served"; "The lake serves recreation"; "The President's wisdom has served the country well"*
S: (v) **serve**, do (spend time in prison or in a labor camp) *"He did six years for embezzlement"*
S: (v) **serve**, attend to, wait on, attend, assist (work for or be a servant to) *"May I serve you?"; "She attends the old lady in the wheelchair"; "Can you wait on our table, please?"; "Is a salesperson assisting you?"; "The minister served the King for many years"*
S: (v) **serve**, process, swear out (deliver a warrant or summons to someone) *"He was processed by the sheriff"*
S: (v) suffice, do, answer, **serve** (be sufficient; be adequate, either in quality or quantity) *"A few words would answer"; "This car suits my purpose well"; "Will $100 do?"; "A \`B' grade doesn't suffice to get me into medical school"; "Nothing else will serve"*
S: (v) **serve** (do military service) *"She served in Vietnam"; "My sons never served, because they are short-sighted"*
S: (v) **serve**, service (mate with) *"male animals serve the females for breeding purposes"*
S: (v) **serve** (put the ball into play) *"It was Agassi's turn to serve"*

Word senses

- They rarely serve red meat
- He served as U.S. ambassador to Norway
- He might have served his time.

Jurafsky & Martin (2008)

Zeugma

- Conjunction (“yoke”) of antagonistic readings; one test for whether word senses are distinct.

Zeugma

- Conjunction (“yoke”) of antagonistic readings; one test for whether word senses are distinct.

- Which flights serve breakfast?

Zeugma

- Conjunction (“yoke”) of antagonistic readings; one test for whether word senses are distinct.

- Which flights serve breakfast?

- Does Midwest Express serve Philadelphia?

Zeugma

- Conjunction ("yoke") of antagonistic readings; one test for whether word senses are distinct.

- Which flights serve breakfast?
- Does Midwest Express serve Philadelphia?
- Does Midwest Express serve breakfast and Philadelphia?

Zeugma

- Can get productive yoking of multiple meanings

Zeugma

- Can get productive yoking of multiple meanings
    - [They] covered themselves with dust and glory. (Mark Twain)

Zeugma

- Can get productive yoking of multiple meanings
  - [They] covered themselves with dust and glory. (Mark Twain)
  - You are free to execute your laws and your citizens as you see fit. (ST:TNG)

Zeugma

- Can get productive yoking of multiple meanings
    - [They] covered themselves with dust and glory. (Mark Twain)
    - You are free to execute your laws and your citizens as you see fit. (ST:TNG)
    - I like my X like I like my Y: Z

Word senses

I’m going to the bank

Je vais à la ________

Word senses

I’m going to the bank

Je vais à la banque

(financial institution)

Word senses

I’m going to the bank

Je vais à la rive

(side of the river)

Relationship between senses

- Synonymy/antonymy
- Hypernymy
- Metonymy
- Meronymy

Synonym

Synonym

- Two senses of different words are synonyms of each other if their meaning is nearly identical*

Synonym

- Two senses of different words are synonyms of each other if their meaning is nearly identical*
- Two words are never exactly the same in their meaning, distribution of use, dialect or other contexts in which they're licensed.

Synonym

- Two senses of different words are synonyms of each other if their meaning is nearly identical*
- Two words are never exactly the same in their meaning, distribution of use, dialect or other contexts in which they're licensed.
- Synonyms can be exchanged for each other without changing the truth conditions of a sentence.

Synonym

- Two senses of different words are synonyms of each other if their meaning is nearly identical*
- Two words are never exactly the same in their meaning, distribution of use, dialect or other contexts in which they're licensed.
- Synonyms can be exchanged for each other without changing the truth conditions of a sentence.

couch	sofa
filbert	hazelnut
car	automobile
fair	impartial
fair	pale

Synonymy

Synonymy holds between word senses, not words

Synonymy

Synonymy holds between word senses, not words

- How big is that plane?

Synonymy

Synonymy holds between word senses, not words

- How big is that plane?
- Would I be flying on a large or small plane?

Synonymy

Synonymy holds between word senses, not words

- How big is that plane?
- Would I be flying on a large or small plane?

- Miss Nelson, for instance, became a kind of big sister to Benjamin

Synonymy

Synonymy holds between word senses, not words

- How big is that plane?
- Would I be flying on a large or small plane?

- Miss Nelson, for instance, became a kind of big sister to Benjamin
- ?Miss Nelson, for instance, became a kind of large sister to Benjamin

Antonymy

- Two senses of different words are synonyms of each other if their meaning is nearly opposite
- All aspects of meaning are nearly identical between antonyms, except one (very much like synonyms in this respect)

long	short	both describe length
big	little	both describe size
fast	slow	both describe speed
cold	hot	both describe temperature
dark	light	both describe luminescence

Hyponymy

hypo = “under”
(e.g., hypothermia)

- Sense A is a hyponym of sense B if A is a subclass of B

- Formally, entailment: for entity x, $A(x) \Rightarrow B(x)$

hyponym/subordinate	hypernym/superordinate
car	vehicle
mango	fruit
chair	furniture
dog	mammal
mammal	animal

Hyponymy

hypo = “under”
(e.g., hypothermia)

- Hyponymy is generally transitive

hyponym/subordinate	hypernym/superordinate
car	vehicle
mango	fruit
chair	furniture
dog	mammal
mammal	animal
dog	animal

Meronymy

- Part-whole relations. A meronym is a part of a holonym.

meronym	holonym
leg	chair
wheel	car
car	automobile

WordNet

- Lexical database for nouns, verbs and adjectives/ adverbs.

- Each word sense is arranged in a synset (category of near-synonyms) and each synset is related to others in terms of their sense relations.

Relations

Relation	Also Called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	$\textit{breakfast}^1 \rightarrow \textit{meal}^1$
Hyponym	Subordinate	From concepts to subtypes	$\textit{meal}^1 \rightarrow \textit{lunch}^1$
Instance Hypernym	Instance	From instances to their concepts	$\textit{Austen}^1 \rightarrow \textit{author}^1$
Instance Hyponym	Has-Instance	From concepts to concept instances	$\textit{composer}^1 \rightarrow \textit{Bach}^1$
Member Meronym	Has-Member	From groups to their members	$\textit{faculty}^2 \rightarrow \textit{professor}^1$
Member Holonym	Member-Of	From members to their groups	$\textit{copilot}^1 \rightarrow \textit{crew}^1$
Part Meronym	Has-Part	From wholes to parts	$\textit{table}^2 \rightarrow \textit{leg}^3$
Part Holonym	Part-Of	From parts to wholes	$\textit{course}^7 \rightarrow \textit{meal}^1$
Substance Meronym		From substances to their subparts	$\textit{water}^1 \rightarrow \textit{oxygen}^1$
Substance Holonym		From parts of substances to wholes	$\textit{gin}^1 \rightarrow \textit{martini}^1$
Antonym		Semantic opposition between lemmas	$\textit{leader}^1 \Longleftrightarrow \textit{follower}^1$
Derivationally Related Form		Lemmas w/same morphological root	$\textit{destruction}^1 \Longleftrightarrow \textit{destroy}^1$

Figure 17.2 Noun relations in WordNet.

Synsets

synset	gloss
mark, grade, score	a number or letter indicating quality
scratch, scrape, scar, mark	an indication of damage
bell ringer, bull's eye, mark, home run	something that exactly succeeds in achieving its goal
chump, fool, gull, mark, patsy, fall guy, sucker, soft touch, mug	a person who is gullible and easy to take advantage of
mark, stigma, brand, stain	a symbol of disgrace or infamy

Synsets

- S: (n) victim, dupe (a person who is tricked or swindled)
  - S: (n) person, individual, someone, somebody, mortal, soul (a human being) *"there was too much for one person to do"*
    - S: (n) organism, being (a living thing that has (or can develop) the ability to act or function independently)
      - S: (n) living thing, animate thing (a living (or once living) entity)
        - S: (n) whole, unit (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is a unit"*
          - S: (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) *"it was full of rackets, balls and other objects"*
            - S: (n) physical entity (an entity that has physical existence)
              - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Hypernyms of {chump, fool, gull, mark, patsy, fall guy, sucker, soft touch, mug} sunset

WordNet + Distributional Semantics

- WordNet encodes human-judged measures of similarity. Learn distributed representations of words that respect WordNet similarities (Faruqui et al. 2015)

- By indexing word senses, we can build annotated resources on top of it for word sense disambiguation.

Semcor

- Semcor: 200K+ words from Brown corpus tagged with Wordnet senses.
    - http://web.eecs.umich.edu/~mihalcea/downloads/semcor/semcor3.0.tar.gz

original	It urged that the city take steps to remedy this problem
lemma sense	It urge^{1} that the city^{2} take^{1} step^{1} to remedy^{1} this problem^{2}
synset number	It urge$^{2:32:00}$ that the city$^{1:15:01}$ take$^{2:41:04}$ step$^{1:04:02}$ to remedy$^{2:30:00}$ this problem$^{1:10:00}$

www.bit.ly/wordnet-serve

- They rarely serve red meat
- Eisenhower served as Supreme Commander of the Allied Expeditionary Forces during WWII.
- He might have served his time.

Word senses

- “Serve”

Word senses

- "Serve"

S: (v) **serve**, function (serve a purpose, role, or function) *"The tree stump serves as a table"; "The female students served as a control group"; "This table would serve very well"; "His freedom served him well"; "The table functions as a desk"*
S: (v) **serve** (do duty or hold offices; serve in a specific function) *"He served as head of the department for three years"; "She served in Congress for two terms"*
S: (v) **serve** (contribute or conduce to) *"The scandal served to increase his popularity"*
S: (v) service, **serve** (be used by; as of a utility) *"The sewage plant served the neighboring communities"; "The garage served to shelter his horses"*
S: (v) **serve**, help (help to some food; help with food or drink) *"I served him three times, and after that he helped himself"*
S: (v) **serve**, serve up, dish out, dish up, dish (provide (usually but not necessarily food)) *"We serve meals for the homeless"; "She dished out the soup at 8 P.M."; "The entertainers served up a lively show"*
S: (v) **serve** (devote (part of) one's life or efforts to, as of countries, institutions, or ideas) *"She served the art of music"; "He served the church"; "serve the country"*
S: (v) **serve**, serve well (promote, benefit, or be useful or beneficial to) *"Art serves commerce"; "Their interests are served"; "The lake serves recreation"; "The President's wisdom has served the country well"*
S: (v) **serve**, do (spend time in prison or in a labor camp) *"He did six years for embezzlement"*
S: (v) **serve**, attend to, wait on, attend, assist (work for or be a servant to) *"May I serve you?"; "She attends the old lady in the wheelchair"; "Can you wait on our table, please?"; "Is a salesperson assisting you?"; "The minister served the King for many years"*
S: (v) **serve**, process, swear out (deliver a warrant or summons to someone) *"He was processed by the sheriff"*
S: (v) suffice, do, answer, **serve** (be sufficient; be adequate, either in quality or quantity) *"A few words would answer"; "This car suits my purpose well"; "Will $100 do?"; "A `B' grade doesn't suffice to get me into medical school"; "Nothing else will serve"*
S: (v) **serve** (do military service) *"She served in Vietnam"; "My sons never served, because they are short-sighted"*
S: (v) **serve**, service (mate with) *"male animals serve the females for breeding purposes"*
S: (v) **serve** (put the ball into play) *"It was Agassi's turn to serve"*

"All-word" Word Sense Disambiguation (WSD)

"Only$_{\text{only1}}$ a relative$_{\text{relative1}}$ handful$_{\text{handful1}}$ of such$_{\text{such0}}$ reports$_{\text{report3}}$ was received$_{\text{receive2}}$"

- For all content words in a sentence, resolve each token to its sense in an fixed sense inventory (e.g., WordNet).

Word Sense Disambiguation (WSD)

- Dictionary methods (Lesk)
- Supervised (machine learning)
- Semi-supervised (Bootstrapping)

Dictionary methods

- Predict the sense a given token that has the highest overlap between the token's context and sense's dictionary gloss.

Dictionary methods

bank^1	Gloss:	a financial institution that accepts deposits and channels the money into lending activities
	Examples:	"he cashed a check at the bank", "that bank holds the mortgage on my home"
bank^2	Gloss:	sloping land (especially the slope beside a body of water)
	Examples:	"they pulled the canoe up on the bank", "he sat on the bank of the river and watched the currents"

"The boat washed up on the the river bank."

Lesk Algorithm

function SIMPLIFIED LESK(*word, sentence*) **returns** best sense of *word*

best-sense ← most frequent sense for *word*
max-overlap ← 0
context ← set of words in *sentence*
for each *sense* **in** senses of *word* **do**
  signature ← set of words in the gloss and examples of *sense*
  overlap ← COMPUTEOVERLAP(*signature*, *context*)
  if *overlap* > *max-overlap* **then**
    max-overlap ← *overlap*
    best-sense ← *sense*
end
return(*best-sense*)

Lesk Algorithm

- Extension (Basile et al. 2014): measure similarity between gloss g = {g_1, … g_G} and context c = {c_1, …, c_C} as cosine similarity between sum of distributed representations

$$\cos\left(\sum_{i=1}^{G} g_i, \sum_{i=1}^{C} c_i\right)$$

Supervised WSD

- We have labeled training data; let's learn from it.
  - Decision trees (Yarowsky 1994)
  - Naive Bayes, log-linear classifiers, support vector machines (Zhong and Ng 2010)
  - Bidirectional LSTM (Raganato et al. 2017)

Supervised WSD

- Collocational: words in specific positions before/after the target word to be disambiguation
- Bag-of-words: words in window around target (without encoding specific position)

feature
w_{i-1} = fish
w_{i-2}= fish
w_{i+1} = fish
w_{i+2} = fish
word in context = fish
…

Supervised learning

“I got money from the bank to buy a bass”

- Pre-processing: part of speech tagging, lemmatization, syntactic parsing (headwords, dependency relations)

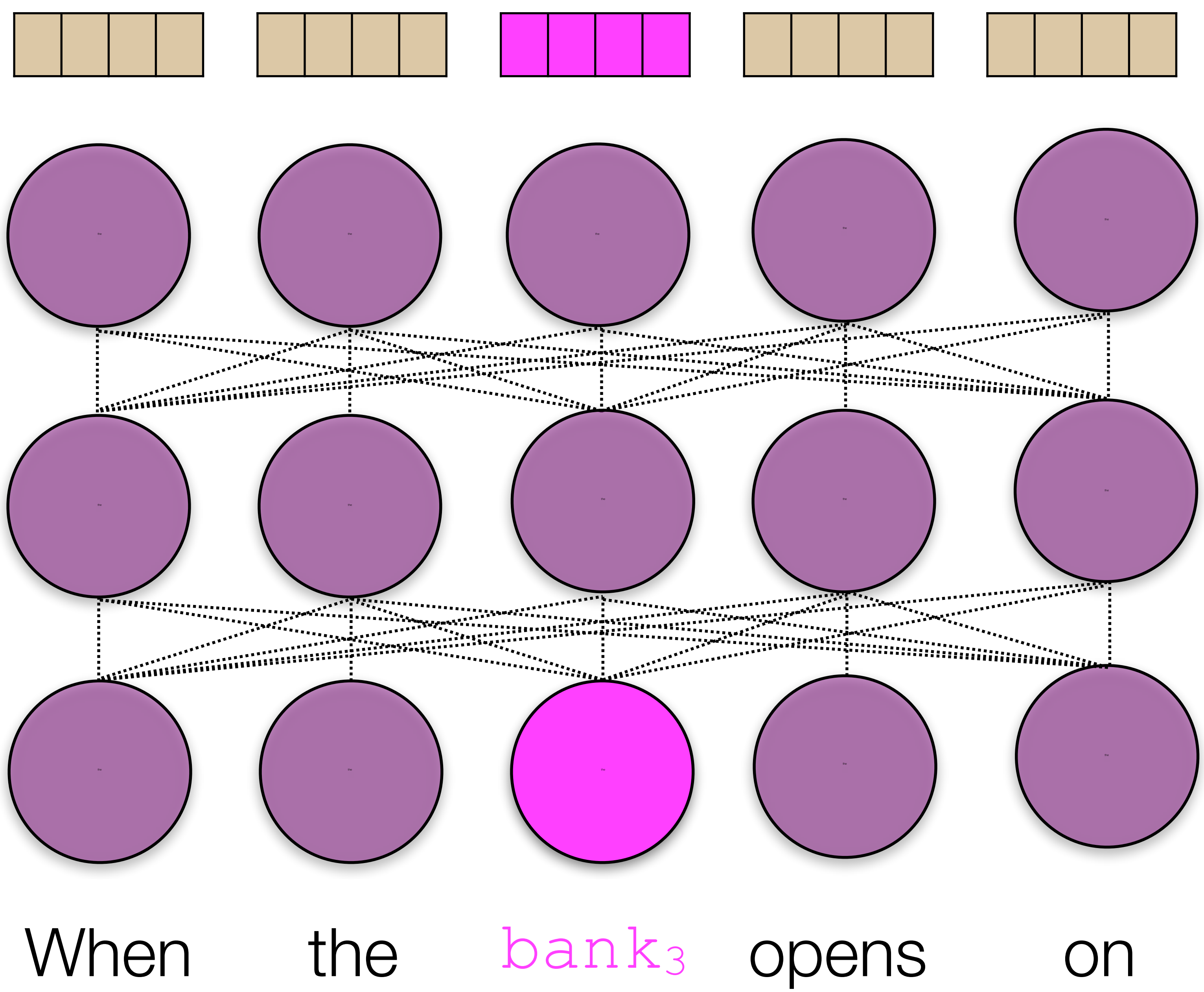
- Collocations:
  - Token 1 word to the left, 1 word to the right
- All words within window of n tokens

	Dev	Test Datasets				Concatenation of All Test Datasets				
	SE07	**SE2**	**SE3**	**SE13**	**SE15**	**Nouns**	**Verbs**	**Adj.**	**Adv.**	**All**
BLSTM	61.8	71.4	68.8	65.6	69.2	70.2	56.3	75.2	**84.4**	68.9
BLSTM + att.	62.4	71.4	**70.2**	66.4	70.8	71.0	**58.4**	75.2	83.5	69.7
BLSTM + att. + LEX	63.7	**72.0**	69.4	66.4	**72.4**	**71.6**	57.1	**75.6**	83.2	**69.9**
BLSTM + att. + LEX + POS	**64.8**	**72.0**	69.1	**66.9**	71.5	71.5	57.5	75.0	83.8	**69.9**
Seq2Seq	60.9	68.5	67.9	65.3	67.0	68.7	54.5	74.0	81.2	67.3
Seq2Seq + att.	62.9	69.9	69.6	65.6	67.7	69.5	57.2	74.5	81.8	68.4
Seq2Seq + att. + LEX	64.6	70.6	67.8	66.5	68.7	70.4	55.7	73.3	82.9	68.5
Seq2Seq + att. + LEX + POS	63.1	70.1	68.5	66.5	69.2	70.1	55.2	75.1	84.4	68.6
IMS	61.3	70.9	69.3	65.3	69.5	70.5	55.8	75.6	82.9	68.9
IMS+emb	**62.6**	**72.2**	**70.4**	65.9	71.5	**71.9**	56.6	**75.9**	**84.7**	**70.1**
Context2Vec	61.3	71.8	69.1	65.6	**71.9**	71.2	**57.4**	75.2	82.7	69.6
Lesk_{ext}+emb	⋆56.7	63.0	63.7	66.2	64.6	70.0	51.1	51.7	80.6	64.2
UKB_{gloss} w2w	42.9	63.5	55.4	⋆62.9	63.3	64.9	41.4	69.5	69.7	61.1
Babelfy	51.6	⋆67.0	63.5	**66.4**	70.3	68.9	50.7	73.2	79.8	66.4
MFS	54.5	65.6	⋆66.0	63.8	⋆67.1	67.7	49.8	73.1	80.5	65.5

Raganato et al. 2017

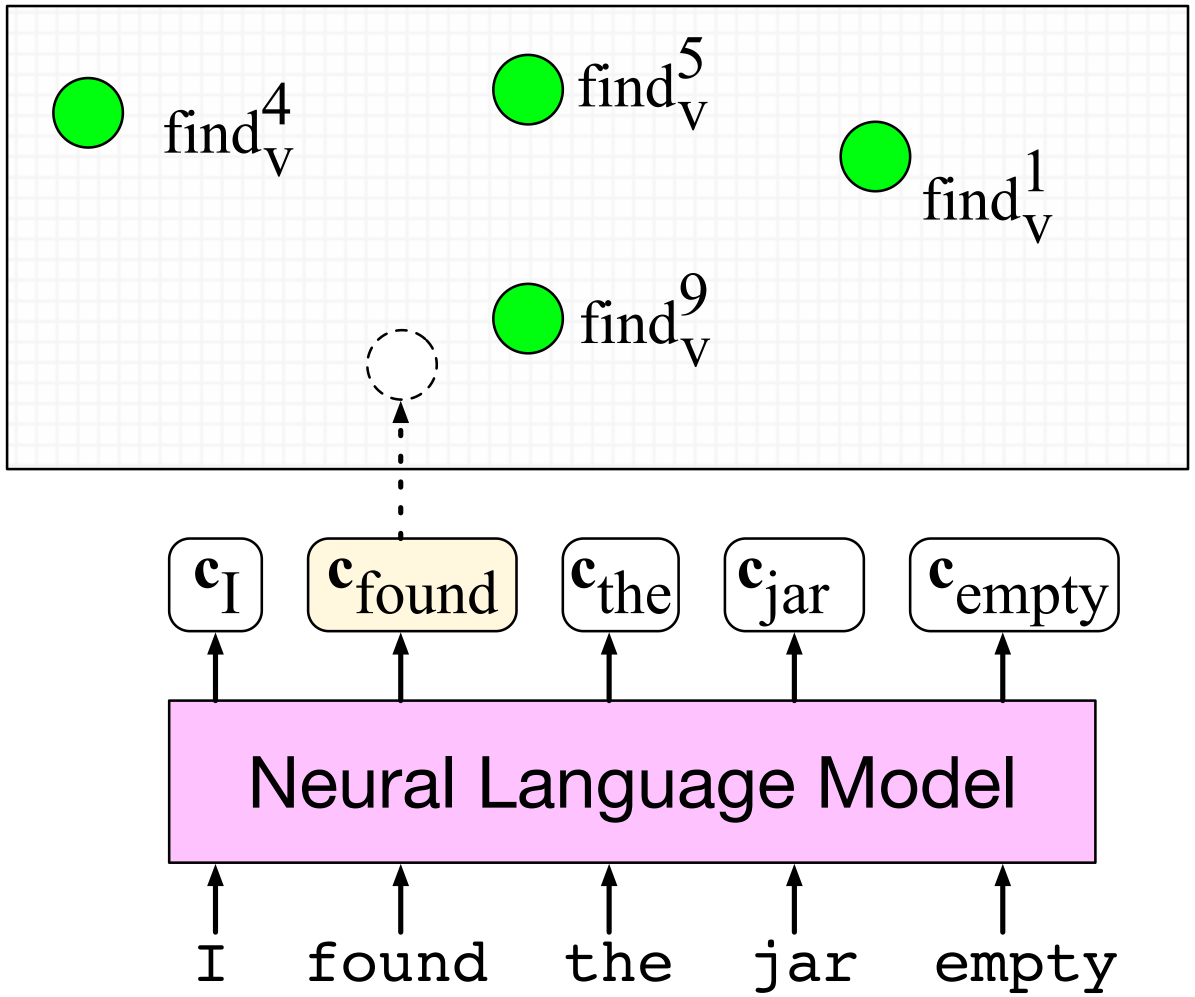
WSD with Neural Language Models

Training: Use contextual embeddings on sense-tagged data to learn what each sense “looks like”



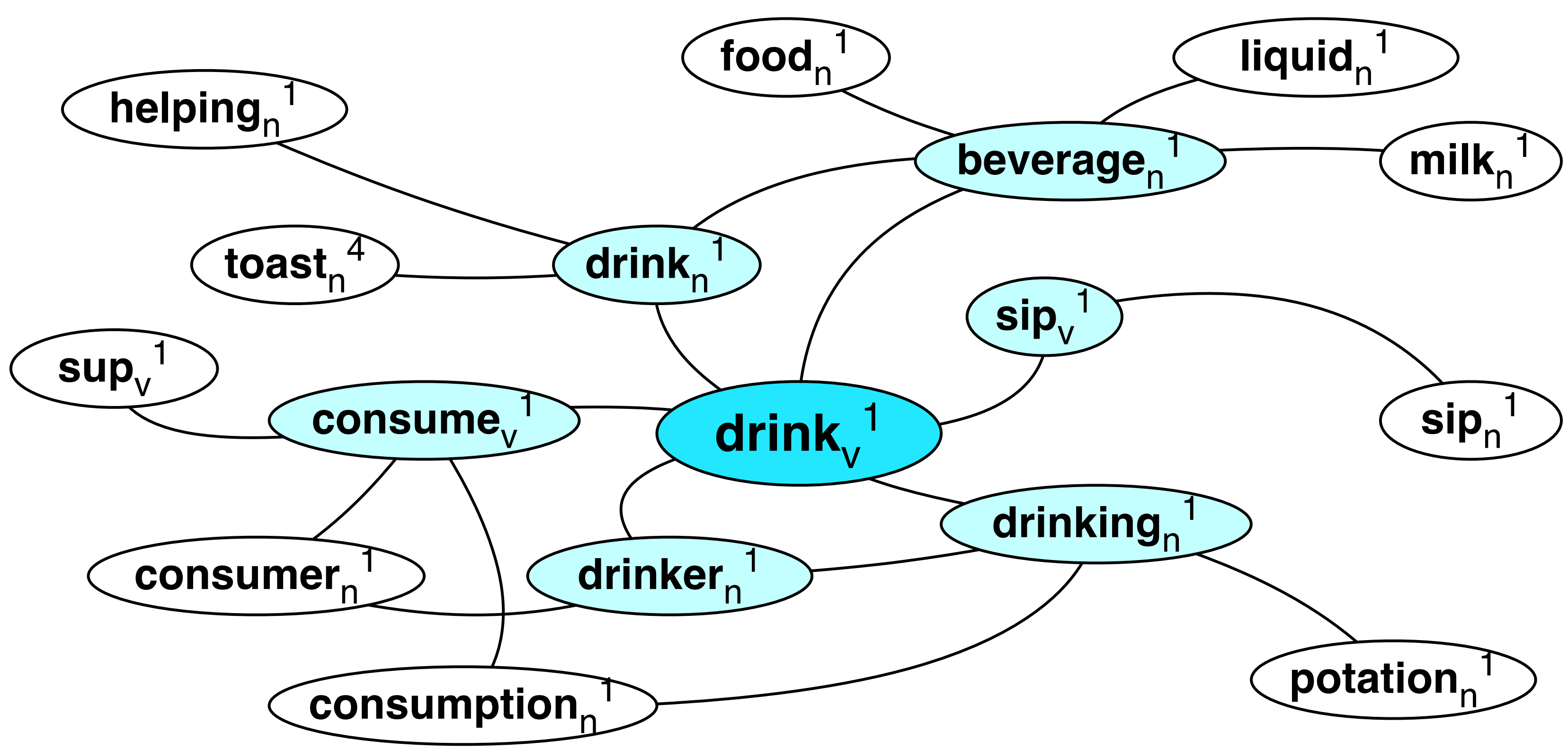
WSD with Neural Language Models

During testing, generate a contextual embedding for the target word and choose the nearest neighbor of the sense-tagged contextual embeddings from training



Graph-based WSD methods

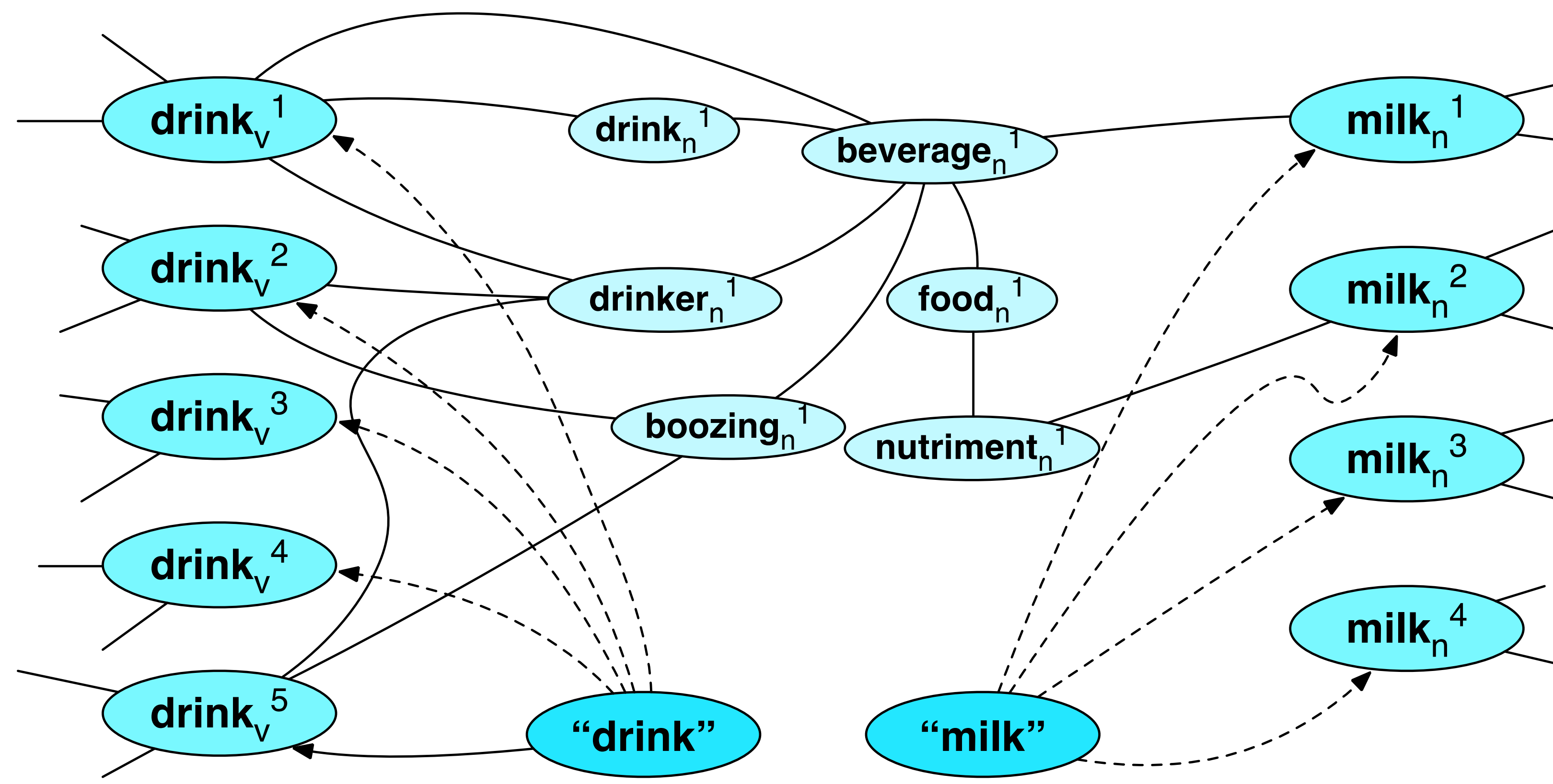
- WordNet can be viewed as a knowledge graph
  - senses are nodes
  - relations (hypernymy, meronymy) are edges
  - Also add edge between word and unambiguous gloss words



How to use the graph for WSD

- Insert target word and words in its sentential context into the graph, with directed edges to their senses

- “She drank some milk”
- Now choose the most central sense
- Add some probability to “drink” and “milk” and compute node with highest PageRank



One sense per discourse

- If a word appears multiple times in a document, it's usually with the same sense. (Gale et al. 1992)
    - Articles about financial banks don't use talk about river banks.

Supervised learning

But WSD is not POS tagging! Why?

Semi-supervised WSD

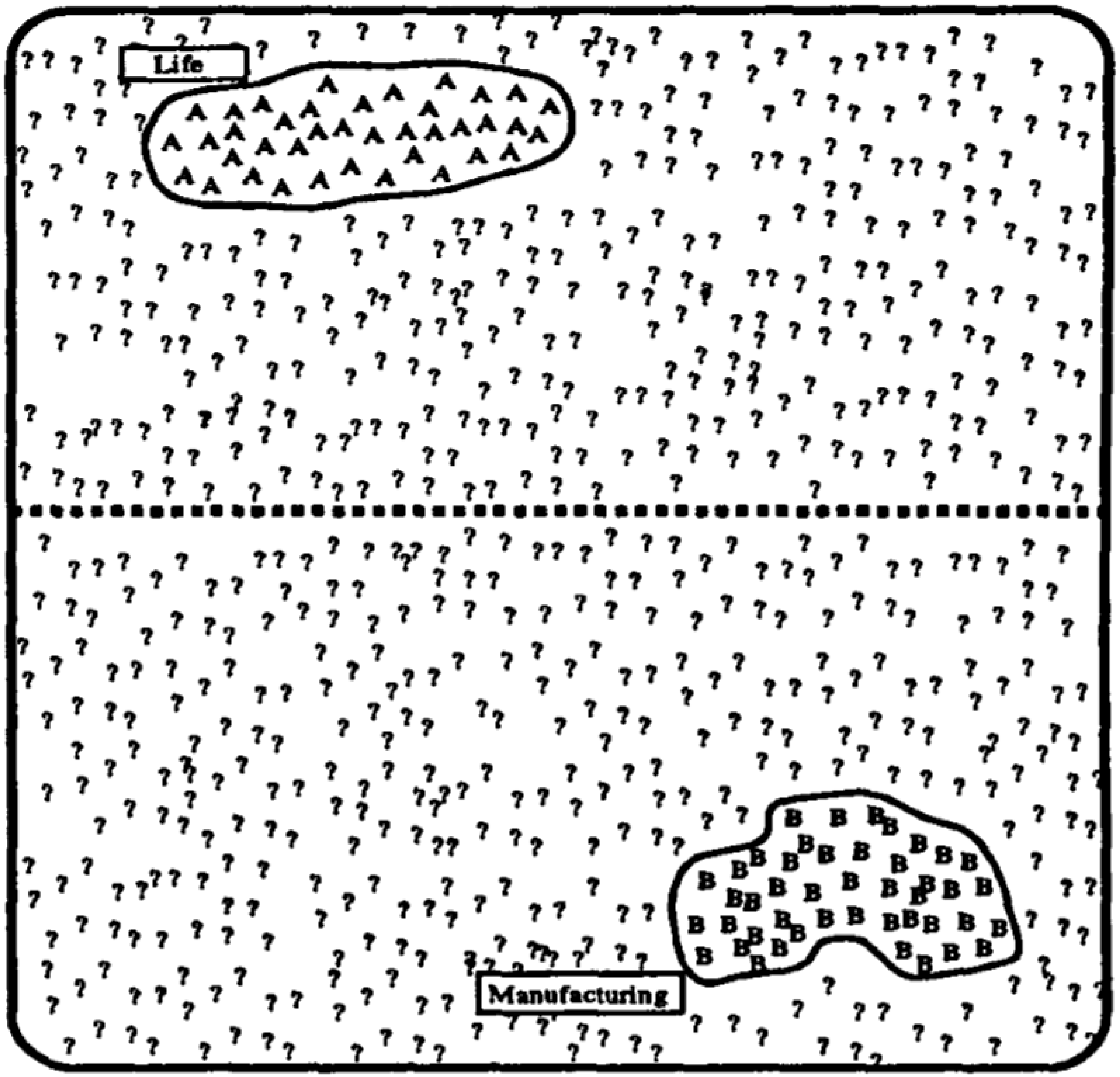
1. Produce seeds (dictionary definitions, single defining collocate, or label common collocates)

2. Repeat until convergence:

    1. Train supervised classifier on labeled examples

    2. Label all examples, and keep labels for high-confidence instances

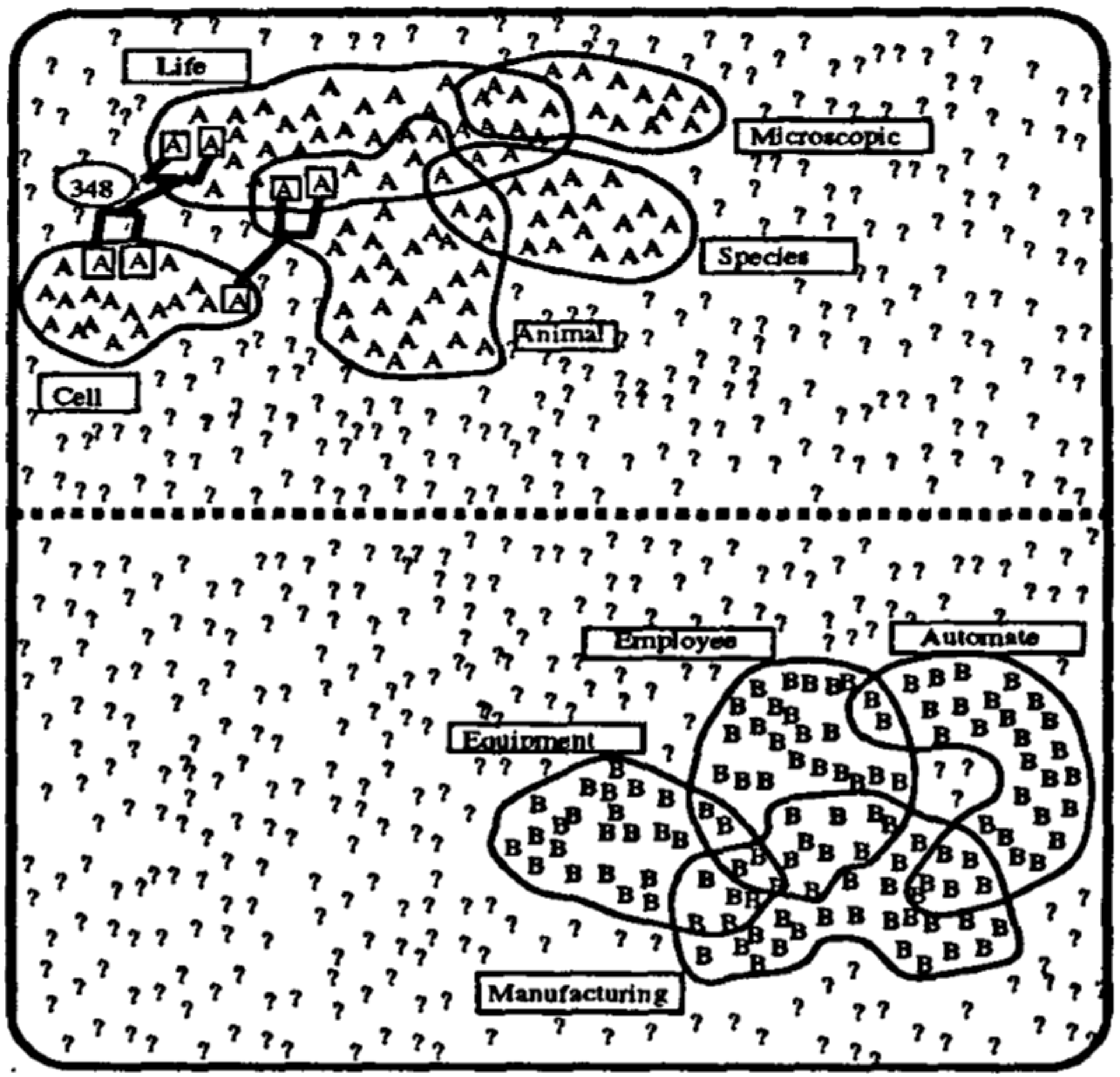
Semi-supervised WSD



"Plant"

A = SENSE-A training example
B = SENSE-B training example
? = currently unclassified training example
$\boxed{\text{Life}}$ = Set of training examples containing the collocation "life".

Semi-supervised WSD



"Plant"

A = SENSE-A training example
B = SENSE-B training example
? = currently unclassified training example
$\boxed{\text{Life}}$ = Set of training examples containing the collocation "life".

Evaluation

- Annotated data; cross-validation.
  - Semcor
  - Ontonotes
- Semeval/Senseval competitions

Hyponymy

Hyponymy

Horse

Hyponymy

Equine

Horse

Hyponymy

Ungulate

Equine

Horse

Hyponymy

Mammal

Ungulate

Equine

Horse

Hyponymy

Vertebrate

Mammal

Ungulate

Equine

Horse

Hyponymy

Animal

Vertebrate

Mammal

Ungulate

Equine

Horse

Hyponymy

… Entity

Animal

Vertebrate

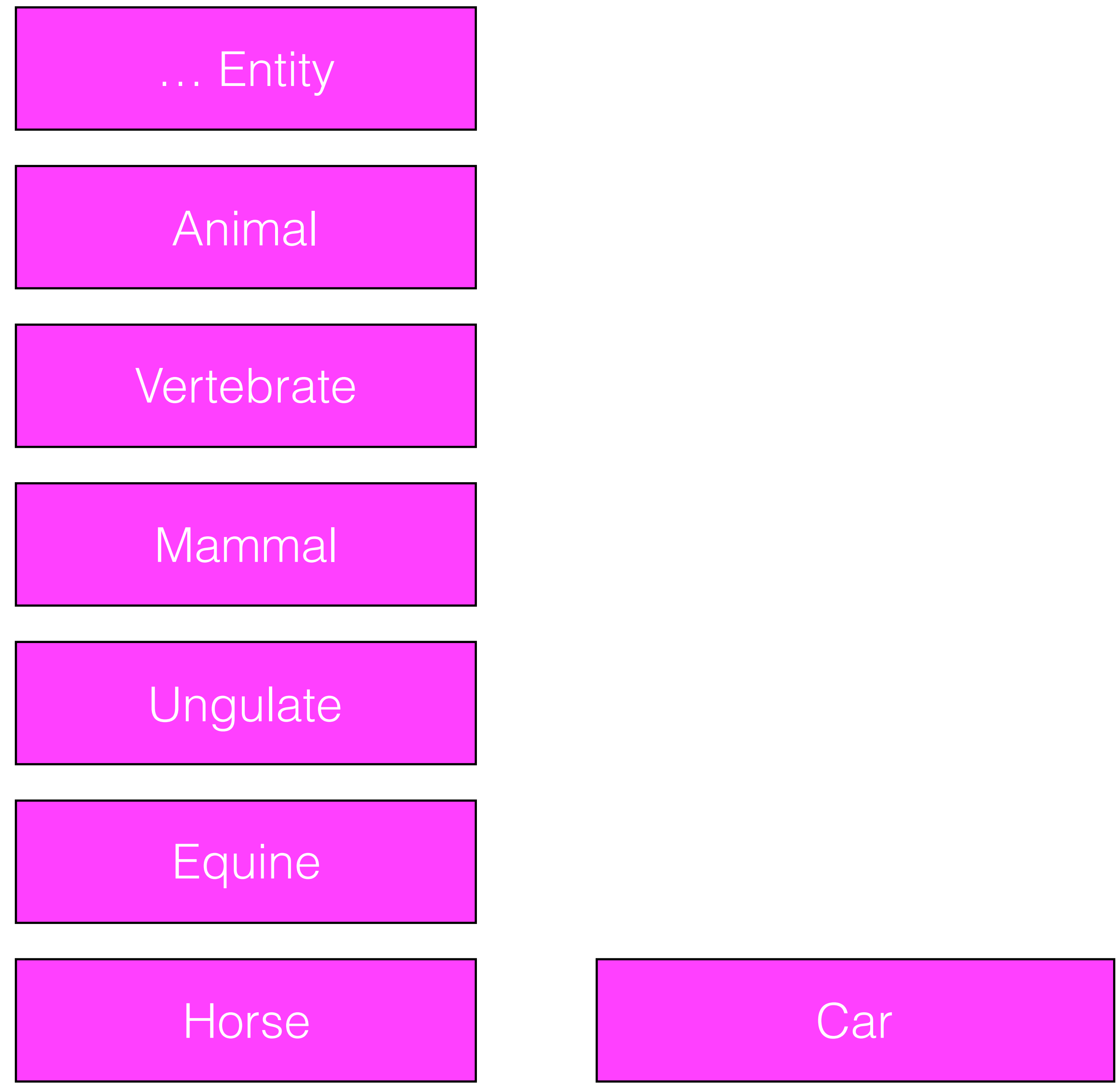
Mammal

Ungulate

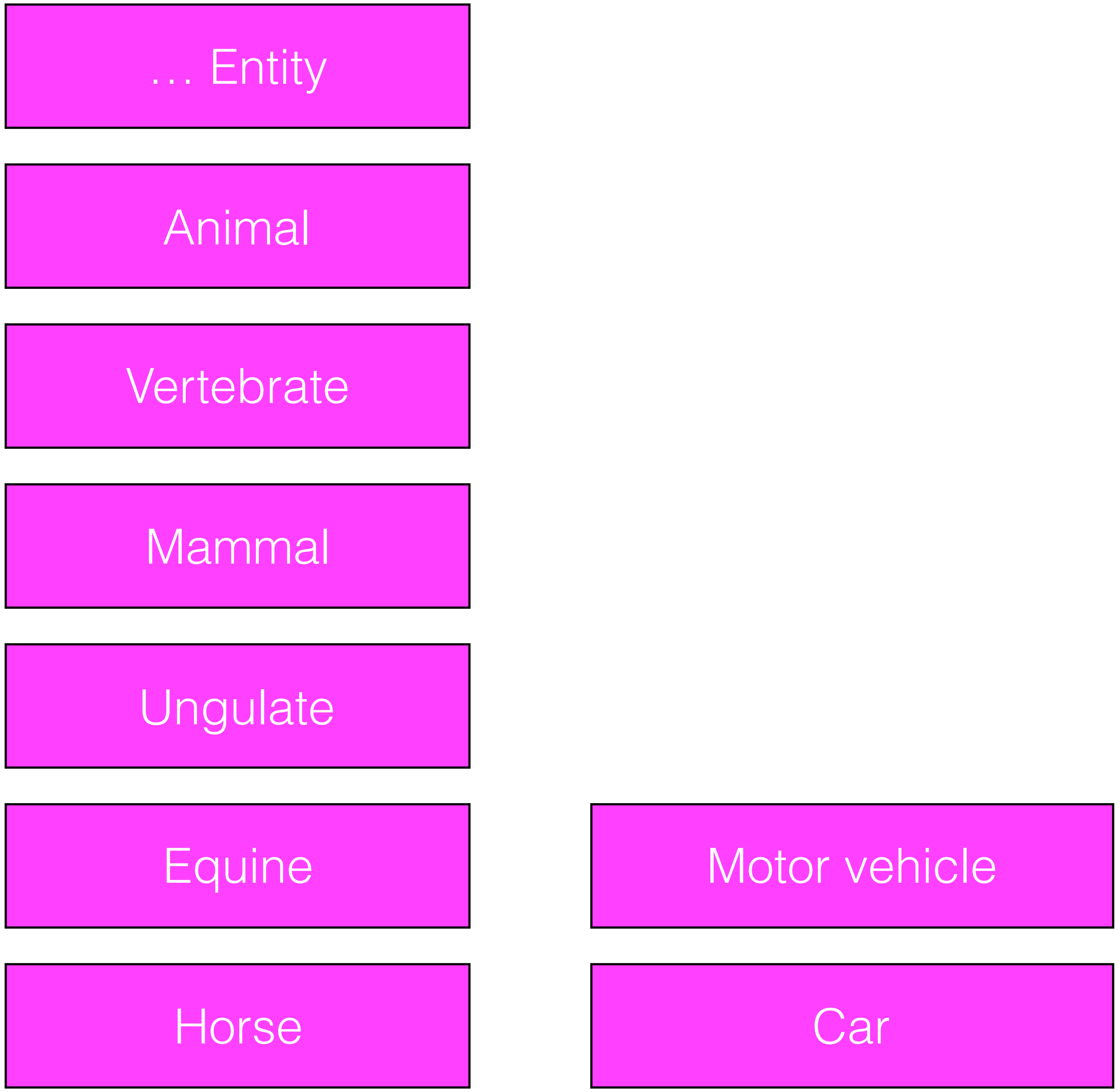
Equine

Horse

Hyponymy


… Entity
Animal
Vertebrate
Mammal
Ungulate
Equine
Horse
Car

Hyponymy


… Entity
Animal
Vertebrate
Mammal
Ungulate
Equine
Horse
Motor vehicle
Car

Hyponymy

… Entity	
Animal	
Vertebrate	
Mammal	
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

Hyponymy

… Entity	
Animal	
Vertebrate	
Mammal	Wheeled vehicle
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

Hyponymy

… Entity	
Animal	
Vertebrate	Conveyance
Mammal	Wheeled vehicle
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

Hyponymy

… Entity

Animal

Vertebrate

Mammal

Ungulate

Equine

Horse

Instrumentality

Conveyance

Wheeled vehicle

Self-propelled vehicle

Motor vehicle

Car

Hyponymy

… Entity	Artifact
Animal	Instrumentality
Vertebrate	Conveyance
Mammal	Wheeled vehicle
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

Hyponymy

	… Entity
… Entity	Artifact
Animal	Instrumentality
Vertebrate	Conveyance
Mammal	Wheeled vehicle
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

NOUNS			
SUPERSENSE	NOUNS DENOTING	SUPERSENSE	NOUNS DENOTING
act	acts or actions	object	natural objects (not man-made)
animal	animals	quantity	quantities and units of measure
artifact	man-made objects	phenomenon	natural phenomena
attribute	attributes of people and objects	plant	plants
body	body parts	possession	possession and transfer of possession
cognition	cognitive processes and contents	process	natural processes
communication	communicative processes and contents	person	people
event	natural events	relation	relations between people or things or ideas
feeling	feelings and emotions	shape	two and three dimensional shapes
food	foods and drinks	state	stable states of affairs
group	groupings of people or objects	substance	substances
location	spatial position	time	time and temporal relations
motive	goals	Tops	abstract terms for unique beginners
VERBS			
SUPERSENSE	VERBS OF	SUPERSENSE	VERBS OF
body	grooming, dressing and bodily care	emotion	feeling
change	size, temperature change, intensifying	motion	walking, flying, swimming
cognition	thinking, judging, analyzing, doubting	perception	seeing, hearing, feeling
communication	telling, asking, ordering, singing	possession	buying, selling, owning
competition	fighting, athletic activities	social	political and social activities and events
consumption	eating and drinking	stative	being, having, spatial relations
contact	touching, hitting, tying, digging	weather	raining, snowing, thawing, thundering
creation	sewing, baking, painting, performing		

Ciarmita and Altun 2003

Supersense tagging

artifact artifact motion time group

The station wagons arrived at noon, a long shining line

motion location location

that coursed through the west campus.

Noun supersenses (Ciarmita and Altun 2003)

Supersense tagging

- Ciarameta and Altun (2006). Trained on data from Semcor (Miller et al. 1993); Brown corpus annotated with WordNet synset labels

- Token-level predictor – each instance of a word has its own supersense tag.

- Maximum-entropy Markov Model (MEMM) trained with averaged perceptron. Features for: word token identity, part-of-speech tag, word shape, previous label + supersense for most frequent synset for word.

- In-domain accuracy: 77.1 F score (cf. 66 F MFS baseline)

Data

- Semcor: 200K+ words tagged with Wordnet senses.
  http://www.cse.unt.edu/~rada/downloads.html#semcor
- WordNet
  https://wordnet.princeton.edu/wordnet/download/
- CROWN
  https://github.com/davidjurgens/crown

https://dofo.org/visit/housebuilding/

Frame Semantics

What do you know about this situation?

The travelers spent a few hours on ________

What do you know about this situation?

The travelers spent a few hours on land________

What do you know about this situation?

The travelers spent a few hours on land______

the ground

What do you know about this situation?

The travelers spent a few hours on land______

the ground

Frame semantics aim to capture this type of background information that readers naturally infer

Frame semantics

Frame semantics

- A semantic frame is a coherent structure of related concepts

Frame semantics

- A semantic frame is a coherent structure of related concepts

- Alternative view: a frame is a data-structure representing a stereotyped situation (Minsky, 1975).

Frame semantics

- A semantic frame is a coherent structure of related concepts
- Alternative view: a frame is a data-structure representing a stereotyped situation (Minsky, 1975).
- Assumes the meaning of a word cannot be understood without access to all the knowledge that relates to that word.

Sally gave Bill $100 for (a) _________

Sally gave Bill $100 for (a) tip __________

ransom

allowance

Sally gave Bill $100 for (a) tip ________

ransom

allowance

refund

Sally gave Bill $100 for (a) tip ________

ransom

allowance

refund

honorarium

Sally gave Bill $100 for (a) tip________

ransom

allowance

refund

honorarium

bounty

Sally gave Bill $100 for (a) tip ________

ransom

allowance

refund

honorarium

bounty

tuition

Sally gave Bill $100 for (a) tip ___________

ransom

allowance

refund

honorarium

bounty

tuition

retainer

Sally gave Bill $100 for (a) tip ________

ransom

allowance

refund

honorarium

bounty

tuition

retainer

bonus

Sally gave Bill $100 for (a) tip ________

ransom
allowance
refund
honorarium
bounty
tuition
retainer
bonus
rent

Sally gave Bill $100 for (a) tip ________

ransom

allowance

refund

honorarium

bounty

tuition

retainer

bonus

rent

fare

Sally gave Bill $100 for (a) tip

ransom

allowance

refund

honorarium

bounty

tuition

retainer

bonus

rent

fare

child support

Sally gave Bill $100 for (a) tip ________

ransom
allowance
refund
honorarium
bounty
tuition
retainer
bonus
rent
fare
child support
bus money

Sally gave Bill $100 for (a) <u>tip</u>

ransom
allowance
refund
honorarium
bounty
tuition
retainer
bonus
rent
fare
child support
bus money
salary

Sally gave Bill $100 for (a) tip ________

ransom
allowance
refund
honorarium
bounty
tuition
retainer
bonus
rent
fare
child support
bus money
salary
reward

Sally gave Bill $100 for (a) tip ________

ransom
allowance
refund
honorarium
bounty
tuition
retainer
bonus
rent
fare
child support
bus money
salary
reward
alimony

Break out session time

- What are the different people, things, events involved in a “ordering food” frame?

Representation of meaning

Representation of meaning

- Constants name individual entities in the world

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities
- Variables refer to entities that have not yet been specified

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities
- Variables refer to entities that have not yet been specified
- Quantifiers bind variables.

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities
- Variables refer to entities that have not yet been specified
- Quantifiers bind variables.
    - ∃ (existential quantifier)

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities
- Variables refer to entities that have not yet been specified
- Quantifiers bind variables.
    - $\exists$ (existential quantifier)
    - $\forall$ (universal quantifier)

Representation of meaning

Representation of meaning

Pat likes Sal

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

- The denotation ⟦likes⟧ = the ordered set of entities for whom the relation is true

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

- The denotation ⟦likes⟧ = the ordered set of entities for whom the relation is true
- likes(Pat, Sal) = true

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

- The denotation ⟦likes⟧ = the ordered set of entities for whom the relation is true
- likes(Pat, Sal) = true
- ⟦likes⟧ = {(Pat, Sal), (…, …)}

Representation of meaning

Representation of meaning

- Quantifiers bind variables.

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)
  - ∀ (universal quantifier)

Representation of meaning

- Quantifiers bind variables.
    - ∃ (existential quantifier)
    - ∀ (universal quantifier)

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)
  - ∀ (universal quantifier)
- Order matters!

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)
  - ∀ (universal quantifier)
- Order matters!
  - $\forall x \exists y$ speaks(x,y)

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)
  - ∀ (universal quantifier)

- Order matters!
  - $\forall x \exists y$ speaks(x,y)
  - $\exists y \forall x$ speaks(x,y)

Representation of meaning

- Relations: likes(x,y) is scoped over two variables
- We can represent the partial representation of meaning with lambda expressions:

$\lambda x.likes(x,Sal)$

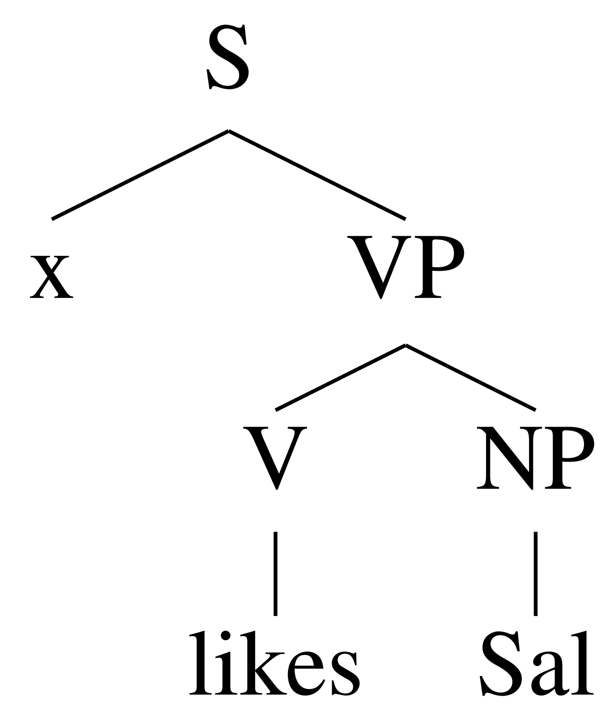
Expect one other argument to complete the meaning of this relation

Church (1940)

Representation of meaning

λx.likes(x,Sal)

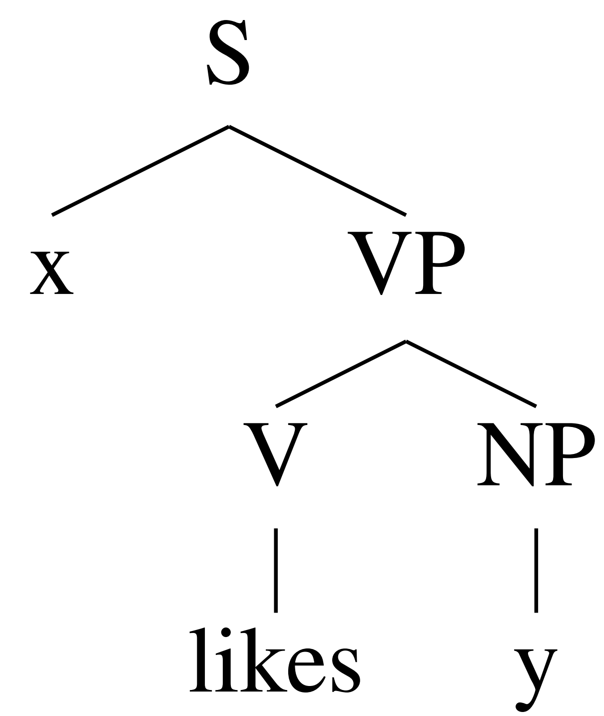
Lambda expressions
let us tie semantics
explicitly to phrases
(subtrees in syntax)



Representation of meaning

λy.λx.likes(x,y)

Lambda expressions let us tie semantics explicitly to phrases (subtrees in syntax)



Event semantics

Pat gives Sal a book

Eisenstein 2017

Event semantics

Pat gives Sal a book

$\exists x.\text{book}(x) \wedge \textsc{give}(\text{Pat},\text{Sal},x)$

Eisenstein 2017

Representation of meaning

- Relations: likes(x,y) is scoped over two variables
- We can represent the partial representation of meaning with lambda expressions:

λx.likes(x,Sal)

Expect one other argument to complete the meaning of this relation

Church (1940)

Event semantics

Yesterday, Pat gives Sal a book reluctantly

Eisenstein 2017

Event semantics

Yesterday, Pat gives Sal a book reluctantly

- One option: extend the arity of the relation (require more arguments)

Eisenstein 2017

Event semantics

Yesterday, Pat gives Sal a book reluctantly

$\exists x.\text{book}(x) \wedge \textsc{give}(\text{Pat}, \text{Sal}, x, \text{yesterday}, \text{reluctantly})$

- One option: extend the arity of the relation (require more arguments)

Eisenstein 2017

Event semantics

Yesterday, Pat gives Sal a book reluctantly

$\exists x.\text{book}(x) \wedge \textsc{give}(\text{Pat}, \text{Sal}, x, \text{yesterday}, \text{reluctantly})$

- One option: extend the arity of the relation (require more arguments)
- But that's not great because we need a separate predicate for every possible combination of arguments (even those that aren't required).

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists e, x.\text{GIVE-EVENT}(e)$

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists e,x.\text{GIVE-EVENT}(e)$
$\wedge\ \text{GIVER}(e,Pat)$

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists e,x.\text{GIVE-EVENT}(e)$
$\wedge\ \text{GIVER}(e,\text{Pat})$
$\wedge\ \text{GIFT}(e,x)$

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists$e,x.GIVE-EVENT(e)
$\wedge$ GIVER(e,Pat)
$\wedge$ GIFT(e,x)
$\wedge$ BOOK(e,x)

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists e, x.\text{GIVE-EVENT}(e)$
$\wedge\ \text{GIVER}(e, \text{Pat})$
$\wedge\ \text{GIFT}(e, x)$
$\wedge\ \text{BOOK}(e, x)$
$\wedge\ \text{RECIPIENT}(e, \text{Sal})$

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists$e,x.GIVE-EVENT(e)
$\wedge$ GIVER(e,Pat)
$\wedge$ GIFT(e,x)
$\wedge$ BOOK(e,x)
$\wedge$ RECIPIENT(e,Sal)
$\wedge$ TIME(e,yesterday)

Eisenstein 2017

Event semantics

We can reify the event to an existentially quantified variable of its own, and then use it as a argument in other relations.

$\exists$e,x.GIVE-EVENT(e)
$\wedge$ GIVER(e,Pat)
$\wedge$ GIFT(e,x)
$\wedge$ BOOK(e,x)
$\wedge$ RECIPIENT(e,Sal)
$\wedge$ TIME(e,yesterday)
$\wedge$ MANNER(e,reluctanctly)

Eisenstein 2017

Event semantics

Neo-Davidson event semantics: the event is central, and relations are predicated of the event. Each argument of an event holds its own relation.

$\exists$e,x.GIVE-EVENT(e)
$\wedge$ GIVER(e,Pat)
$\wedge$ GIFT(e,x)
$\wedge$ BOOK(e,x)
$\wedge$ RECIPIENT(e,Sal)
$\wedge$ TIME(e,yesterday)
$\wedge$ MANNER(e,reluctanctly)

Eisenstein 2017

Event semantics

Neo-Davidson event semantics: the event is central, and relations are predicated of the event. Each argument of an event holds its own relation.

$\exists$e,x.GIVE-EVENT(e)
∧ GIVER(e,Pat)
∧ GIFT(e,x)
∧ BOOK(e,x)
∧ RECIPIENT(e,Sal)
∧ TIME(e,yesterday)
∧ MANNER(e,reluctanctly)

In model-theoretic semantics, each of these has a denotation in a world model

Eisenstein 2017

Event semantics

Sasha broke the window

SLP3

Event semantics

$\exists e,y.$BREAKING-EVENT(e)

Sasha broke the window

SLP3

Event semantics

Sasha broke the window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$

SLP3

Event semantics

Sasha broke the window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$

SLP3

Event semantics

Sasha broke the window

$\exists$e,y.BREAKING-EVENT(e)
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

SLP3

Event semantics

Sasha broke the window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

Pat opened the door

SLP3

Event semantics

Sasha broke the window

$\exists$e,y.BREAKING-EVENT(e)
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

Pat opened the door

$\exists$e,y.OPENING-EVENT(e)

SLP3

Event semantics

Sasha broke the window

$\exists e,y.\mathrm{BREAKING\text{-}EVENT}(e)$
$\wedge\ \mathrm{BREAKER}(e,\mathrm{Sasha})$
$\wedge\ \mathrm{BROKEN\text{-}THING}(e,y)$
$\wedge\ \mathrm{WINDOW}(e,y)$

Pat opened the door

$\exists e,y.\mathrm{OPENING\text{-}EVENT}(e)$
$\wedge\ \mathrm{OPENER}(e,\mathrm{Pat})$

Event semantics

Sasha broke the window

$\exists$e,y.BREAKING-EVENT(e)
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

Pat opened the door

$\exists$e,y.OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)

SLP3

Event semantics

Sasha broke the window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

Pat opened the door

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

Event semantics

In model-theoretic semantics, each of these has some denotation in the world model.

Example: WINDOW has a identifier in some knowledge base (e.g., Freebase) uniquely identifying its properties.

$\exists$e,y.BREAKING-EVENT(e)
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists$e,y.OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

Event semantics

This requires a comprehensive representation of the world

$\exists e,y.$BREAKING-EVENT(e)
$\wedge$ BREAKER$(e,Sasha)$
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.$OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

Shallow semantics

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

Shallow semantics

$\exists e,y.$BREAKING-EVENT(e)
$\wedge$ BREAKER$(e,Sasha)$
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.$OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Sasha})$

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge$ CAUSER-OF-ACTION(e,Sasha)
$\wedge$ RECIPIENT-OF-ACTION(e,y)

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge$ BREAKER(e,Sasha)
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Sasha})$
$\wedge\ \text{RECIPIENT-OF-ACTION}(e,y)$
$\wedge$ “window”(y)

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

∃e,y.EVENT(e)
∧ CAUSER-OF-ACTION(e,Sasha)
∧ RECIPIENT-OF-ACTION(e,y)
∧ “window”(y)

∃e,y.BREAKING-EVENT(e)
∧ BREAKER(e,Sasha)
∧ BROKEN-THING(e,y)
∧ WINDOW(e,y)

∃e,y.EVENT(e)

∃e,y.OPENING-EVENT(e)
∧ OPENER(e,Pat)
∧ OPENED-THING(e,y)
∧ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Sasha})$
$\wedge\ \text{RECIPIENT-OF-ACTION}(e,y)$
$\wedge$ "window"(y)

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Pat})$

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Sasha})$
$\wedge\ \text{RECIPIENT-OF-ACTION}(e,y)$
$\wedge$ “window”(y)

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{CAUSER-OF-ACTION}(e,\text{Pat})$
$\wedge\ \text{RECIPIENT-OF-ACTION}(e,y)$

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

∃e,y.EVENT(e)
∧ CAUSER-OF-ACTION(e,Sasha)
∧ RECIPIENT-OF-ACTION(e,y)
∧ “window”(y)

∃e,y.BREAKING-EVENT(e)
∧ BREAKER(e,Sasha)
∧ BROKEN-THING(e,y)
∧ WINDOW(e,y)

∃e,y.EVENT(e)
∧ CAUSER-OF-ACTION(e,Pat)
∧ RECIPIENT-OF-ACTION(e,y)
∧ “door”(y)

∃e,y.OPENING-EVENT(e)
∧ OPENER(e,Pat)
∧ OPENED-THING(e,y)
∧ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Shallow semantics

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{AGENT}(e,\text{Sasha})$
$\wedge\ \text{THEME}(e,y)$
$\wedge$ “window”(y)

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

$\exists e,y.\text{EVENT}(e)$
$\wedge\ \text{AGENT}(e,\text{Pat})$
$\wedge\ \text{THEME}(e,y)$
$\wedge$ “door”(y)

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

Shallow semantics

Agent: Sasha
Theme: window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

Agent: Pat
Theme: door

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

Thematic roles

SLP3

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

SLP3

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

- John broke the window

SLP3

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

- John broke the window
- The window was broken by John

SLP3

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

- John broke the window
- The window was broken by John

SLP3

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

- John broke the window
- The window was broken by John

Different syntactic roles, but the same thematic role.

SLP3

Thematic roles

SLP3

Thematic roles

Agent	The volitional causer of an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

SLP3

Thematic roles

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond*...

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond*...
Content	Mona asked *"You met Mary Ann at a supermarket?"*

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond*...
Content	Mona asked *"You met Mary Ann at a supermarket?"*
Instrument	He poached catfish, stunning them with *a shocking device*...

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond...*
Content	Mona asked *"You met Mary Ann at a supermarket?"*
Instrument	He poached catfish, stunning them with *a shocking device*...
Beneficiary	Whenever Ann makes hotel reservations for *her boss*...

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond*...
Content	Mona asked *"You met Mary Ann at a supermarket?"*
Instrument	He poached catfish, stunning them with *a shocking device*...
Beneficiary	Whenever Ann makes hotel reservations for *her boss*...
Source	I flew in from *Boston*.

SLP3

Thematic roles

Agent	*The waiter* spilled the soup.
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond...*
Content	Mona asked *"You met Mary Ann at a supermarket?"*
Instrument	He poached catfish, stunning them with *a shocking device*...
Beneficiary	Whenever Ann makes hotel reservations for *her boss*...
Source	I flew in from *Boston*.
Goal	I drove to *Portland*.

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John
- John broke the window with a rock

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John
- John broke the window with a rock
- The rock broke the window

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John
- John broke the window with a rock
- The rock broke the window
- The window broke

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John
- John broke the window with a rock
- The rock broke the window
- The window broke

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Doris gave the book to Cary

Doris gave Cary the book

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent

Doris gave the book to Cary

Doris gave Cary the book

SLP3

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent

Theme

Doris gave the book to Cary

Doris gave Cary the book

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent Theme Goal

Doris gave the book to Cary

Doris gave Cary the book

SLP3

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent Theme Goal

Doris gave the book to Cary

Agent

Doris gave Cary the book

SLP3

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent Theme Goal

Doris gave the book to Cary

Agent Goal

Doris gave Cary the book

SLP3

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent Theme Goal

Doris gave the book to Cary

Agent Goal Theme

Doris gave Cary the book

SLP3

Thematic roles

- Thematic roles are very useful but different to formally define AGENT, THEME, etc.
- At the same time, they may be too coarse for some applications.

Thematic roles

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

- The cook opened the jar with the new gadget

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

- The cook opened the jar with the new gadget
- The new gadget opened the jar

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

- The cook opened the jar with the new gadget
- The new gadget opened the jar
- Shelly ate the sliced banana with a fork

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

- The cook opened the jar with the new gadget
- The new gadget opened the jar

- Shelly ate the sliced banana with a fork
- *The fork ate the sliced banana

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

Intermediary instruments can be subjects

- The cook opened the jar with the new gadget
- The new gadget opened the jar

- Shelly ate the sliced banana with a fork
- *The fork ate the sliced banana

Levin and Rappaport Hovav 2005; SLP3

Thematic roles

Intermediary instruments can be subjects

- The cook opened the jar with the new gadget
- The new gadget opened the jar

- Shelly ate the sliced banana with a fork
- *The fork ate the sliced banana

Enabling instruments cannot

Levin and Rappaport Hovav 2005; SLP3

Coarsening: Proto-roles

- Proto-roles = generalized thematic roles
- Proto-agent: causing an event, having volition wrt event, moving, acting with intention
- Proto-patient: change of state, causally affected by event)

Propbank

- Sentences from the Penn Treebank annotated with proto-roles, along with lexical entries for each sense of a verb identifying the specific meaning of each proto-role for that verb sense.

https://propbank.github.io

Propbank

(22.11) **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].

(22.12) **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%].

Propbank

- Verb-specific argument structures lets us map the commonalities among the different surface forms
    - [$_{\text{Arg0}}$ 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%].

Propbank

- Verb-specific argument structures lets us map the commonalities among the different surface forms
  - [$_{Arg0}$ Big Fruit Co.] increased [$_{Arg1}$ the price of bananas].
  - [$_{Arg1}$ The price of bananas] was increased again [$_{Arg0}$ by Big Fruit Co.]
  - [$_{Arg1}$ The price of bananas] increased [$_{Arg2}$ 5%].

- [$_{\text{Arg1}}$ 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

- Propbank maps argument structure for individual verb senses
- FrameNet maps argument structure for frames, which are evoked by a lexical unit (typically a verb)

https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

Frames

AI

- Schank and Abelson 1975, 1977
- Minksky 1974

Linguistics

- Fillmore 1975, 1982, Tannen 1979

Cognitive Psychology

- Rumelhart 1975, 1980

Sociology

- Goffman 1975

Media Studies

- Entman 1993

Frames

John went into a restaurant. He ordered a hamburger and coke. He asked the waitress for the check and left.

(Schank & Abelson 75)

Frames

- “A frame is a data-structure for representing a stereotyped situation” (Minsky 1975)
- By the term ‘frame’ I have in mind any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structured is introduced … all of the others are automatically made available.’’ (Fillmore 1982)

Who did what to whom?

Who did what to whom?

- John bought the car at the dealership

Who did what to whom?

- John bought the car at the dealership
- The car was bought by John

Who did what to whom?

- John bought the car at the dealership
- The car was bought by John
- John’s purchase of the car

Who did what to whom?

- John bought the car at the dealership
- The car was bought by John
- John’s purchase of the car
- The sale of the car cleared their inventory.

Who did what to whom?

commercial_transaction

- John bought the car at the dealership
- The car was bought by John
- John's purchase of the car
- The sale of the car cleared their inventory.

Who did what to whom?

Buyer

- John bought the car at the dealership
- The car was bought by John
- John's purchase of the car
- The sale of the car cleared their inventory.

Who did what to whom?

Thing bought

- John bought the car at the dealership
- The car was bought by John
- John's purchase of the car
- The sale of the car cleared their inventory.

Semantic Frame

APPLY_HEAT

- Lexical units:

bake.v, barbecue.v, blanch.v, boil.v, braise.v, broil.v, brown.v, char.v, coddle.v, cook.v, deep fry.v, fry.v, grill.v, microwave.v, parboil.v, plank.v, poach.v, roast.v, saute.v, scald.v, sear.v, simmer.v, singe.v, steam.v, steep.v, stew.v, toast.v

- Core Frame Elements:

Cook	The Cook applies heat to the Food.
Food	Food is the entity to which heat is applied by the Cook.
Heating instrument	The entity that directly supplies heat to the Foo
Container	The Container holds the Food to which heat is applied.
Temperature setting	The Temperature_setting of the Heating_instrument for the Food.

Semantic Frame

DESTROY

- Lexical units:

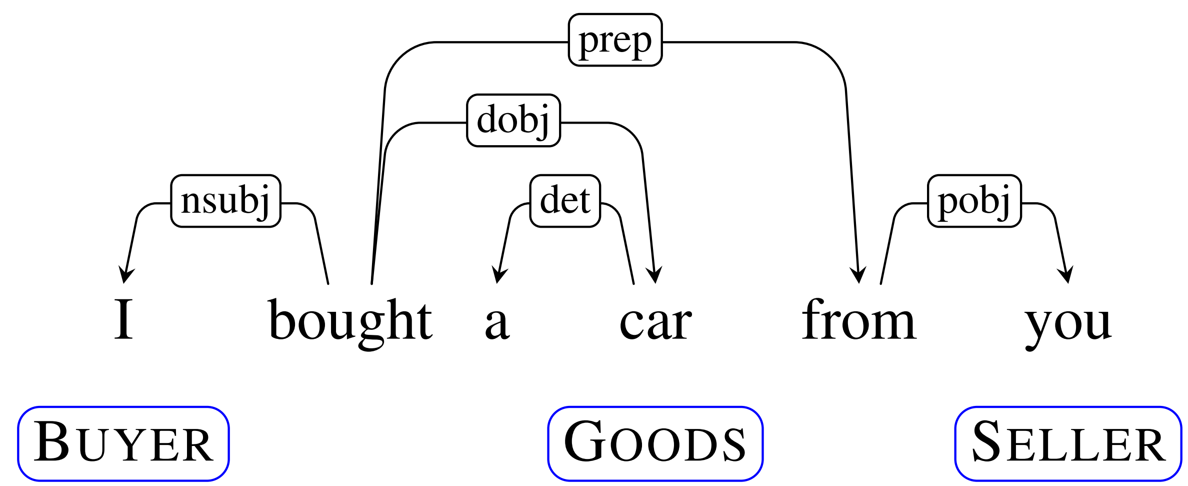
annihilate.v, annihilation.n, blast.v, blow up.v, demolish.v, demolition.n, destroy.v, destruction.n, destructive.a, devastate.v, devastation.n, dismantle.v, dismantlement.n, lay waste.v, level.v, obliterate.v, obliteration.n, raze.v, ruin.v, take out.v, unmake.v, vaporize.v

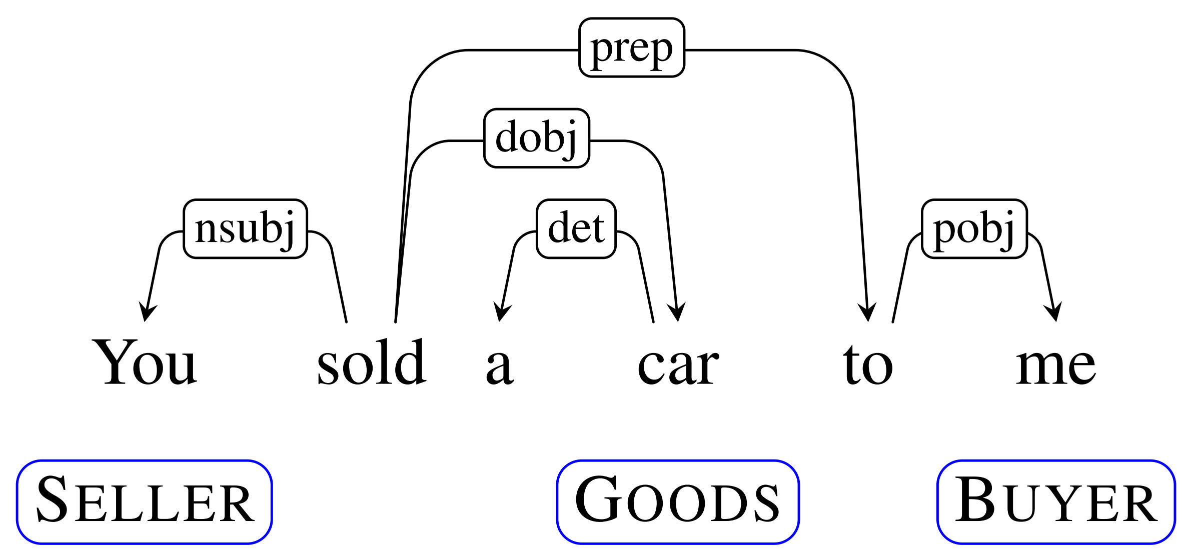
- Core Frame Elements:

Cause	The event or entity which is responsible for the destruction of the Patient.
Destroyer	The conscious entity, generally a person, that performs the intentional action that results in the Patient's destruction.
Patient	The entity which is destroyed by the Destroyer.

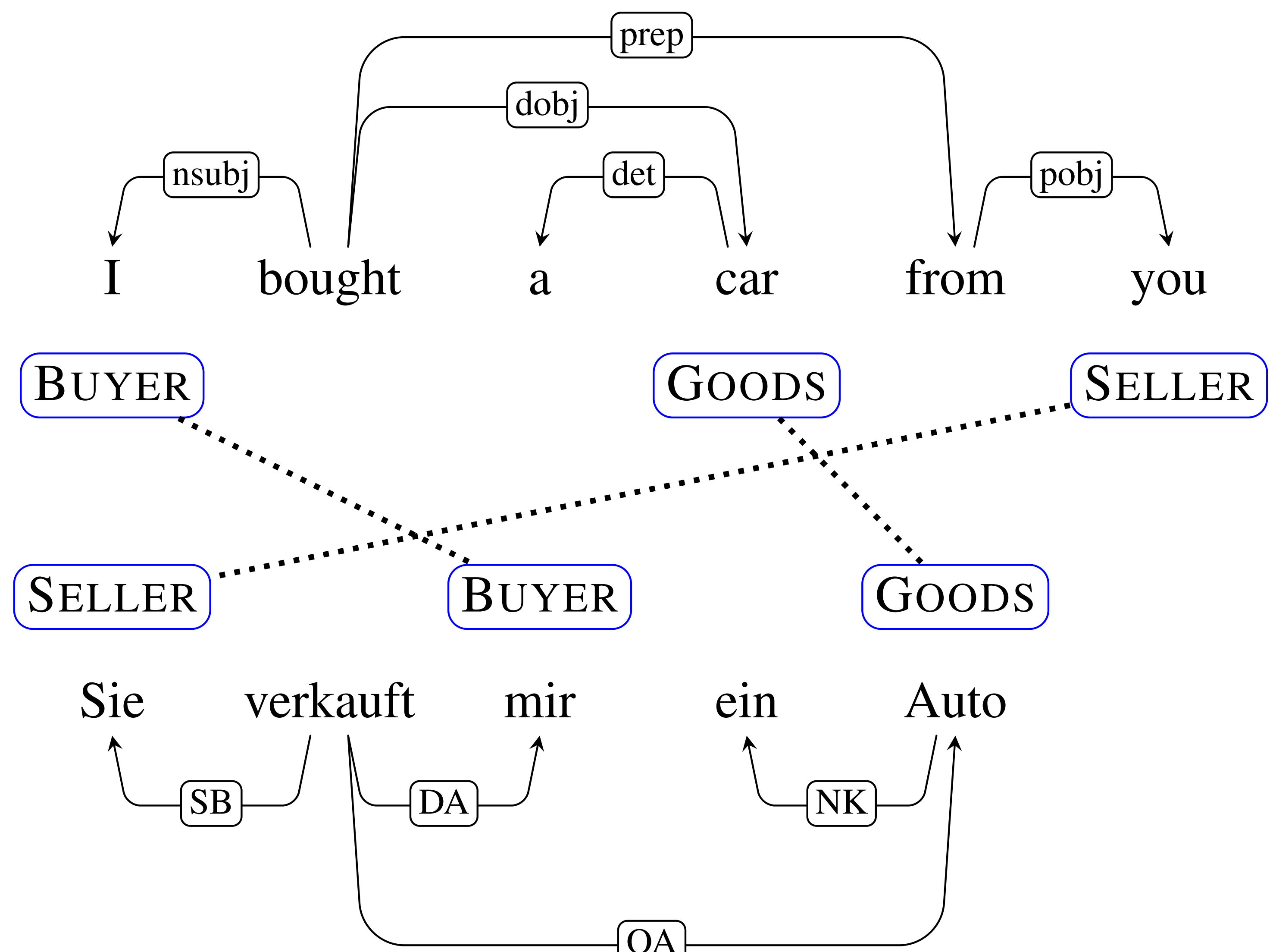
Semantic representations

Two different perspectives on a commercial transaction





Multilingual frames



Multilingual frames

- French
- Chinese
- Brazilian Portuguese
- German
- Spanish
- Japanese
- Swedish
- Korean

https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages

Semantic role labeling

Semantic role labeling

- Input: a sentence

Semantic role labeling

- Input: a sentence
- Output:

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)
    - a span

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)
    - a span
    - a set of arguments, each containing:

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)
    - a span
    - a set of arguments, each containing:
      - a label (thematic role, FrameNet role)

Smith 2017

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)
    - a span
    - a set of arguments, each containing:
      - a label (thematic role, FrameNet role)
      - a span

Smith 2017

Semantic role labeling

FrameNet	[You]	can't	[blame]	[the program]	[for being unable to identify it]
	COGNIZER		TARGET	EVALUEE	REASON
PropBank	[The San Francisco Examiner]	issued	[a special edition]	[yesterday]	
	ARG0	TARGET	ARG1	ARGM-TMP	

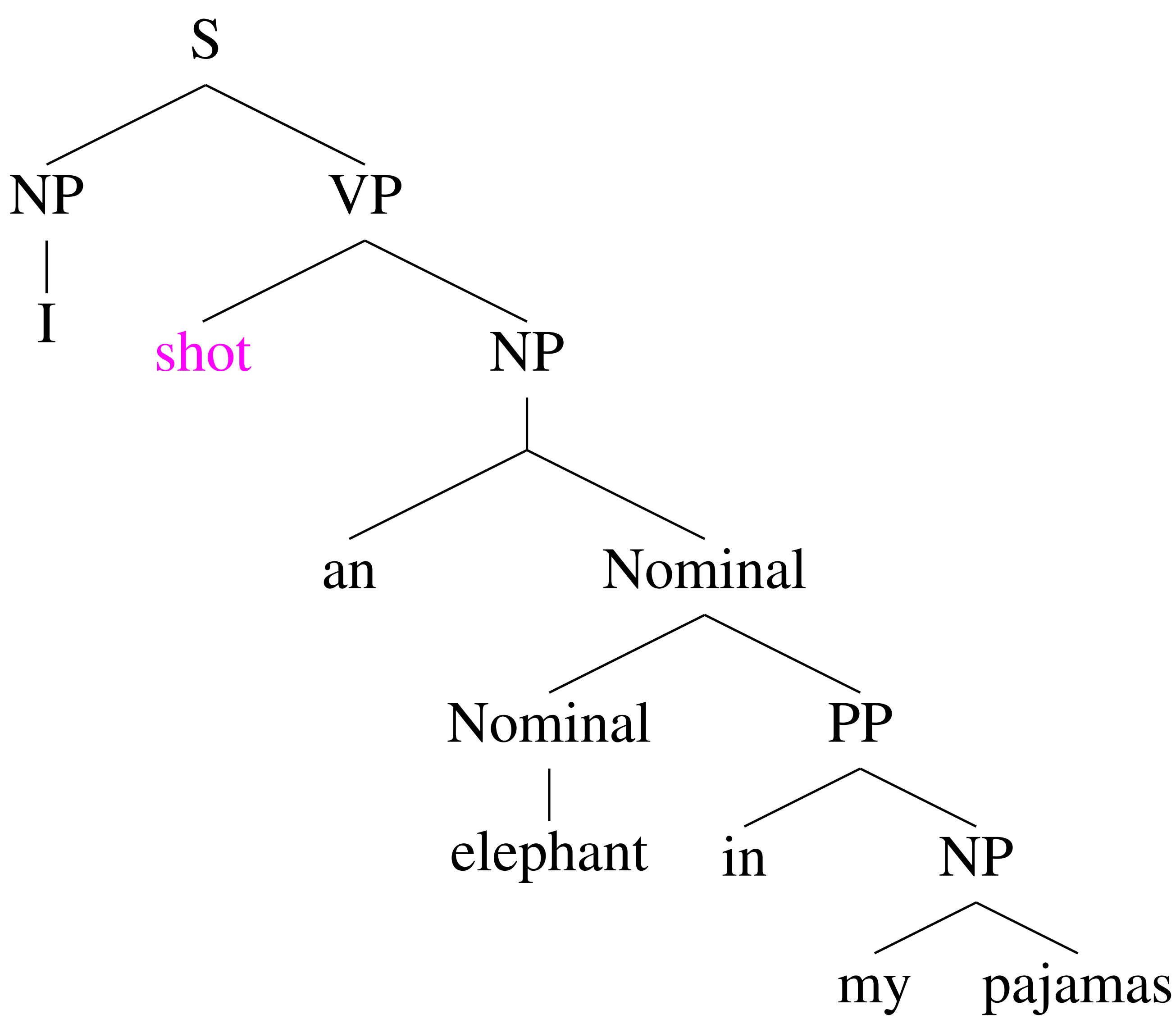
Semantic role labeling

```
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)
```


Gildea and Jurafsky 2002; SLP3

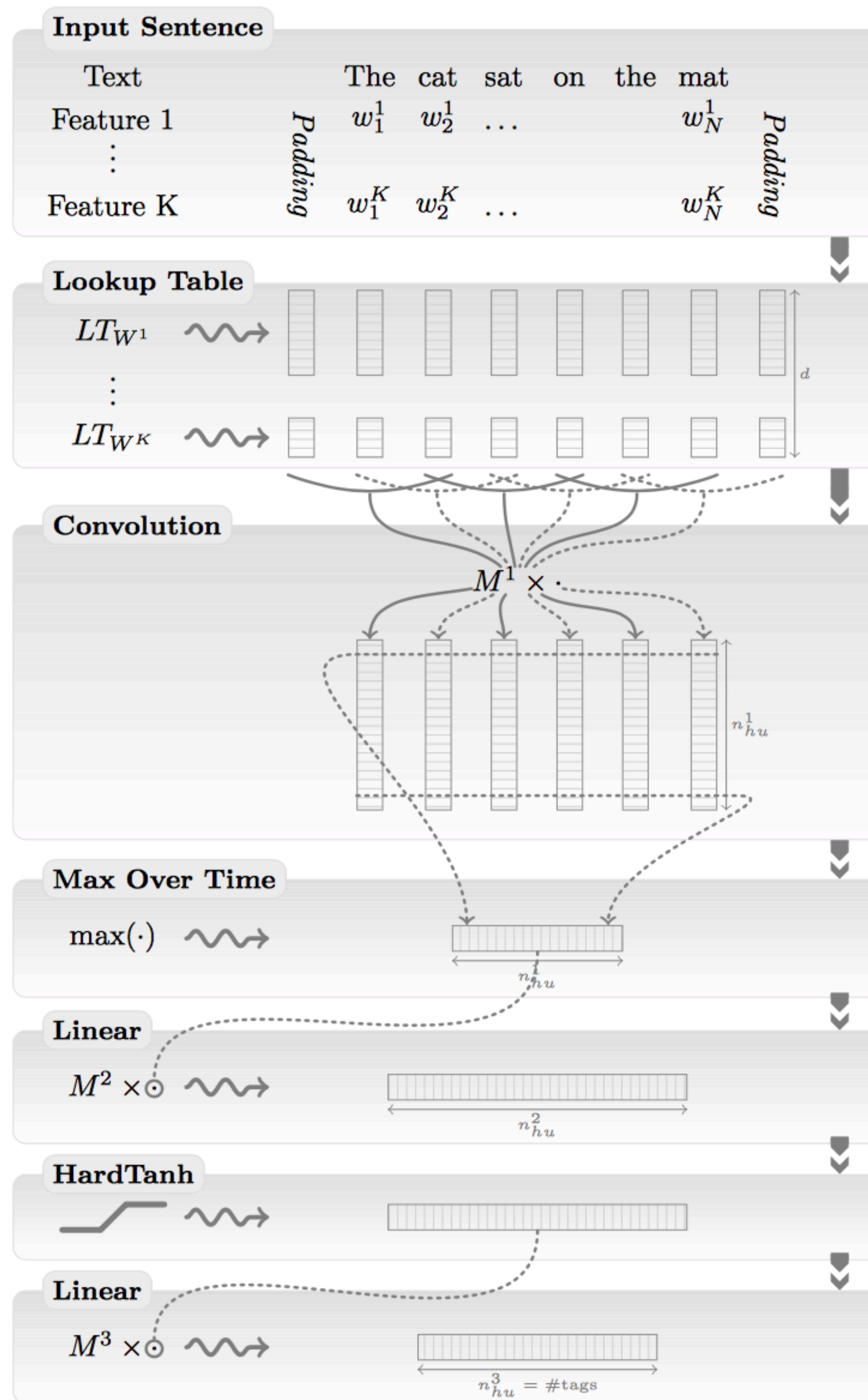
Semantic role labeling



feature
predicate: shot
phrase type = NP
headword of phrase = elephant
path = NP↑S↓VP
voice of verb = active
voice of verb = passive
phrase before verb?
first/last words of phrase

Semantic role labeling

Collobert et al. (2011), Natural Language Processing (Almost) from Scratch



Semantic role labeling

- Sentence-level constraints:
  - Arguments can't overlap
  - For a given predicate, typically only one argument of each type (e.g., ARG0, BUYER)
- Approximate joint decoding (Das et al. 2010)
- Constrained optimization (e.g., ILP)

Smith 2017

Data

- CCGBank [through UCB Library]
  http://groups.inf.ed.ac.uk/ccg/ccgbank.html

- PropBank
  https://propbank.github.io

- FrameNet
  https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

https://www.worldtravelguide.net/guides/europe/greece/

Semantic Parsing

Semantic Parsing

- “Semantic Parsing” is, ironically, a semantically ambiguous term
    - Semantic role labeling
    - Finding generic relations in text
    - Transforming a natural language sentence into its meaning representation

Semantic Parsing

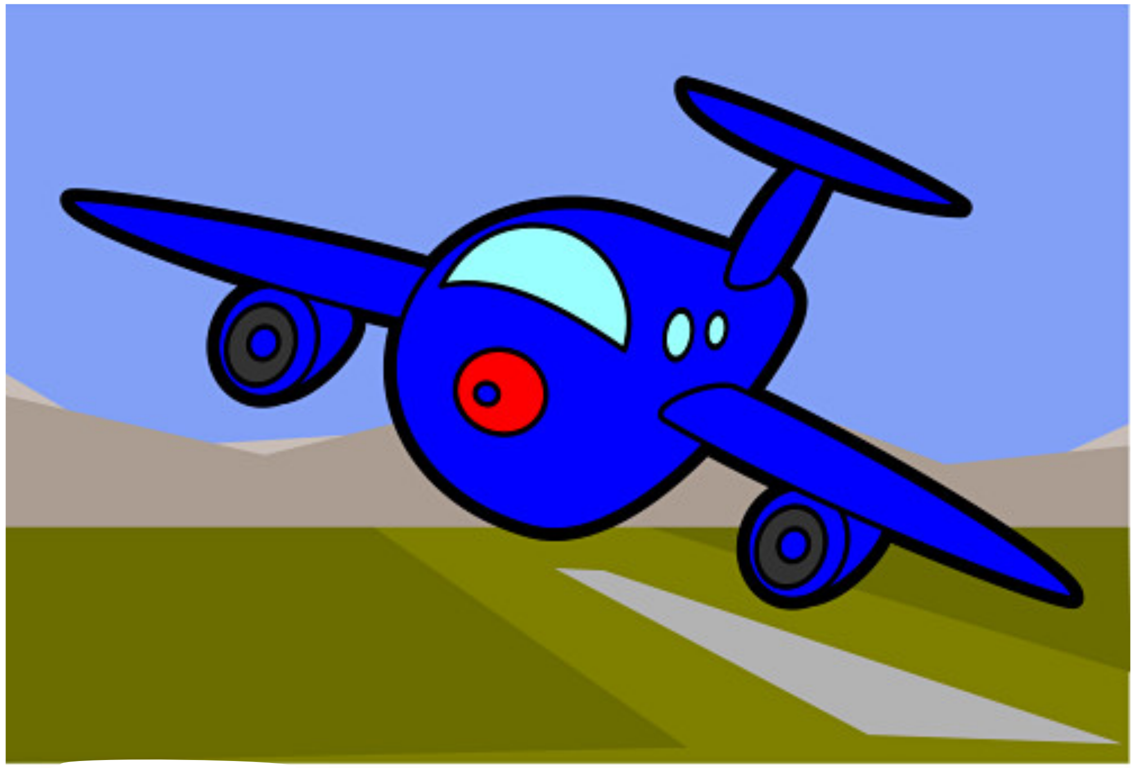
- "Semantic Parsing" is, ironically, a semantically ambiguous term
  - Semantic role labeling
  - Finding generic relations in text
  - → Transforming a natural language sentence into its meaning representation

Semantic Parsing

- Semantic Parsing: Transforming natural language (NL) sentences into computer executable complete meaning representations (MRs) for domain-specific applications
- Realistic semantic parsing currently entails domain dependence
- Example application domains
  - ATIS: Air Travel Information Service
  - CLang: Robocup Coach Language
  - Geoquery: A Database Query Application

ATIS: Air Travel Information Service

- Interface to an air travel database [Price, 1990]
- Widely-used benchmark for spoken language understanding

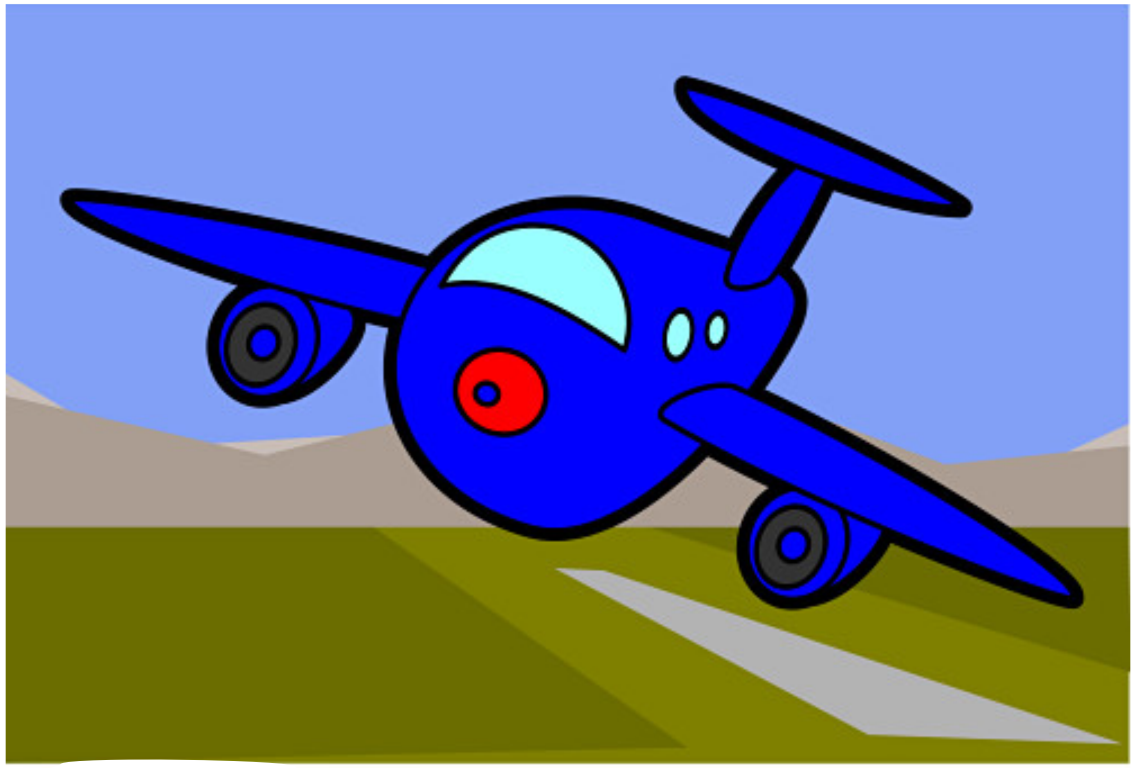


May I see all the flights from Cleveland to Dallas?

Air-Transportation
Show: (Flight-Number)
Origin: (City "Cleveland")
Destination: (City "Dallas")

ATIS: Air Travel Information Service

- Interface to an air travel database [Price, 1990]
- Widely-used benchmark for spoken language understanding



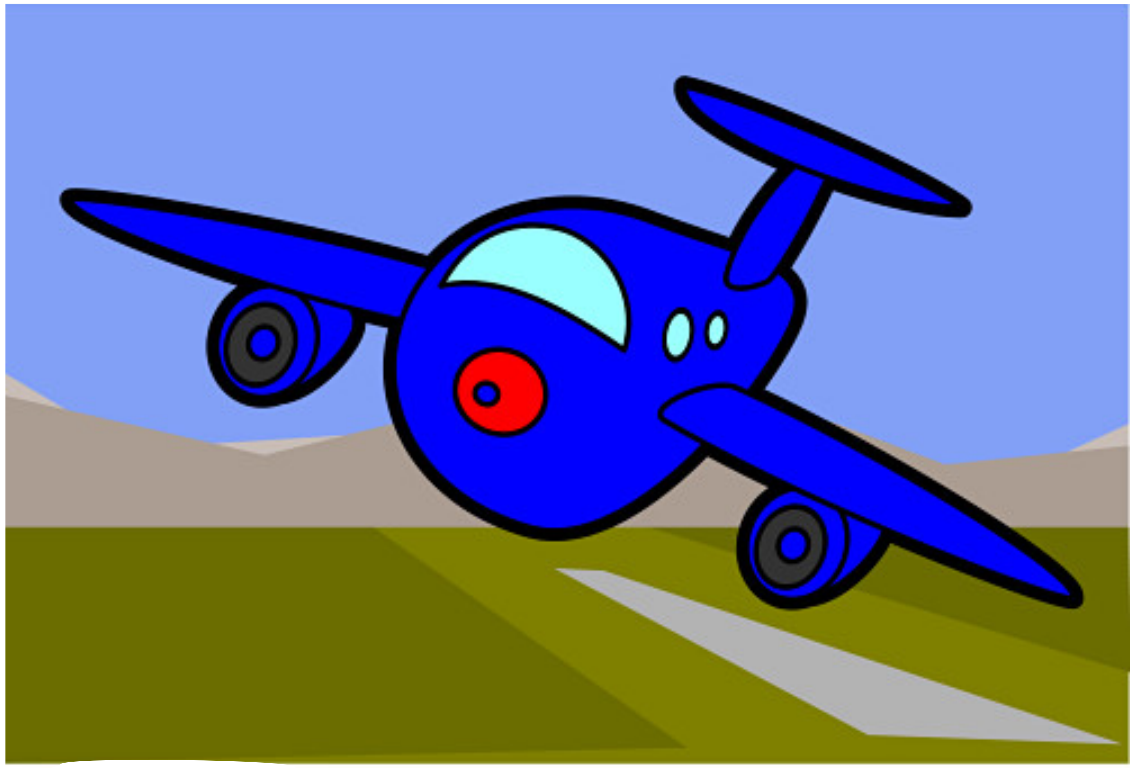
May I see all the flights from Cleveland to Dallas?

Semantic Parsing

Air-Transportation
Show: (Flight-Number)
Origin: (City "Cleveland")
Destination: (City "Dallas")

ATIS: Air Travel Information Service

- Interface to an air travel database [Price, 1990]
- Widely-used benchmark for spoken language understanding



May I see all the flights from Cleveland to Dallas?

Semantic Parsing

Air-Transportation
Show: (Flight-Number)
Origin: (City "Cleveland")
Destination: (City "Dallas")

Query

NA 1439,
TQ 23,
…

Geoquery: A Database Query Application

Which rivers run through the states bordering Texas?

Geoquery: A Database Query Application

Which rivers run through the states bordering Texas?

Semantic Parsing

answer(traverse(next_to(stateid('texas'))))

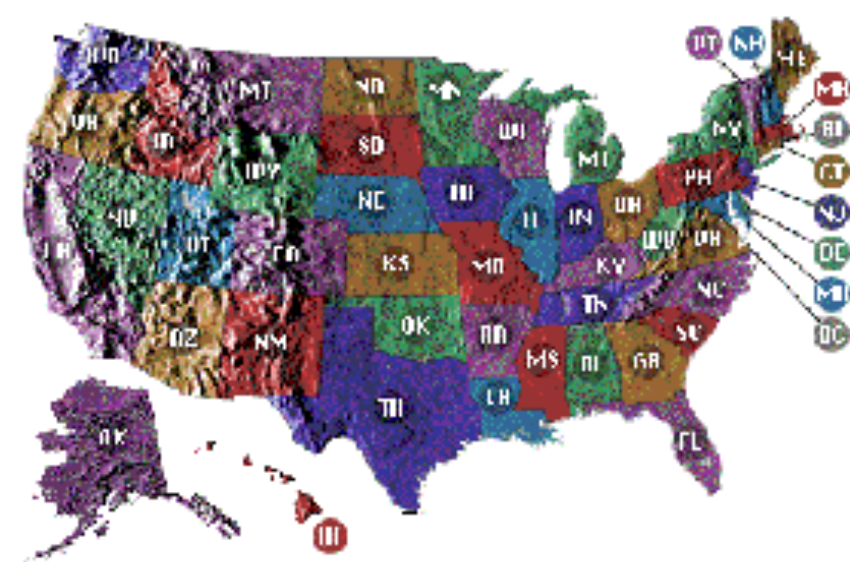
Geoquery: A Database Query Application

Which rivers run through the states bordering Texas?

Semantic Parsing

answer(traverse(next_to(stateid('texas'))))

Query



Geoquery: A Database Query Application

Which rivers run through the states bordering Texas?

Semantic Parsing

answer(traverse(next_to(stateid('texas'))))

Query

Answer

Arkansas, Canadian, Cimarron, Gila, Mississippi, Rio Grande …

Meaning Representation Languages

- Meaning representation language (MRL) for an application is assumed to be present

- MRL is designed by the creators of the application to suit the application’s needs independent of natural language

- CLang was designed by RoboCup community to send formal coaching instructions to simulated players

- Geoquery’s MRL was based on the Prolog database

Engineering Motivation for Semantic Parsing

- Applications of domain-dependent semantic parsing
  - Natural language interfaces to computing systems
  - Communication with robots in natural language
  - Personalized software assistants
  - Question-answering systems
- Machine learning makes developing semantic parsers for specific applications more tractable
- Training corpora can be easily developed by tagging natural-language glosses with formal statements

Distinctions from Other NLP Tasks

- Shallow semantic processing
  - Information extraction
  - Semantic role labeling
- Intermediate linguistic representations
  - Part-of-speech tagging
  - Syntactic parsing
  - Semantic role labeling
- Output meant for humans
  - Question answering
  - Summarization
  - Machine translation


ZENITH
STRATOS

Midterm Content Review

(at a high level)

How do you go out and solve new problems involving text?

1. Language has structure

"… is a film which still causes real, not figurative, chills to run along my spine, and it is certainly the bravest and most ambitious fruit of Coppola's genius"

Roger Ebert, Apocalypse Now

"I hated this movie. Hated hated hated hated hated this movie. Hated it. Hated every simpering stupid vacant audience-insulting moment of it. Hated the sensibility that thought anyone would like it."

Roger Ebert, North

Bag of words

Representation of text only as the counts of words that it contains

	Apocalypse now	North
the	1	1
of	0	0
hate	0	9
genius	1	0
bravest	1	0
stupid	0	1
like	0	1
…		

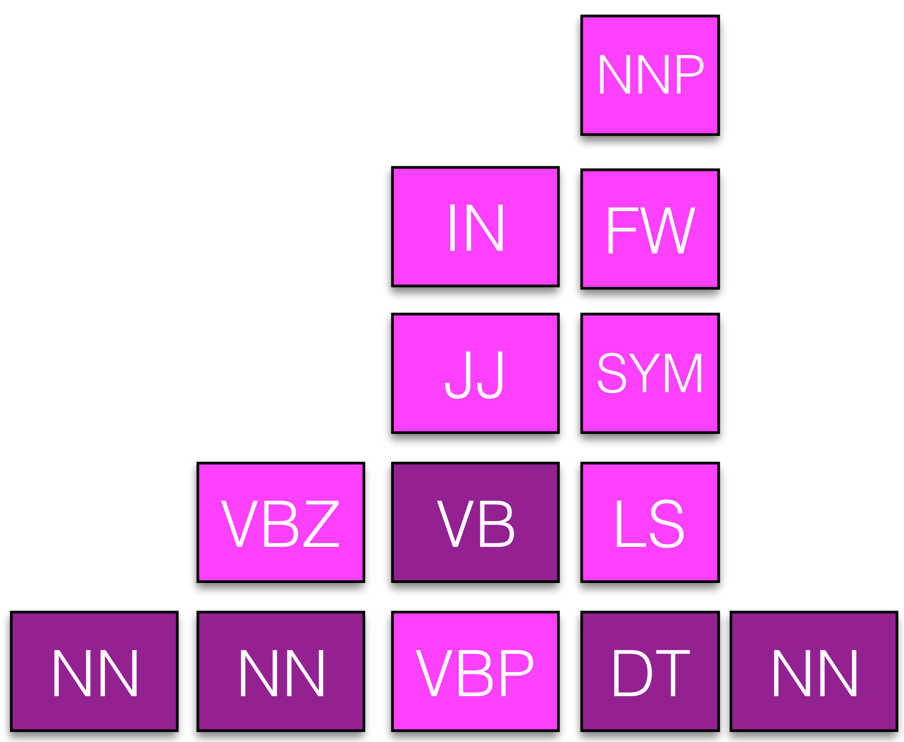
Language model

“Hillary Clinton seemed to add Benghazi to her already-long list of culprits to blame for her upset loss to Donald ________”

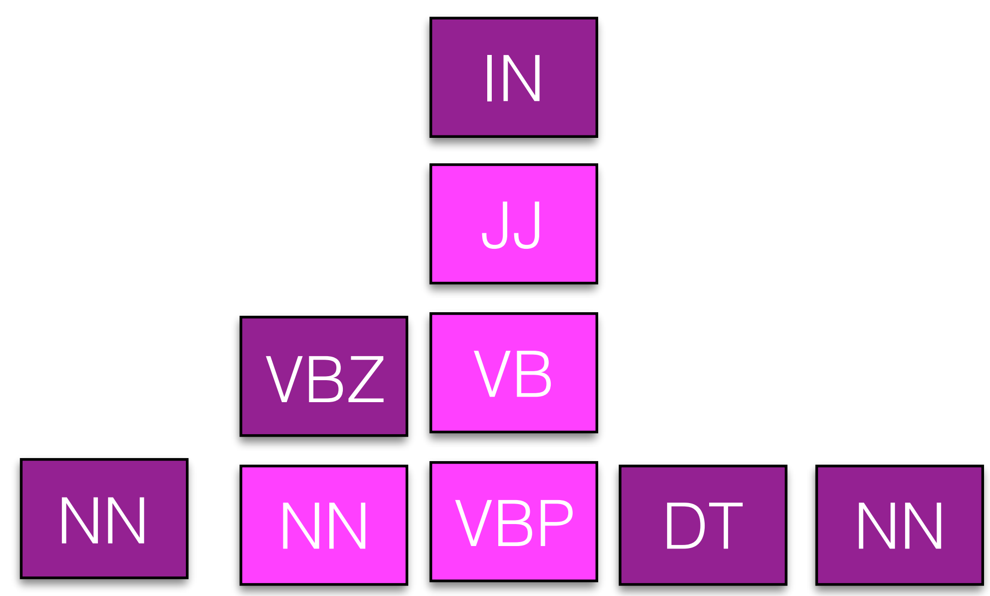
http://www.foxnews.com/politics/2017/09/13/clinton-laments-how-benghazi-tragedy-hurt-her-politically.html

POS tagging

Labeling the tag that's correct for the context.



Fruit flies like a banana



Time flies like an arrow

(Just tags in evidence within the Penn Treebank — more are possible!)

Word senses

original	It urged that the city take steps to remedy this problem
lemma sense	It urge^{1} that the city^{2} take^{1} step^{1} to remedy^{1} this problem^{2}
synset number	It urge$^{2:32:00}$ that the city$^{1:15:01}$ take$^{2:41:04}$ step$^{1:04:02}$ to remedy$^{2:30:00}$ this problem$^{1:10:00}$

Supersense tagging

artifact artifact motion time group

The station wagons arrived at noon, a long shining line

motion location location

that coursed through the west campus.

Noun supersenses (Ciarmita and Altun 2003)

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

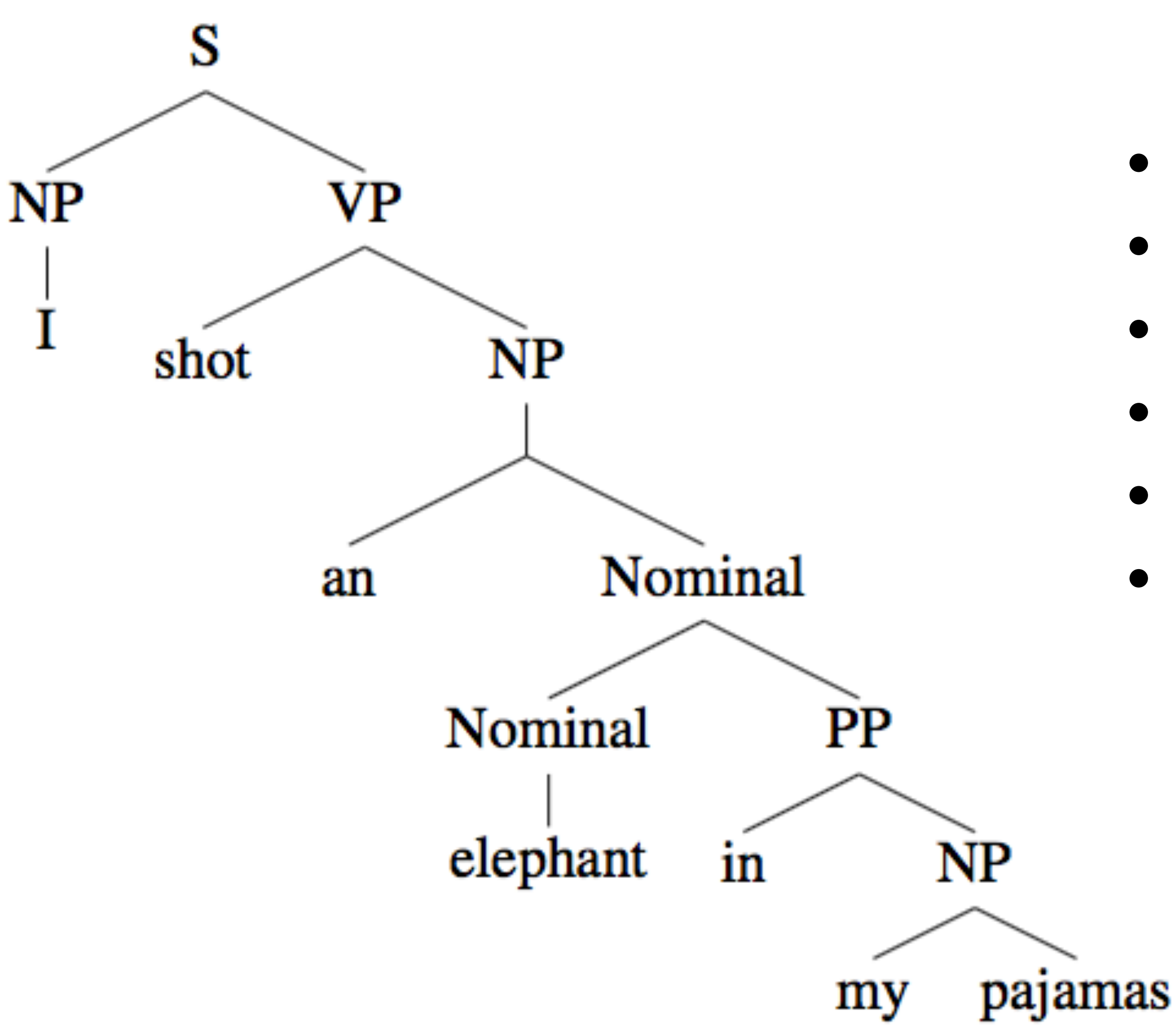
Agent Theme Goal

Doris gave the book to Cary

Agent Goal Theme

Doris gave Cary the book

Phrase-structure syntax

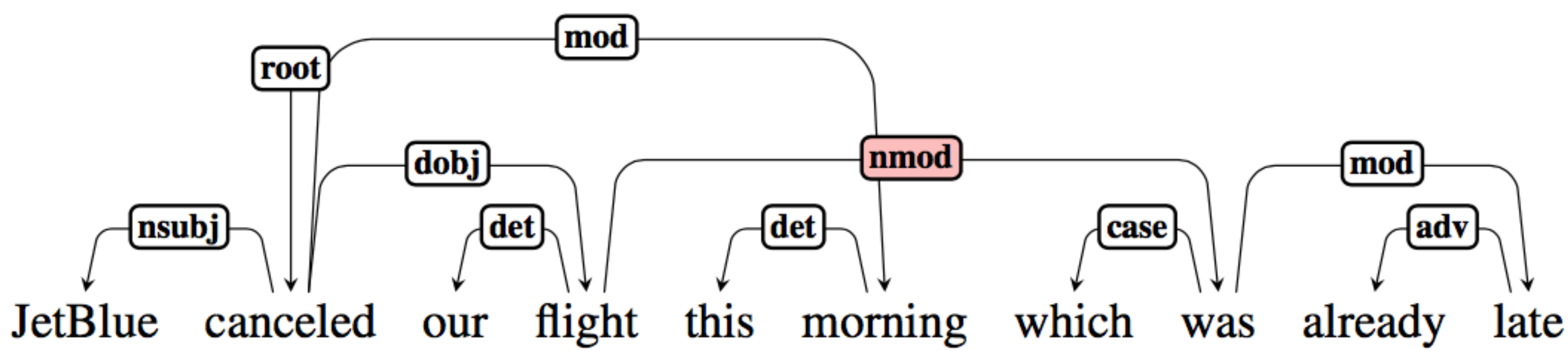


Every internal node is a phrase

- my pajamas
- in my pajamas
- elephant in my pajamas
- an elephant in my pajamas
- shot an elephant in my pajamas
- I shot an elephant in my pajamas

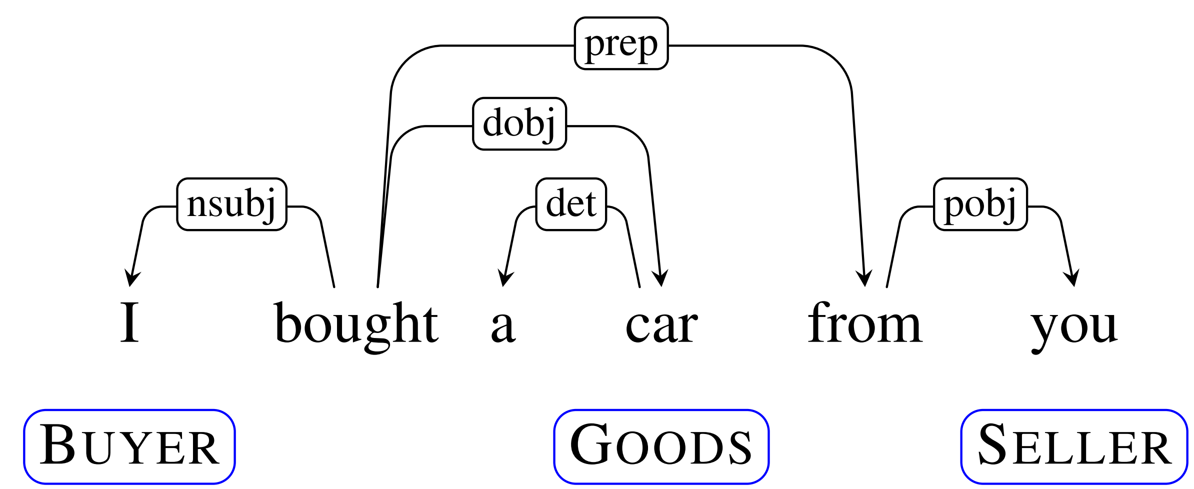
Each phrase could be replaced by another of the same type of constituent

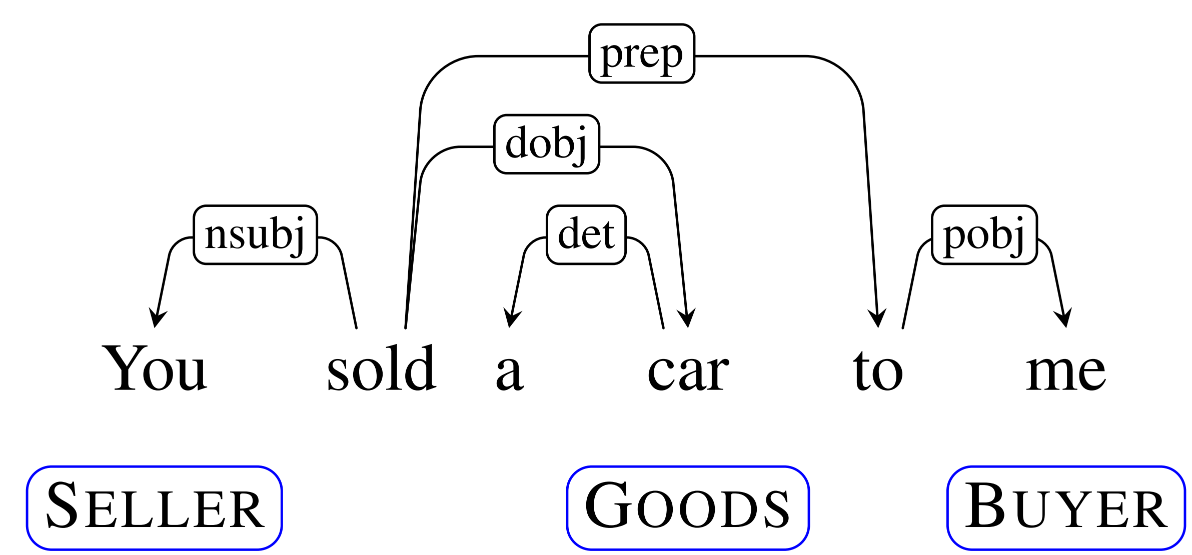
Dependency syntax



Semantic Frame

Two different perspectives on a commercial transaction





Compositional semantics

Utah borders Idaho

Compositional semantics

NP	(S\NP)/NP	NP
utah	**λx.λy(borders(y,x)**	**idaho**
Utah	borders	Idaho

Compositional semantics

$$\dfrac{\text{(S\textbackslash NP)/NP} \quad \text{NP}}{\text{S\textbackslash NP}}$$

NP	(S\NP)/NP	NP
utah	**λx.λy(borders(y,x))**	**idaho**
Utah	borders	Idaho

Combined (borders + Idaho): S\NP — **λy(borders(y,idaho))**

Compositional semantics

S
borders(utah,idaho)

S\NP
λy(borders(y,idaho)

NP	(S\NP)/NP	NP
utah	**λx.λy(borders(y,x)**	**idaho**
Utah	borders	Idaho

Discourse structure

BECAUSE

- John hid Bill’s car keys. He was drunk
- John hid Bills car keys. He likes spinach.

With its distant orbit,

Mars experiences frigid weather.

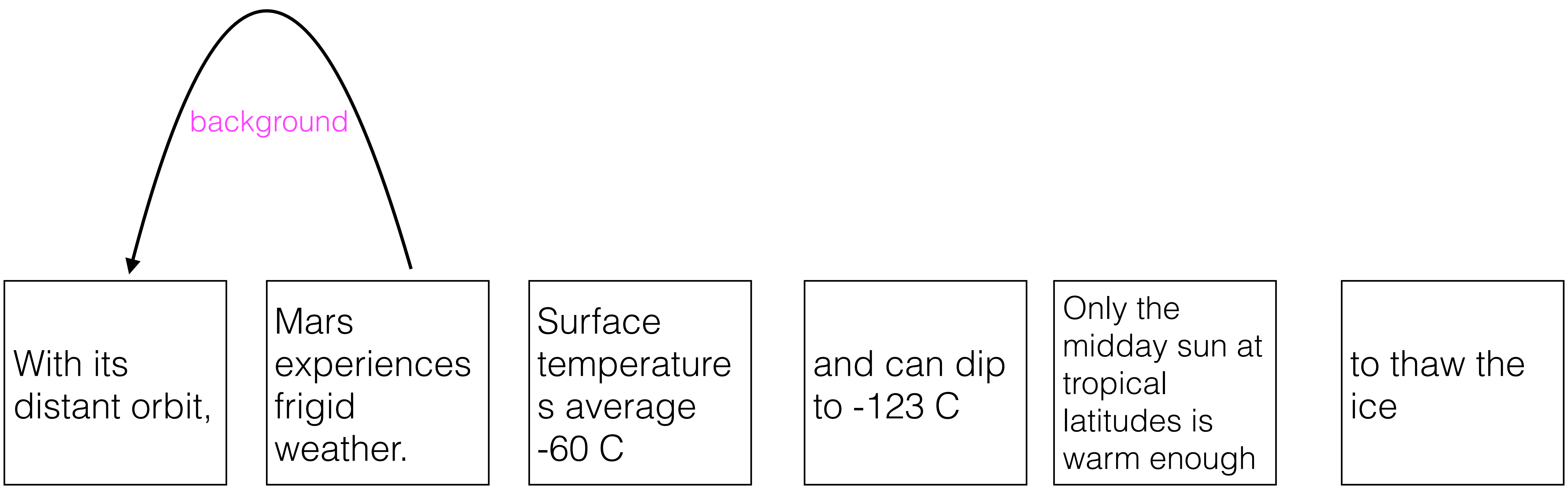
Surface temperatures average -60 C

and can dip to -123 C

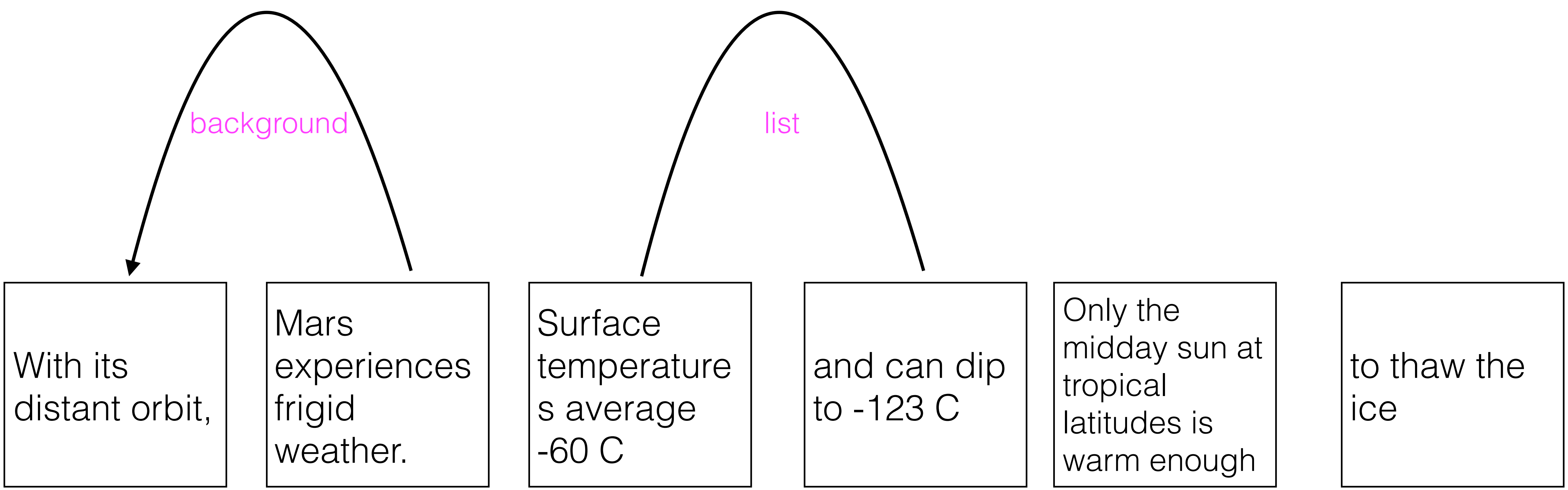
Only the midday sun at tropical latitudes is warm enough

to thaw the ice

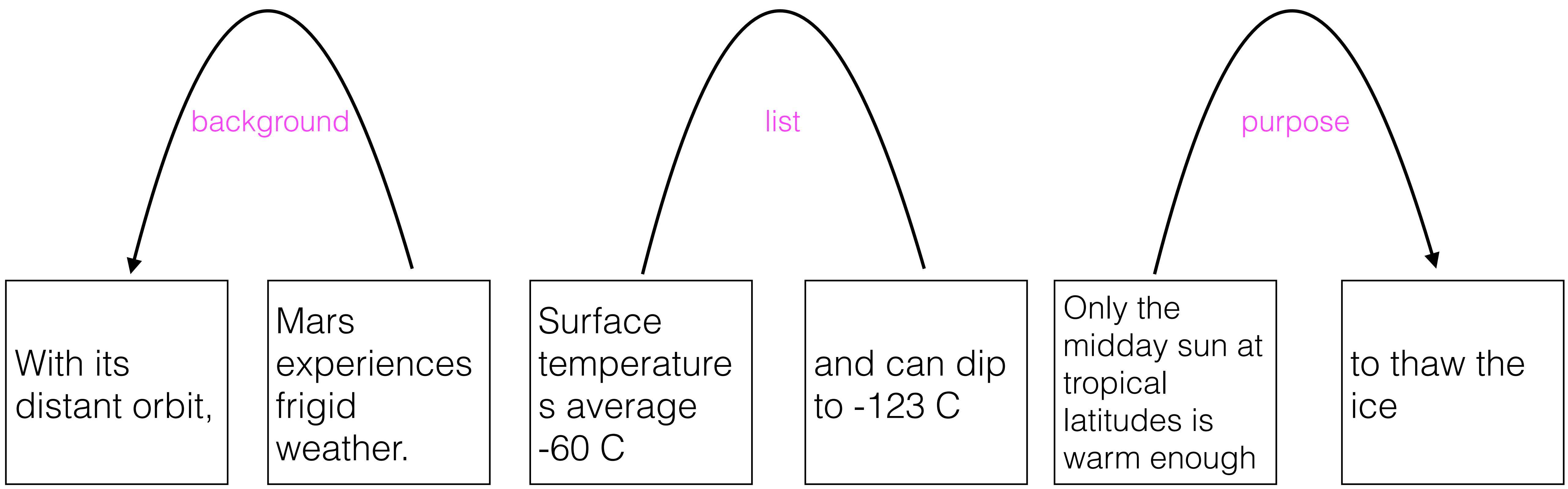
Arrows are directed from nucleus to satellite (undirected edges are nuclei)



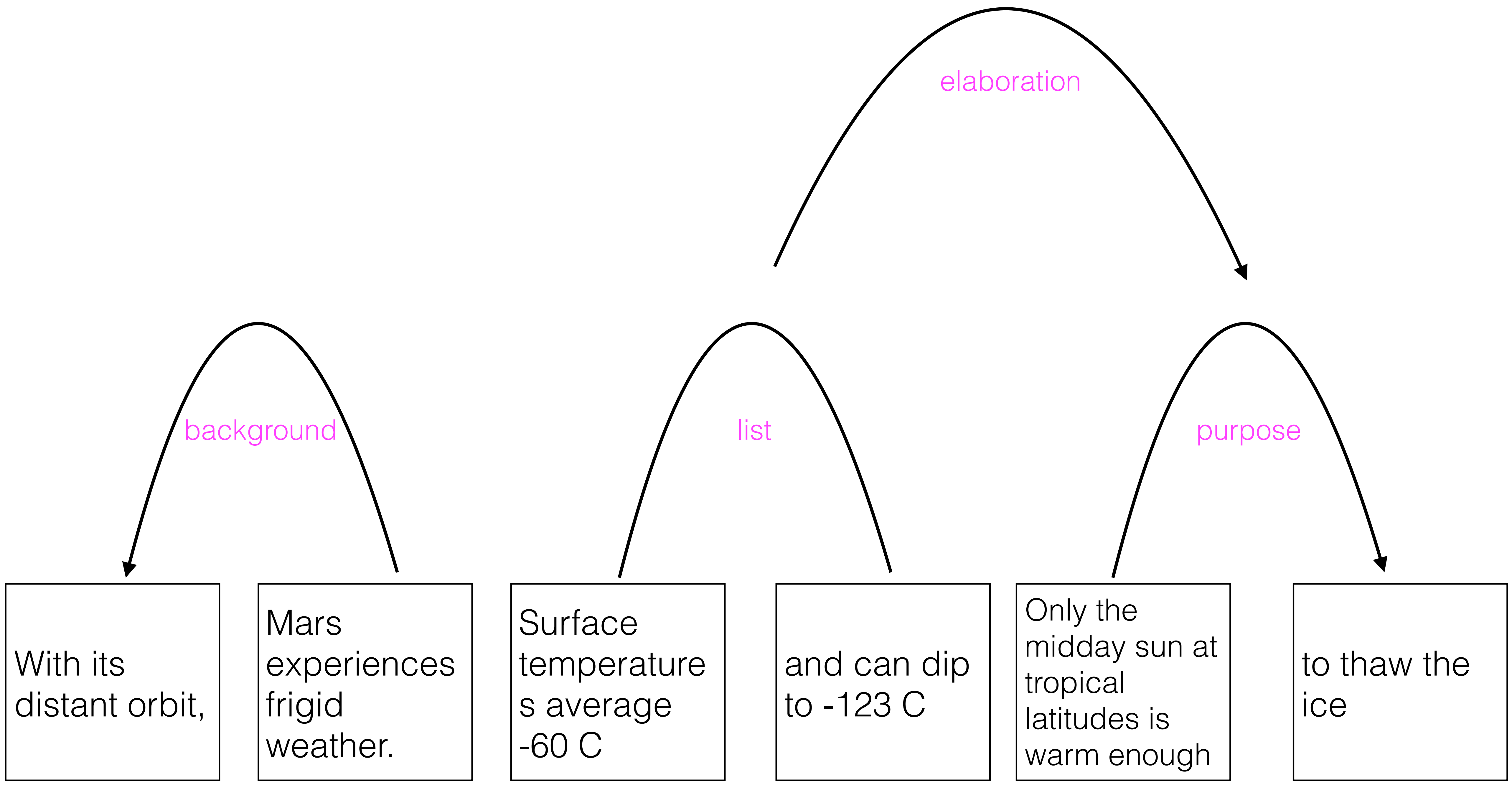
Arrows are directed from nucleus to satellite (undirected edges are nuclei)



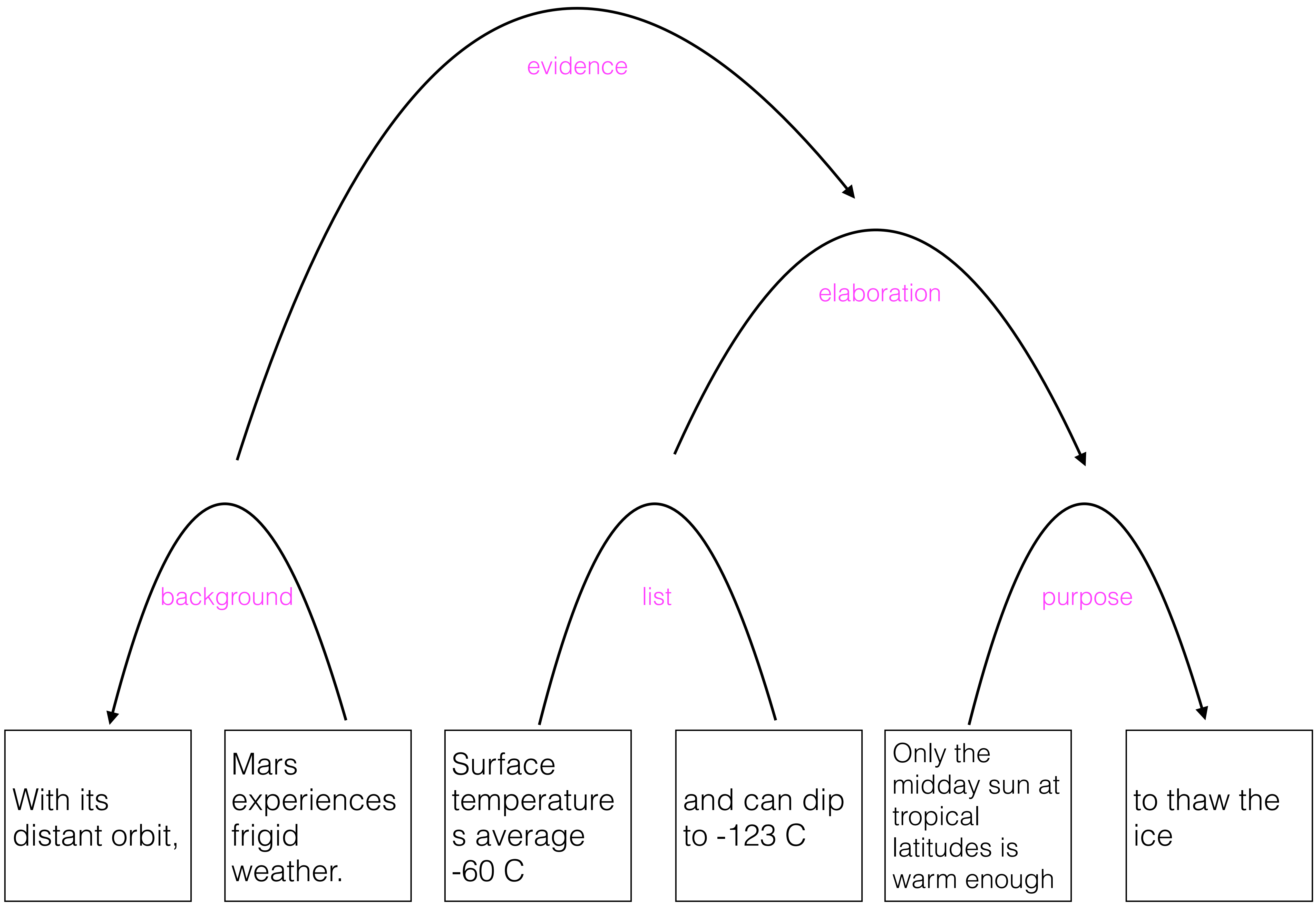
Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Coreference

LUKE

I'll never join you!

VADER

If you only knew the power of the
dark side. Obi-Wan never told
you what happened to your father.

LUKE

He told me enough! It was you
who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know
it to be true.

LUKE

No! No! No!

Information extraction

The trade deadline is getting much more exciting, though.

We got a taste when the Jaguars traded for two-time Pro Bowl defensive tackle Marcell Dareus on Friday, and the Seahawks traded for three-time Pro Bowl offensive tackle Duane Brown on Monday.

But the bombshell came Monday night when the Patriots sent Jimmy Garoppolo to become the new franchise quarterback of the 49ers.

- Who is on the current roster of every NFL team?
- If there are trades between teams, when and who was involved?

2. Most new problems can solved with a familiar class of algorithms

Many computation tools available!

- Classification
- Sequence labeling
- Trees
- Graphs

Many computation tools available!

- Classification
- Sequence labeling
- Trees
- Graphs

- Counting and normalizing (NB, PCFG, HMM, LDA)
- Loglinear (logistic regression, MEMM, CRF)
- Neural (CNN, RNN, LSTM, Transformer)

Classification

Bayes' Rule

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{\sum_y P(Y = y)P(X = x|Y = y)}$$

Bayes' Rule

Prior belief that Y = positive
(before you see any data)

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{\sum_y P(Y = y)P(X = x|Y = y)}$$

Bayes' Rule

Prior belief that Y = positive
(before you see any data)

Likelihood of "really really the worst movie ever"
given that Y= positive

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{\sum_y P(Y = y)P(X = x|Y = y)}$$

Bayes' Rule

Prior belief that Y = positive
(before you see any data)

Likelihood of "really really the worst movie ever"
given that Y= positive

$$P(Y=y|X=x) = \frac{P(Y=y)P(X=x|Y=y)}{\sum_y P(Y=y)P(X=x|Y=y)}$$

This sum ranges over
y=positive + y=negative
(so that it sums to 1)

Bayes' Rule

Prior belief that Y = positive (before you see any data)

Likelihood of "really really the worst movie ever" given that Y= positive

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{\sum_y P(Y = y)P(X = x|Y = y)}$$

Posterior belief that Y=positive given that X="really really the worst movie ever"

This sum ranges over y=positive + y=negative (so that it sums to 1)

Naive Bayes training

Training a Naive Bayes classifier consists of estimating these two quantities from training data for all classes y

$$P(Y=y|X=x) = \frac{P(Y=y)P(X=x|Y=y)}{\sum_y P(Y=y)P(X=x|Y=y)}$$

At test time, use those estimated probabilities to calculate the posterior probability of each class y and select the class with the highest probability

Independence Assumption

really really the worst movie ever

We will assume the features are independent:

$$P(x_1, x_2, x_3, x_4, x_6, x_7 \mid c) = P(x_1 \mid c)P(x_2 \mid c) \ldots P(x_7 \mid c)$$

$$P(x_i \ldots x_n \mid c) = \prod_{i=1}^{N} P(x_i \mid c)$$

Logistic regression

$$P(y = 1 \mid x, \beta) = \frac{1}{1 + \exp\left(-\sum_{i=1}^{F} x_i \beta_i\right)}$$

output space $\mathcal{Y} = \{0, 1\}$

x = feature vector

Feature	Value
the	0
and	0
bravest	0
love	0
loved	0
genius	0
not	0
fruit	1
BIAS	1

β = coefficients

Feature	β
the	0.01
and	0.03
bravest	1.4
love	3.1
loved	1.2
genius	0.5
not	-3.0
fruit	-0.8
BIAS	-0.1

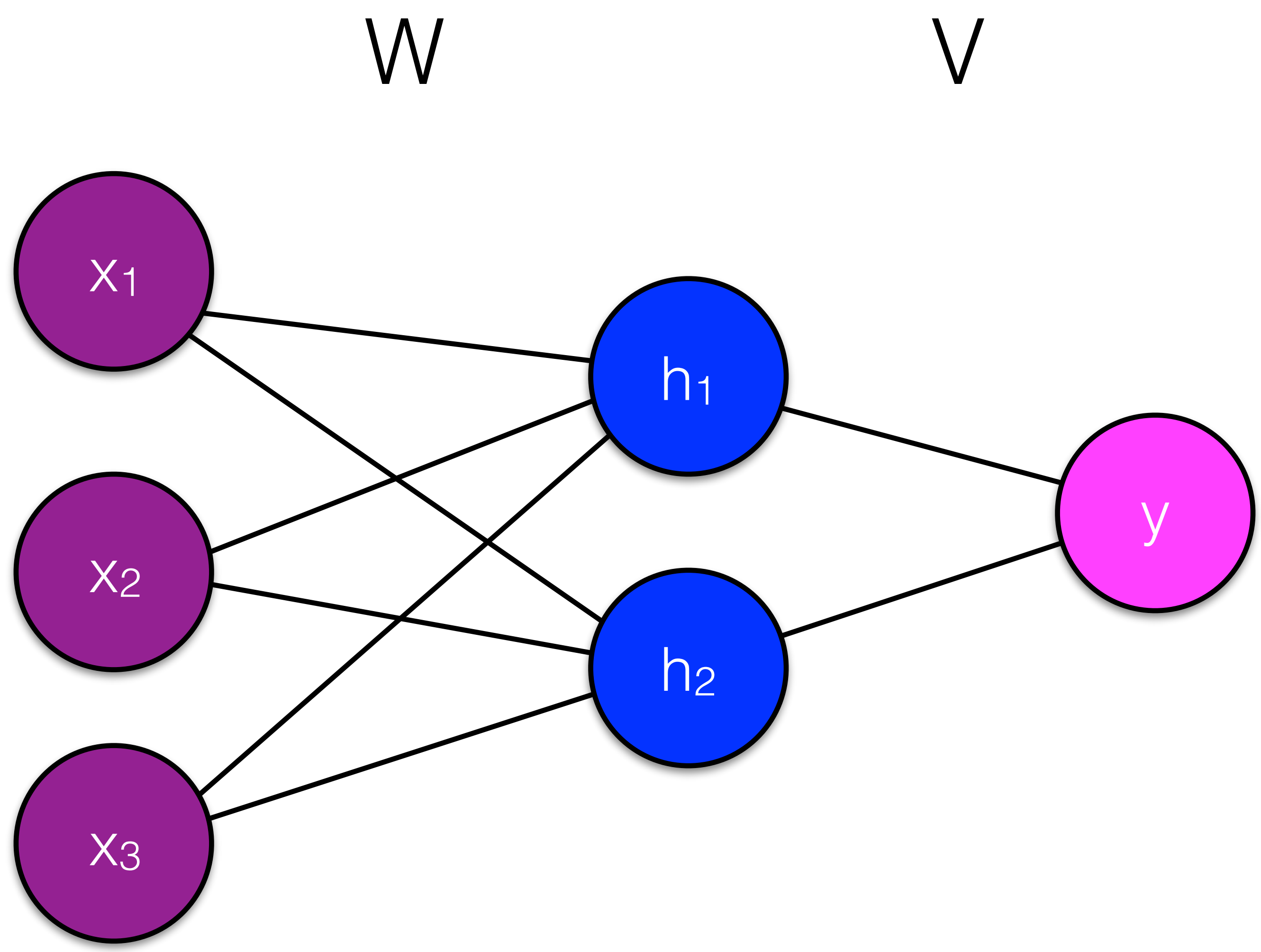
Features

- As a discriminative classifier, logistic regression doesn't assume features are independent like Naive Bayes does.

- Its power partly comes in the ability to create richly expressive features with out the burden of independence.

- We can represent text through features that are not just the identities of individual words, but any feature that is scoped over the entirety of the input.

features
contains like
has word that shows up in positive sentiment dictionary
review begins with "I like"
at least 5 mentions of positive affectual verbs (like, love, etc.)



$$\hat{y} = \sigma\left[V_1\left(\sigma\left(\sum_{i}^{F} x_i W_{i,1}\right)\right) + V_2\left(\sigma\left(\sum_{i}^{F} x_i W_{i,2}\right)\right)\right]$$

we can express y as a function only of the input x and the weights W and V

Sequences

Sequence labeling

$$x = \{x_1, \ldots, x_n\}$$

$$y = \{y_1, \ldots, y_n\}$$

- For a set of inputs x with n sequential time steps, one corresponding label y_i for each x_i
- Model the structure that exists between within y

HMM

$$P(x_1, \ldots, x_n, y_1, \ldots, y_n) \approx \prod_{i=1}^{n+1} P(y_i \mid y_{i-1}) \prod_{i=1}^{n} P(x_i \mid y_i)$$

Hidden Markov Model

Prior probability of label sequence

$$P(y) = P(y_1, \ldots, y_n)$$

$$P(y_1, \ldots, y_n) \approx \prod_{i=1}^{n+1} P(y_i \mid y_{i-1})$$

- We'll make a first-order Markov assumption and calculate the joint probability as the product the individual factors conditioned only on the previous tag.

Hidden Markov Model

$$P(x \mid y) = P(x_1, \ldots, x_n \mid y_1, \ldots, y_n)$$

$$P(x_1, \ldots, x_n \mid y_1, \ldots, y_n) \approx \prod_{i=1}^{N} P(x_i \mid y_i)$$

- Here again we'll make a strong assumption: the probability of the word we see at a given time step is only dependent on its label

MEMM

$$\arg\max_y P(y \mid x, \beta)$$

$$\arg\max_y \prod_{i=1}^{n} P(y_i \mid y_{i-1}, x)$$

MEMM

General maxent form

$$\arg\max_y P(y \mid x, \beta)$$

$$\arg\max_y \prod_{i=1}^{n} P(y_i \mid y_{i-1}, x)$$

MEMM

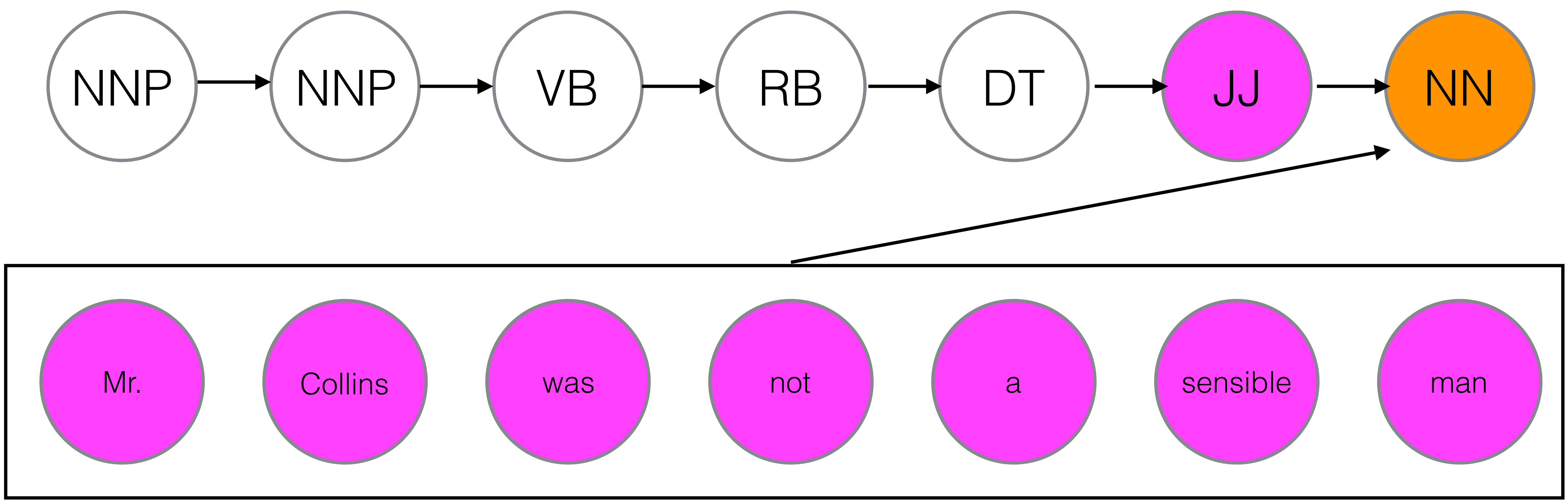
General maxent form

$$\arg\max_y P(y \mid x, \beta)$$

Maxent with first-order Markov assumption: Maximum Entropy Markov Model

$$\arg\max_y \prod_{i=1}^{n} P(y_i \mid y_{i-1}, x)$$

MEMM



Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

Features are scoped over the previous predicted tag and the entire observed input

Features

feature	example

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

Features are scoped over the previous predicted tag and the entire observed input

Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

feature	example
x_i = man	1

Features are scoped over the previous predicted tag and the entire observed input

Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

feature	example
x_i = man	1
t_{i-1} = JJ	1

Features are scoped over the previous predicted tag and the entire observed input

Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

Features are scoped over the previous predicted tag and the entire observed input

feature	example
x_i = man	1
t_{i-1} = JJ	1
i=n (last word of sentence)	1

Features

$$f(t_i, t_{i-1}; x_1, \ldots, x_n)$$

Features are scoped over the previous predicted tag and the entire observed input

feature	example
x_i = man	1
t_{i-1} = JJ	1
i=n (last word of sentence)	1
x_i ends in -ly	0

MEMM Training

$$\prod_{i=1}^{n} P(y_i \mid y_{i-1}, x, \beta)$$

Locally normalized — at each time step, each conditional distribution sums to 1

Conditional random fields

- We can solve this problem using global normalization (over the entire sequences) rather than locally normalized factors.

MEMM

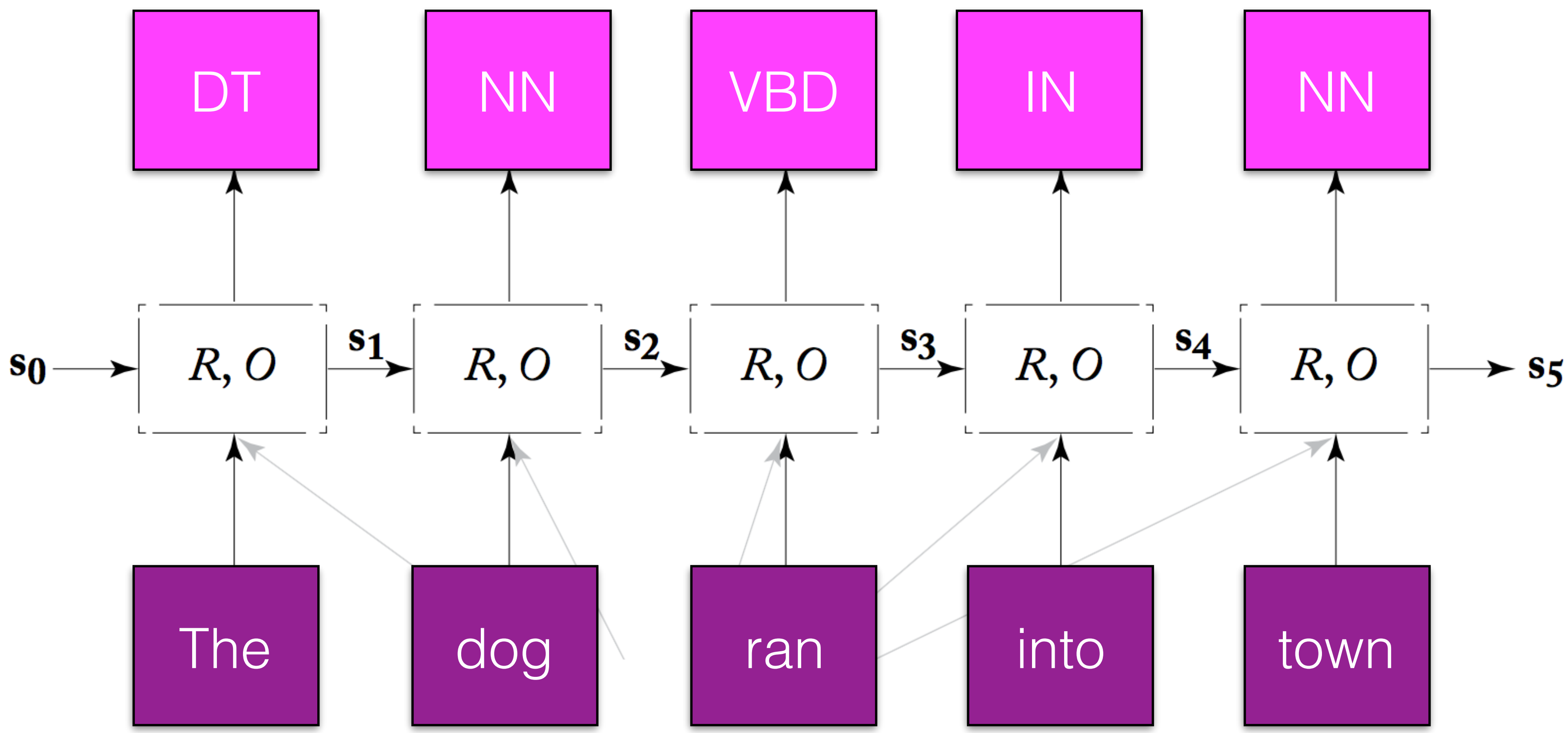
$$P(y \mid x, \beta) = \prod_{i=1}^{n} P(y_i \mid y_{i-1}, x, \beta)$$

CRF

$$P(y \mid x, \beta) = \frac{\exp(\Phi(x, y)^\top \beta)}{\sum_{y' \in \mathcal{Y}} \exp(\Phi(x, y')^\top \beta)}$$

Recurrent neural network

- For POS tagging, predict the tag conditioned on the context

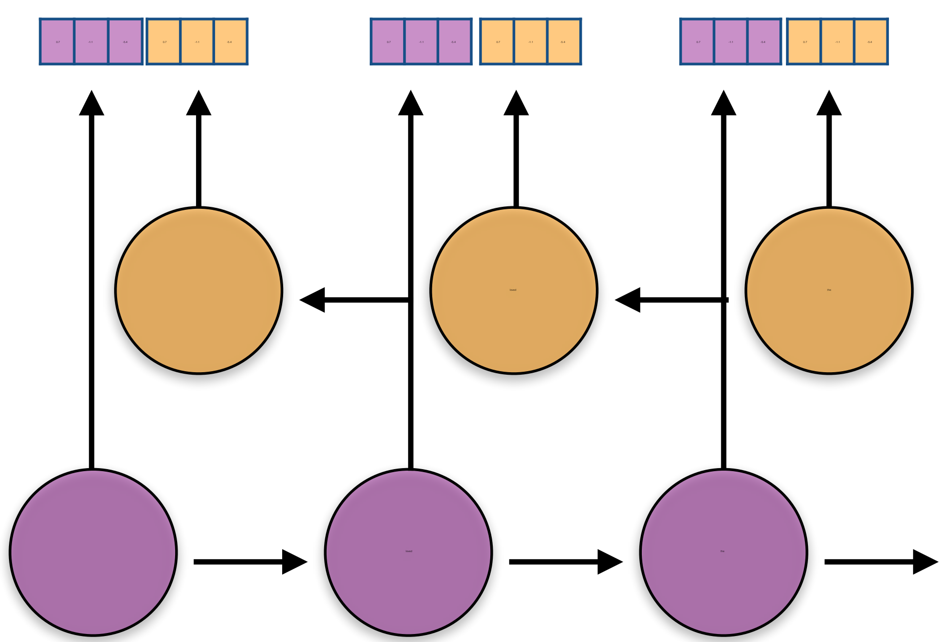


Bidirectional RNN

- A powerful alternative is make predictions conditioning both on the past and the future.
- Two RNNs
  - One running left-to-right
  - One right-to-left
- Each produces an output vector at each time step, which we concatenate

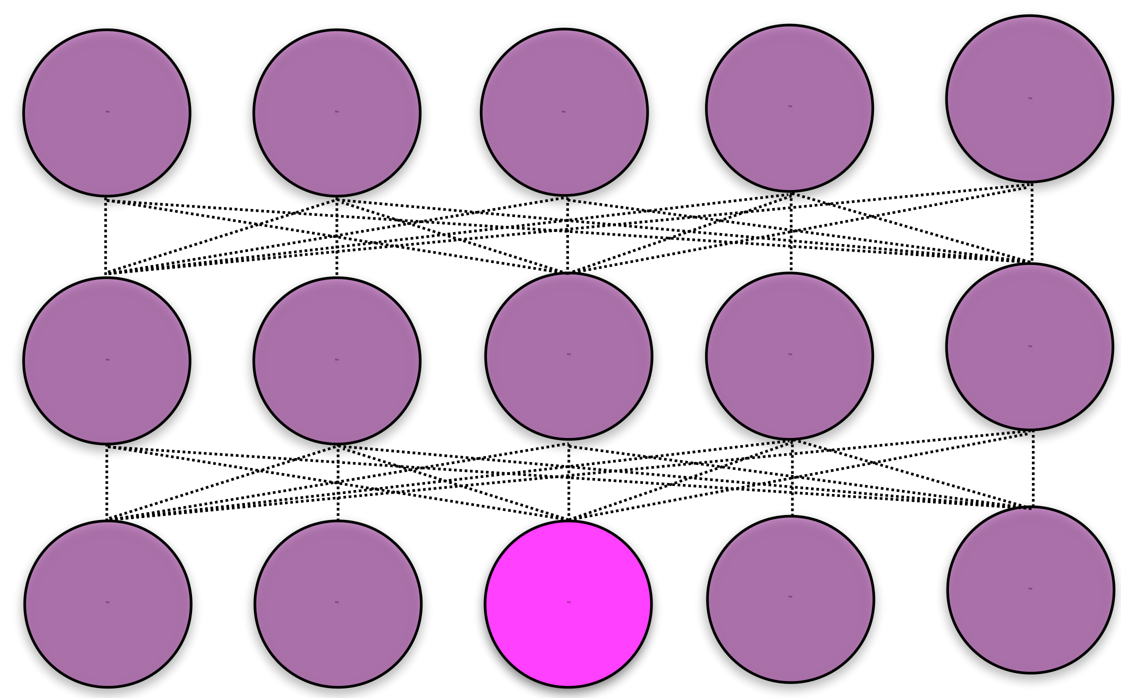
ELMo and BERT

Stacked BiRNN trained to predict next word in language modeling task



Peters et al. 2018

Transformer-based model to predict masked word using bidirectional context + next sentence prediction.



Devlin et al. 2019

Trees

PCFG

PCFG

- Probabilistic context-free grammar: each production is also associated with a probability.

PCFG

- Probabilistic context-free grammar: each production is also associated with a probability.

- This lets us calculate the probability of a parse for a given sentence; for a given parse tree T for sentence S comprised of n rules from R (each $A \rightarrow \beta$):

PCFG

- Probabilistic context-free grammar: each production is also associated with a probability.

- This lets us calculate the probability of a parse for a given sentence; for a given parse tree T for sentence S comprised of n rules from R (each A → β):

$$P(T, S) = \prod_{i}^{n} P(\beta \mid A)$$

Estimating PCFGs

$$\sum_{\beta} P(\beta \mid A) = \frac{C(A \rightarrow \beta)}{\sum_{\gamma} C(A \rightarrow \gamma)}$$

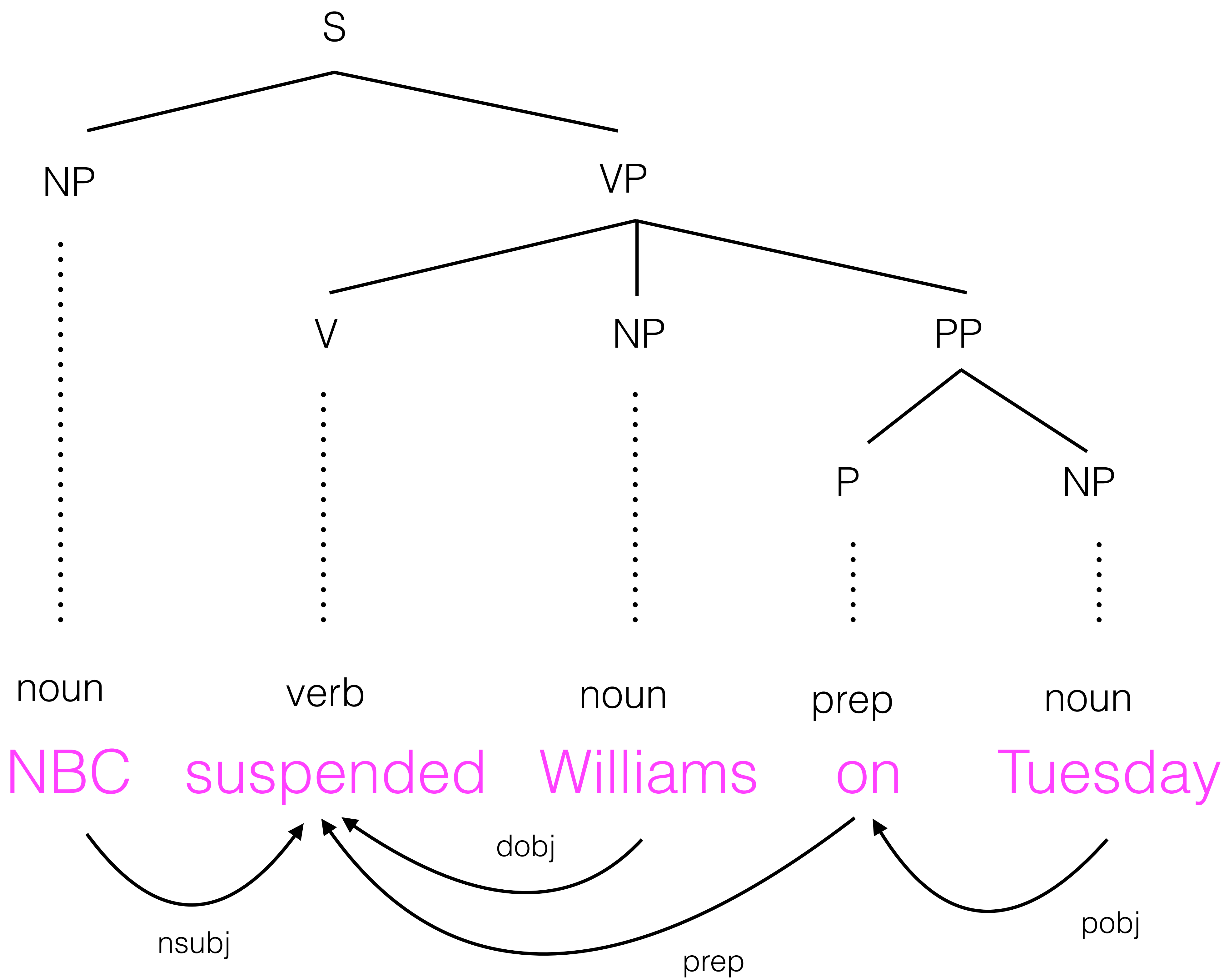
(equivalently)

$$\sum_{\beta} P(\beta \mid A) = \frac{C(A \rightarrow \beta)}{C(A)}$$

A		β	P(β | NP)
NP	→	NP PP	0.092
NP	→	DT NN	0.087
NP	→	NN	0.047
NP	→	NNS	0.042
NP	→	DT JJ NN	0.035
NP	→	NNP	0.034
NP	→	NNP NNP	0.029
NP	→	JJ NNS	0.027
NP	→	QP -NONE-	0.018
NP	→	NP SBAR	0.017
NP	→	NP PP-LOC	0.017
NP	→	JJ NN	0.015
NP	→	DT NNS	0.014
NP	→	CD	0.014
NP	→	NN NNS	0.013
NP	→	DT NN NN	0.013
NP	→	NP CC NP	0.013

Dependency parsing

- Transition-based parsing (HW3!)
  - O(n)
  - Only projective structures (pseudo-projective [Nivre and Nilsson 2005])
- Graph-based parsing
  - $O(n^2)$
  - Projective and non-projective trees

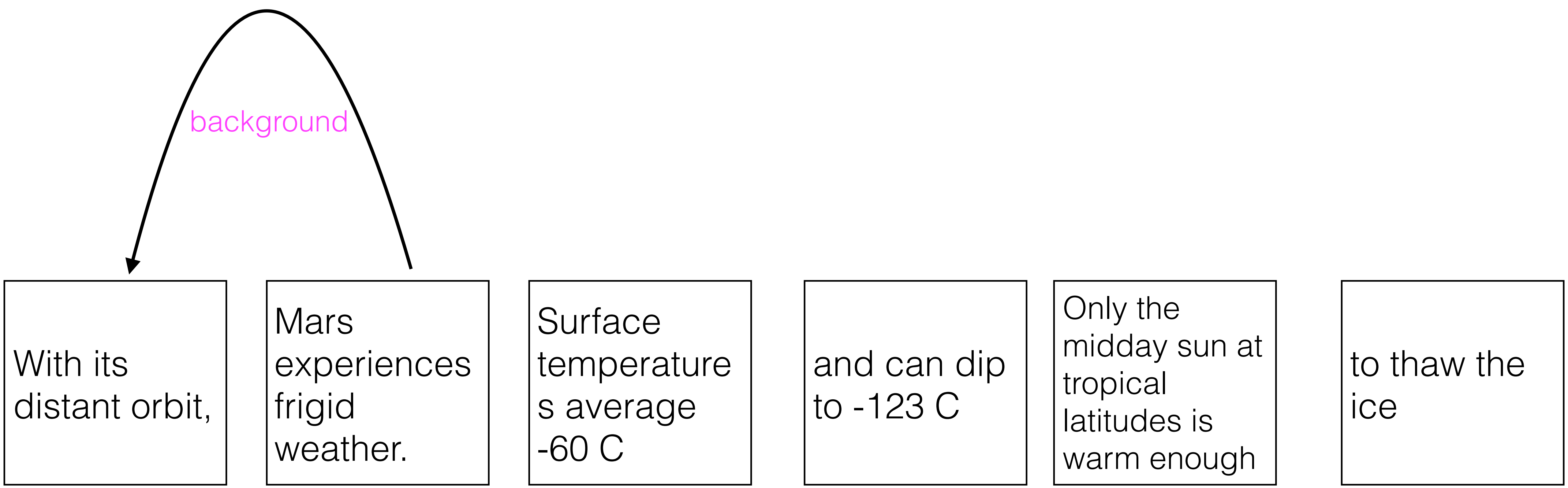

S
NP
VP
V
NP
PP
P
NP
noun
verb
noun
prep
noun
NBC
suspended
Williams
on
Tuesday
dobj
nsubj
prep
pobj

Rhetorical Structure Theory (RST)

With its distant orbit,	Mars experiences frigid weather.	Surface temperatures average -60 C	and can dip to -123 C	Only the midday sun at tropical latitudes is warm enough	to thaw the ice

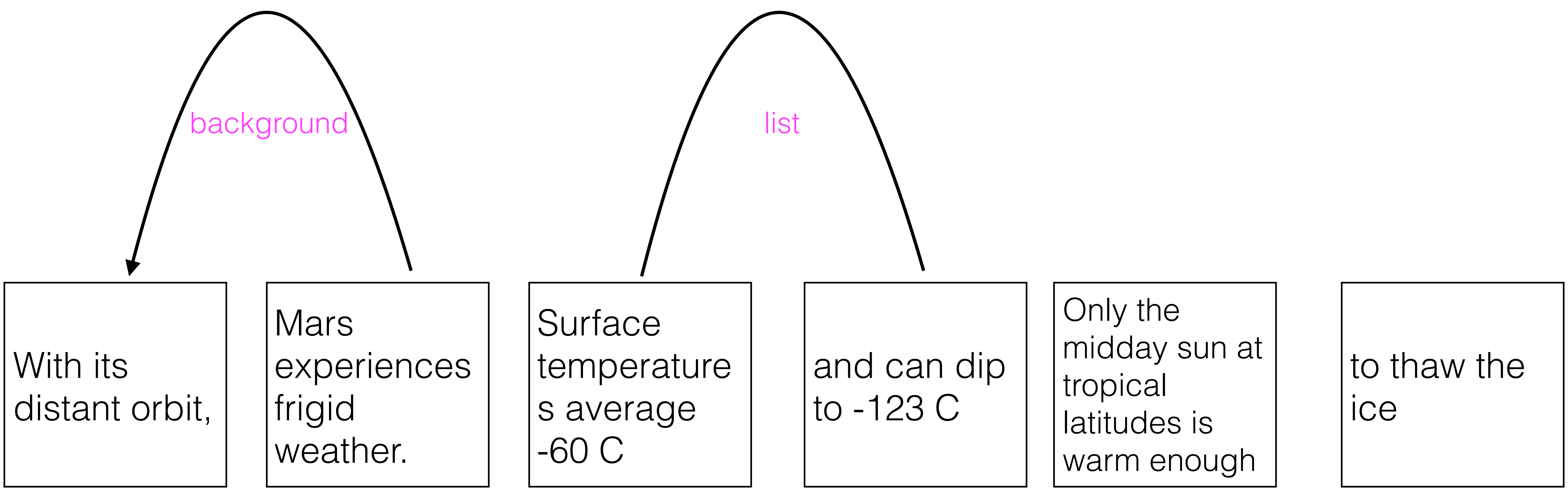
Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Rhetorical Structure Theory (RST)



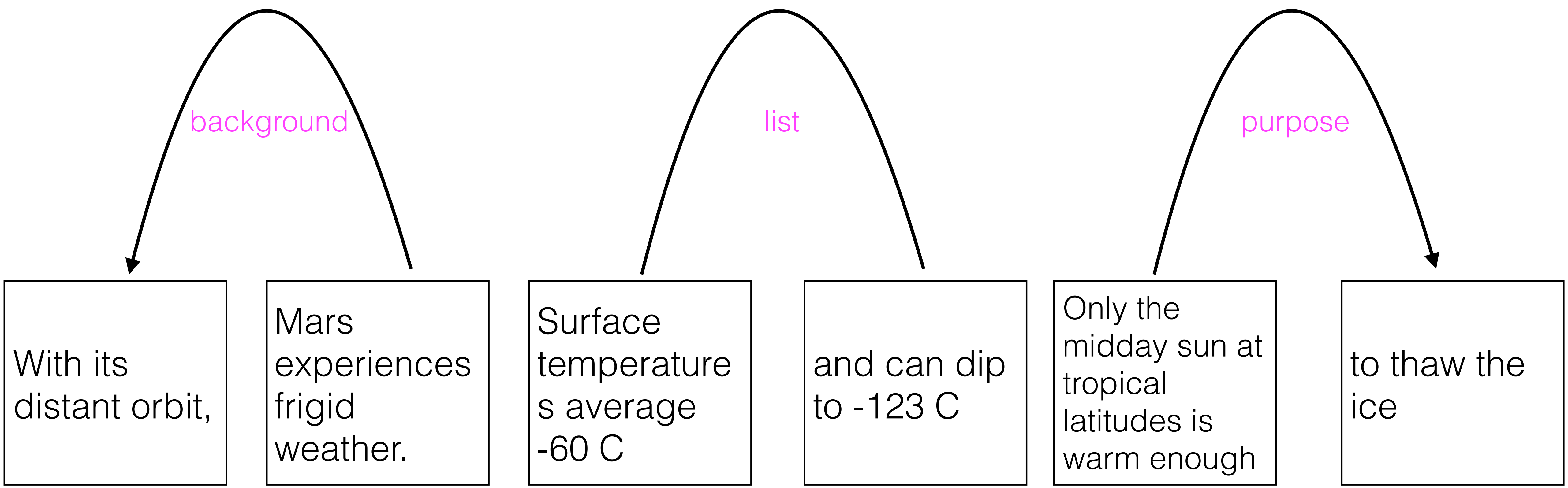
Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Rhetorical Structure Theory (RST)



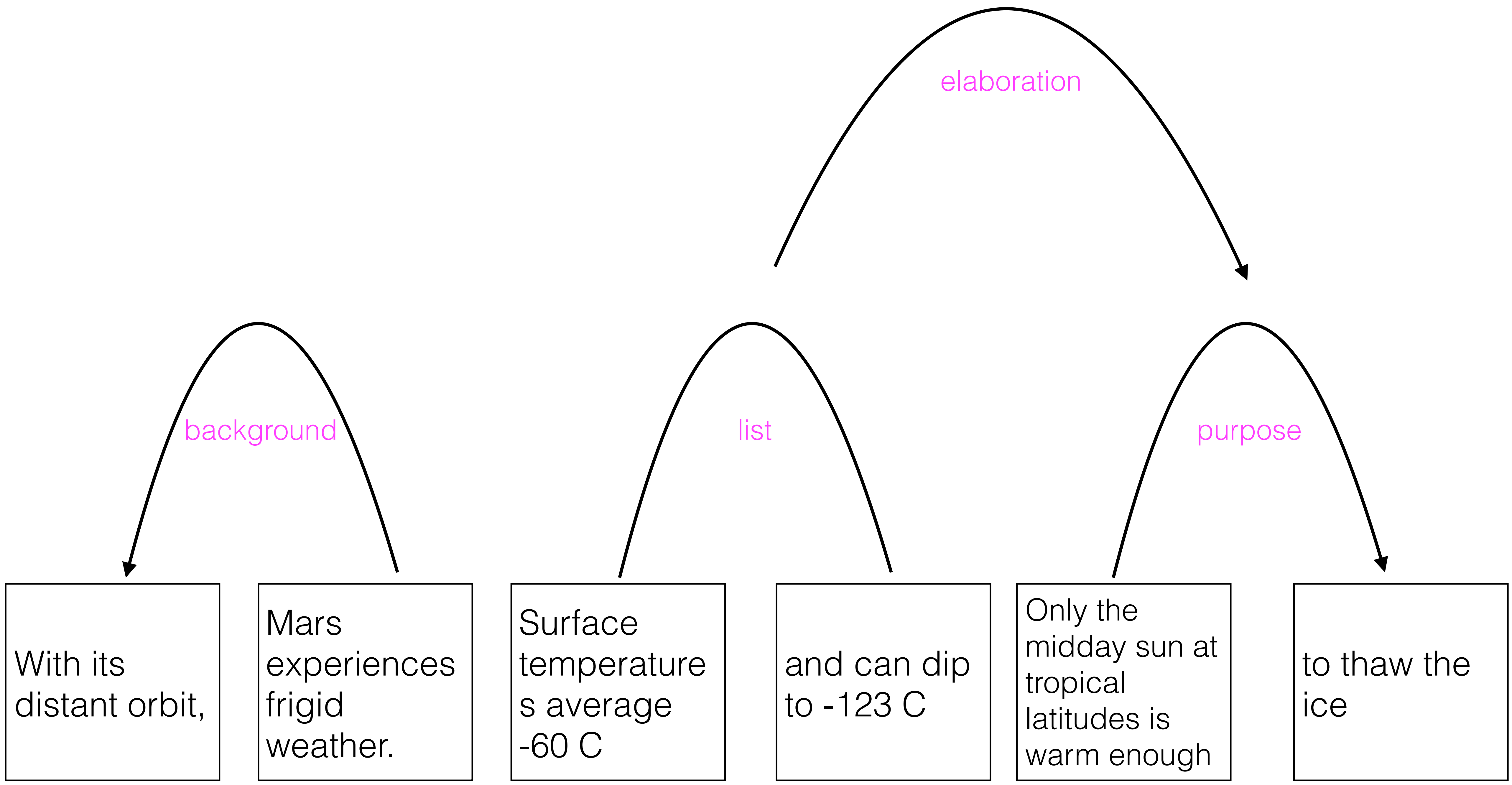
Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Rhetorical Structure Theory (RST)



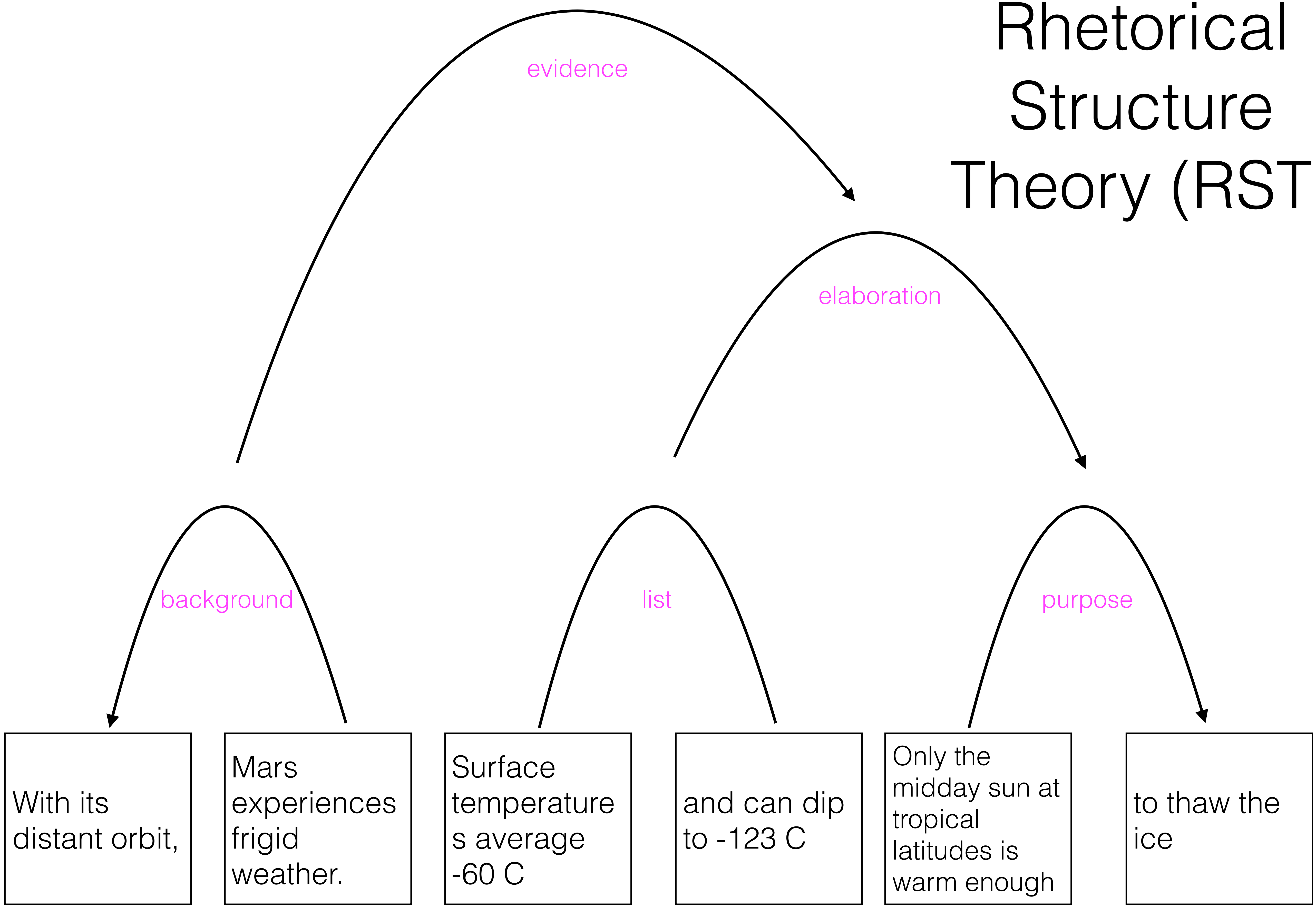
Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Rhetorical Structure Theory (RST)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Rhetorical Structure Theory (RST)



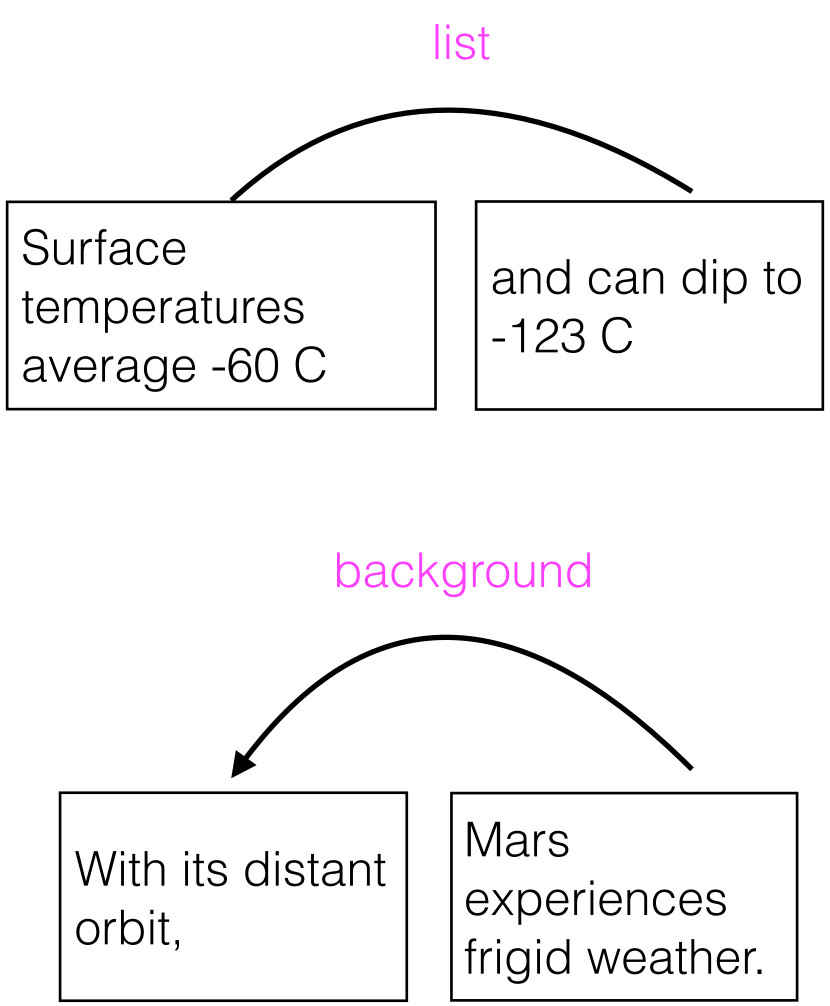
Arrows are directed from nucleus to satellite (undirected edges are nuclei)

RST parsing

- Shift-reduce (Marcu 1999, Sagae 2009)
- Neural-network shift reduce (Ji and Eisenstein 2014)
- Recursive neural network (Li et al. 2014)

RST parsing

stack

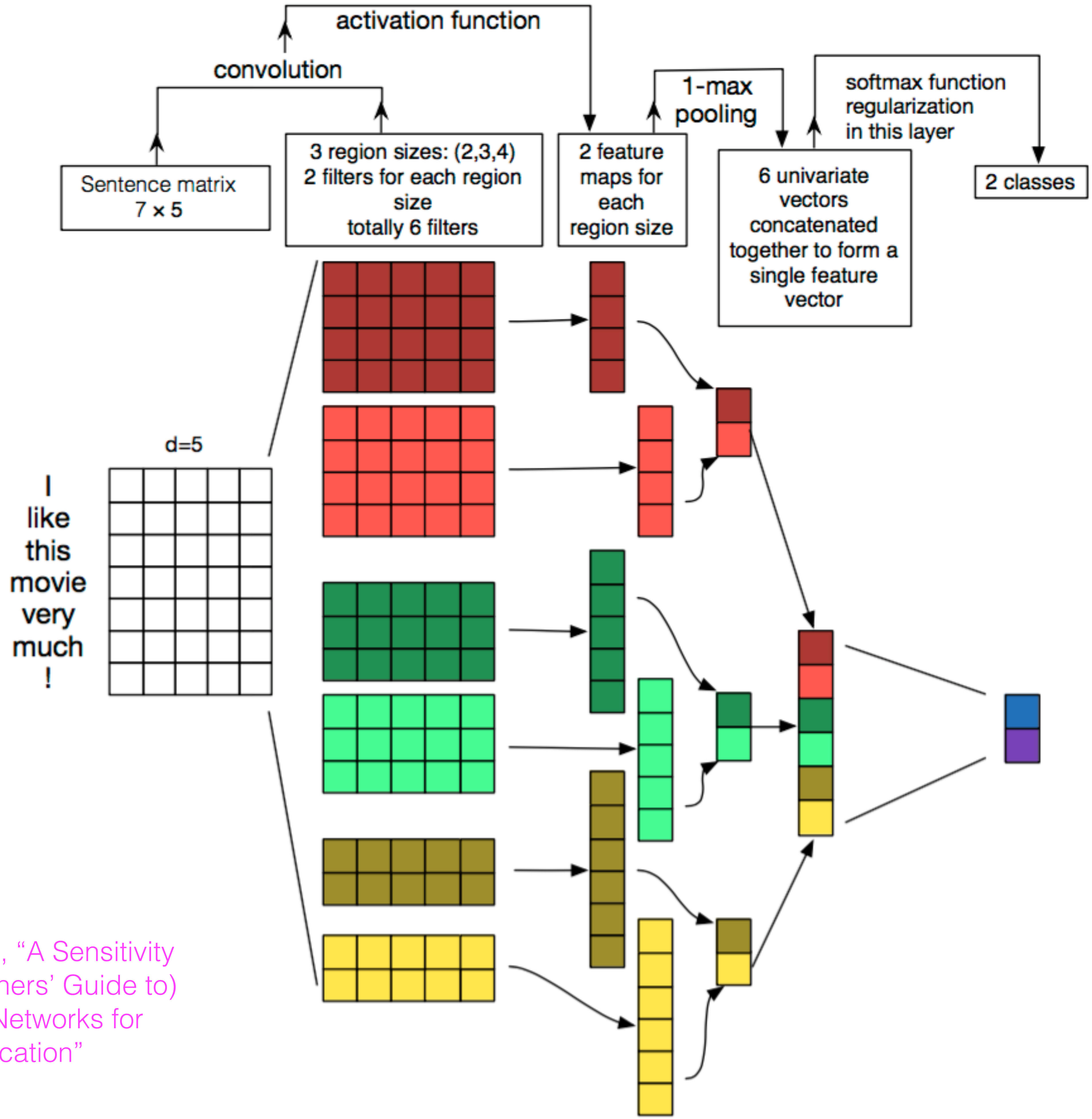


action

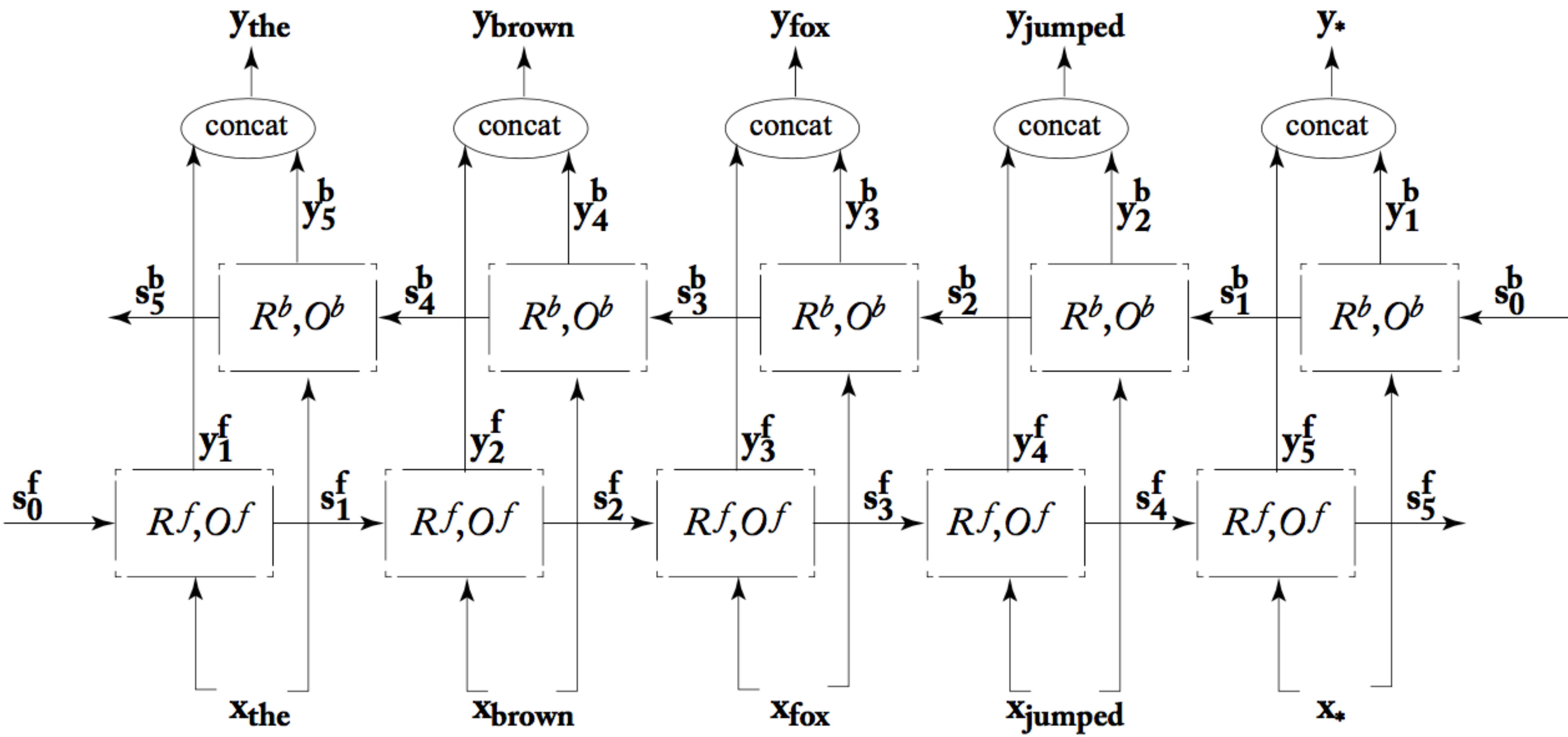
Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

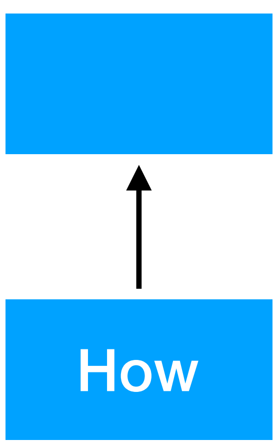
3. Neural methods are generally* better

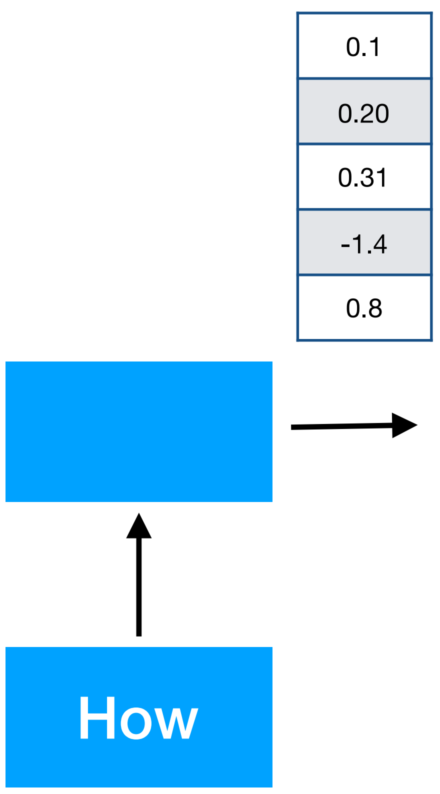


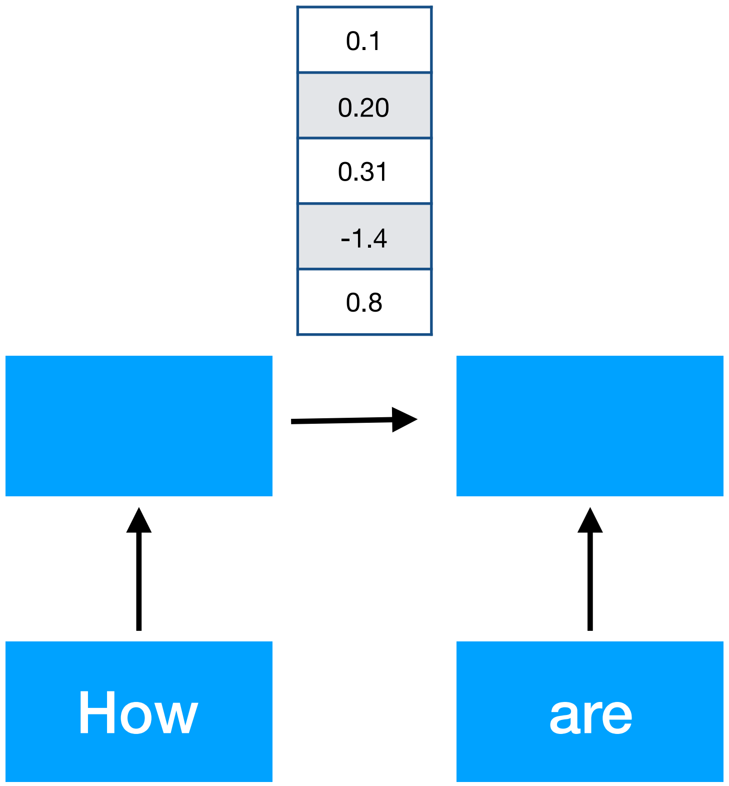
Zhang and Wallace 2016, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification"

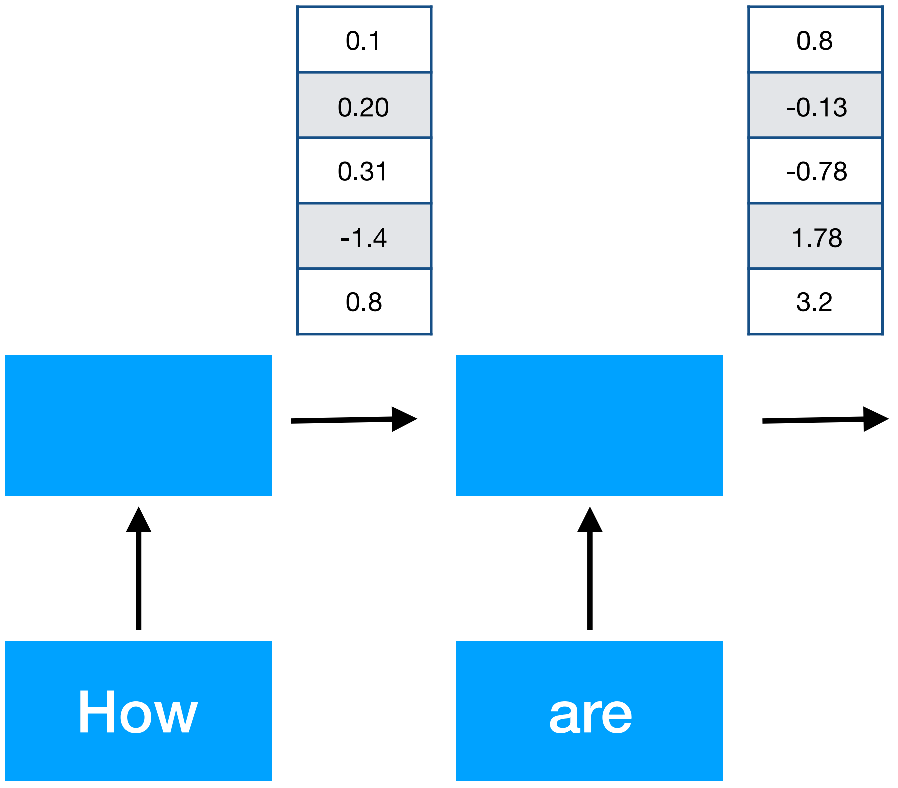


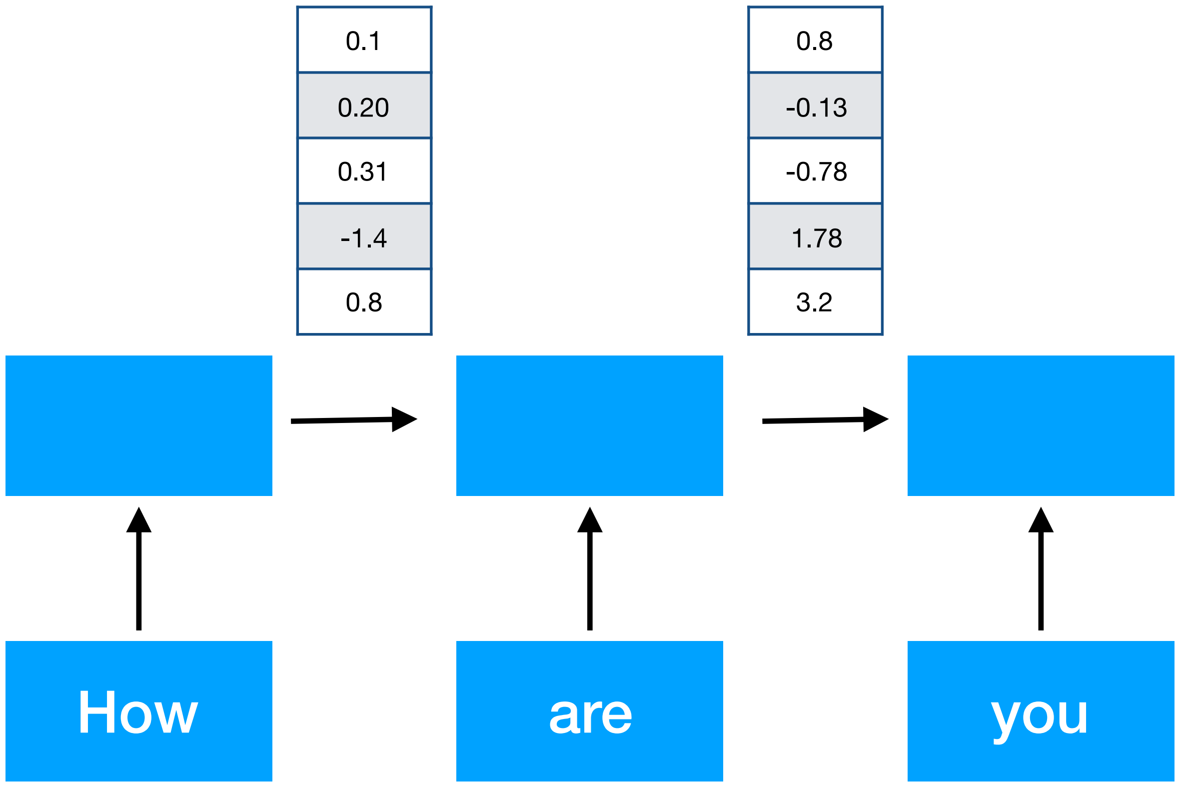
Goldberg 2017

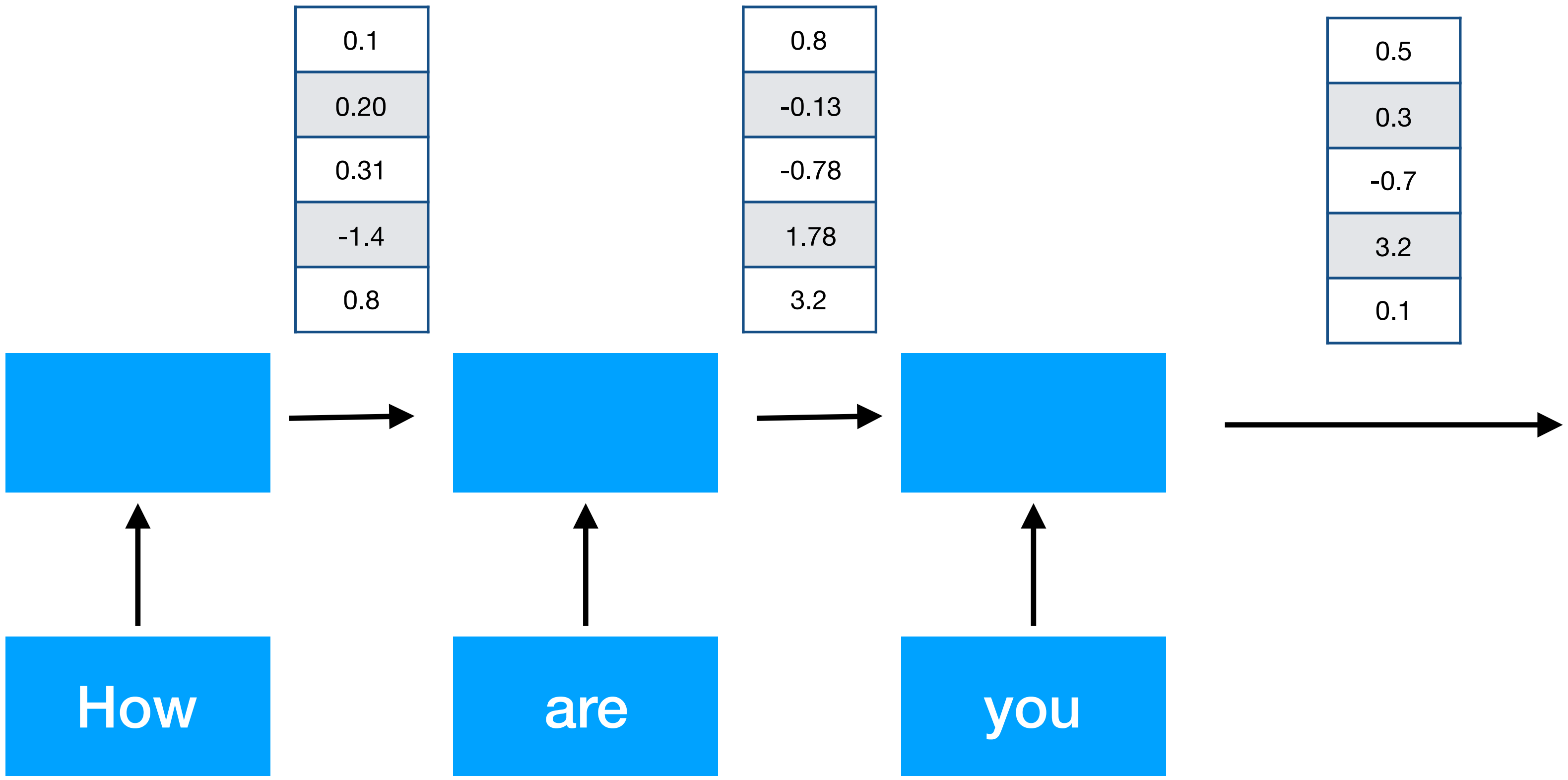

How

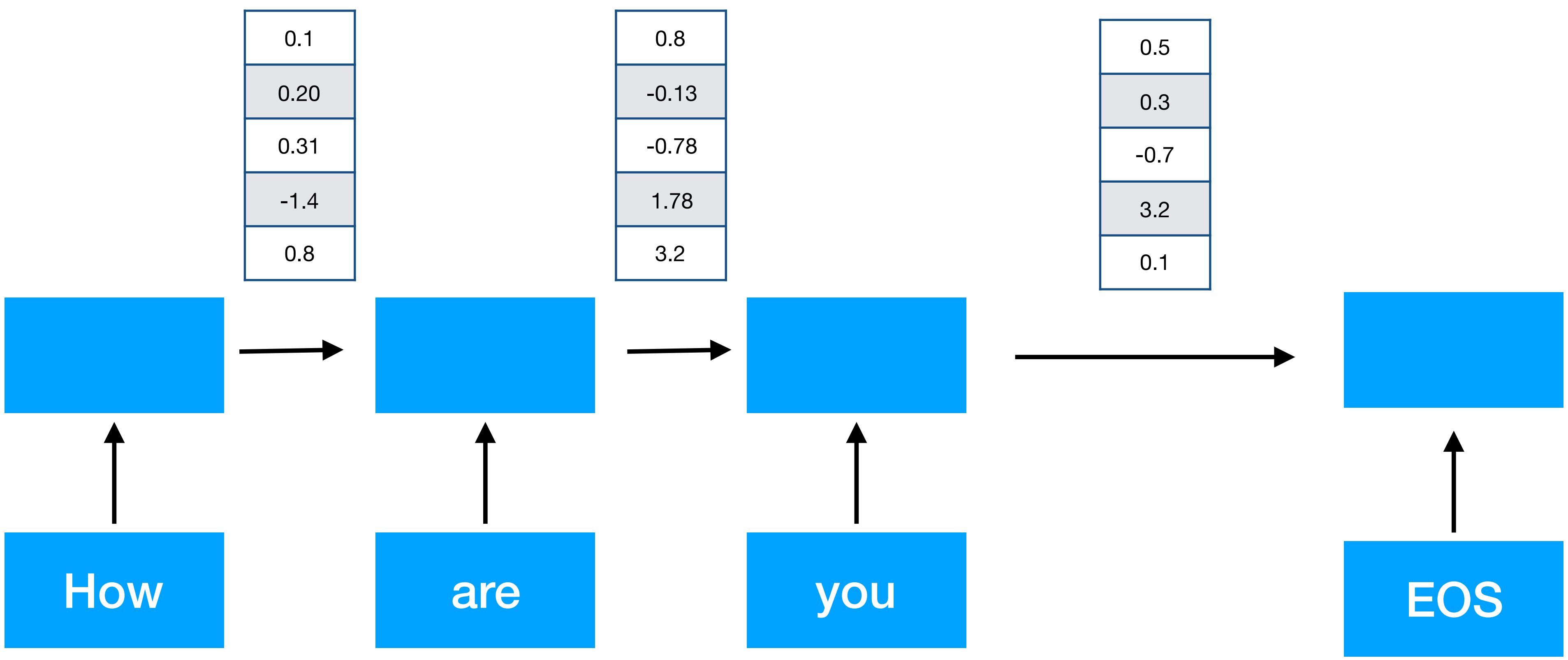

0.1
0.20
0.31
-1.4
0.8
How

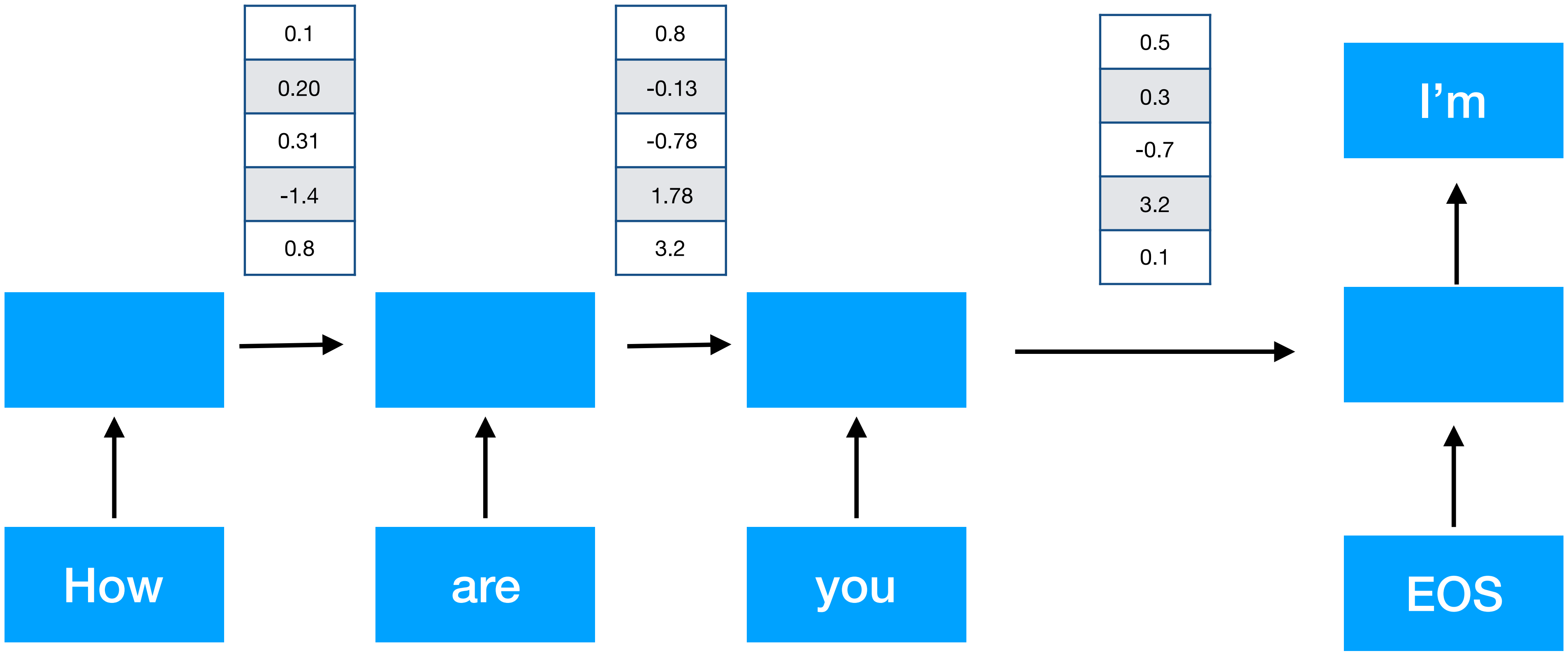

0.1
0.20
0.31
-1.4
0.8
How
are

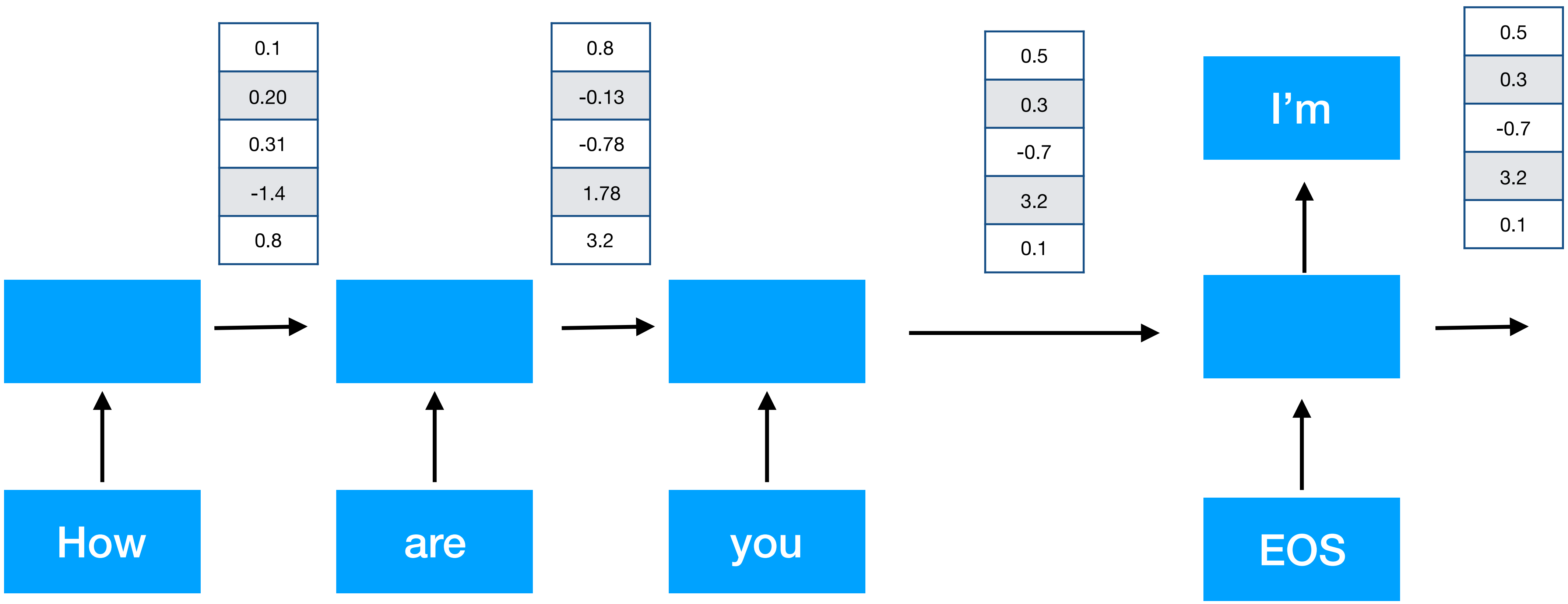

0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
How
are

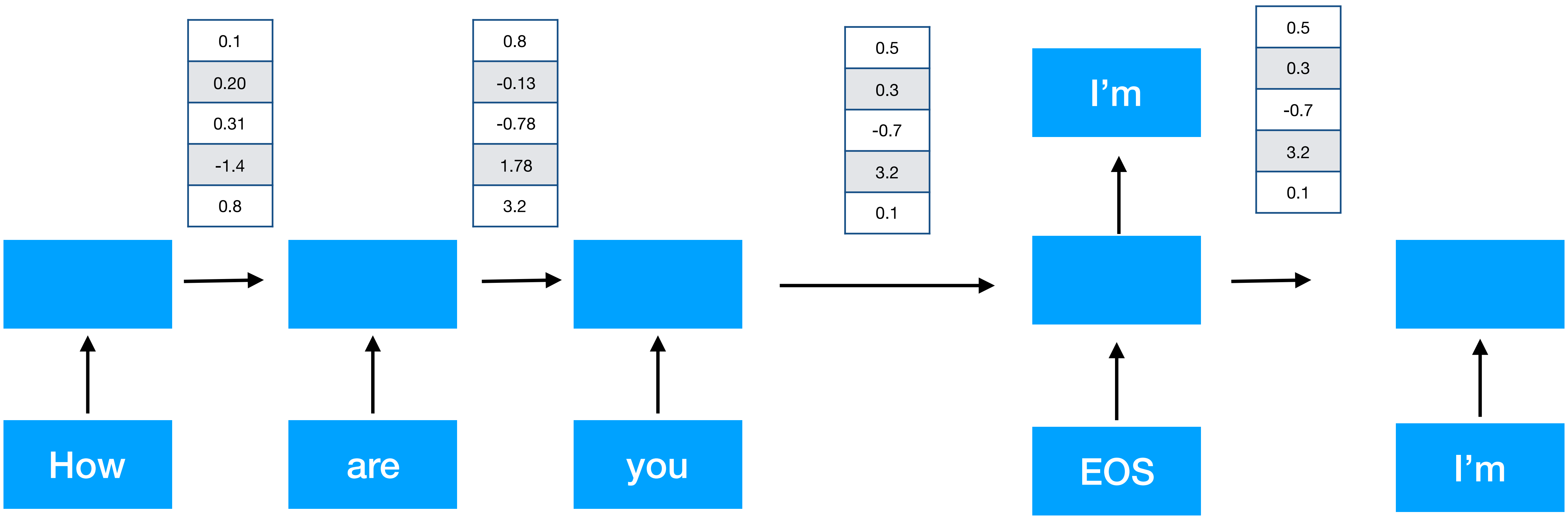

0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
How
are
you

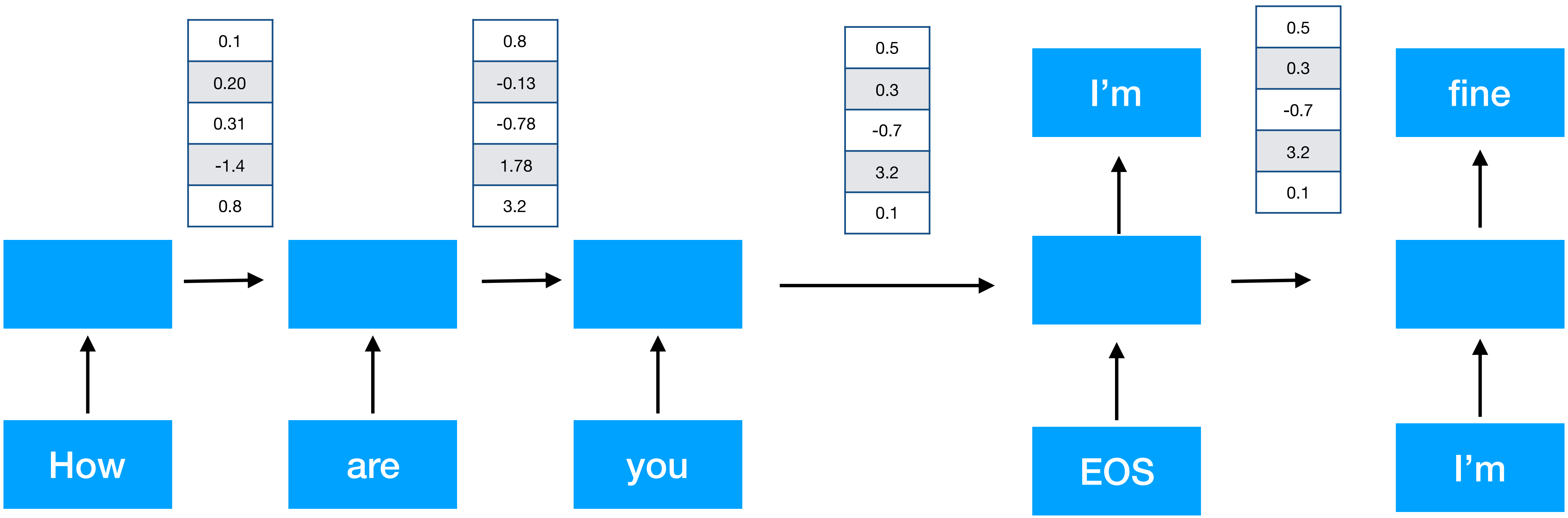

0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
How
are
you

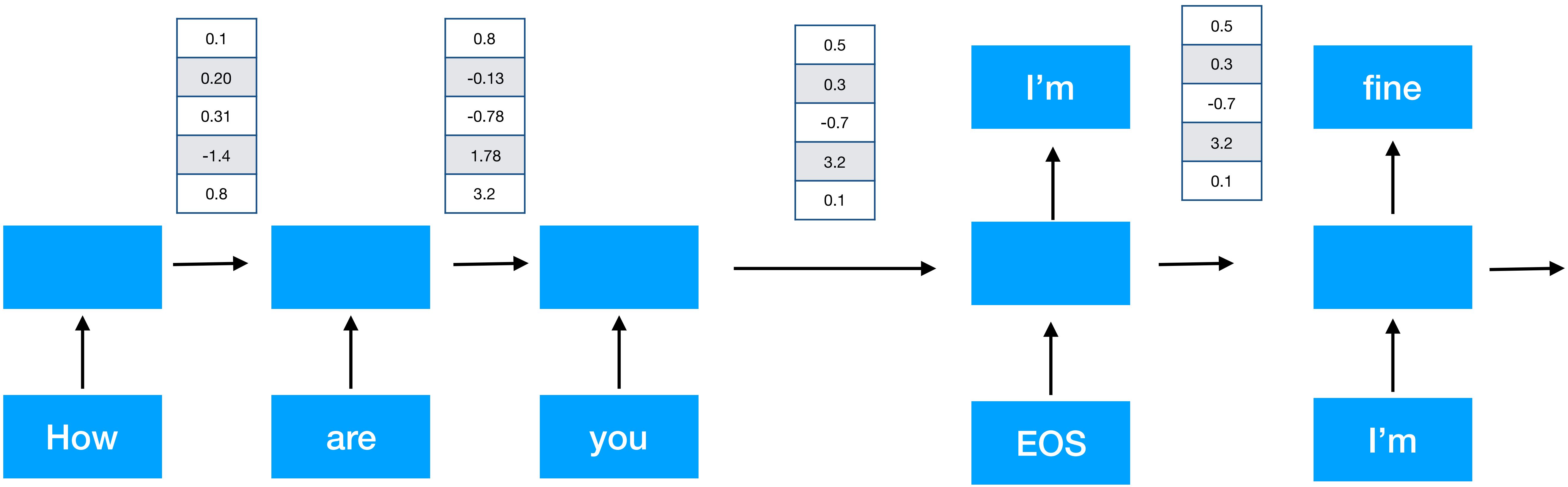

0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
How
are
you
EOS


0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
I’m
How
are
you
EOS


0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
I’m
0.5
0.3
-0.7
3.2
0.1
How
are
you
EOS


0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
I’m
0.5
0.3
-0.7
3.2
0.1
How
are
you
EOS
I’m


0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
I’m
0.5
0.3
-0.7
3.2
0.1
fine
How
are
you
EOS
I’m


0.1
0.20
0.31
-1.4
0.8
0.8
-0.13
-0.78
1.78
3.2
0.5
0.3
-0.7
3.2
0.1
I’m
0.5
0.3
-0.7
3.2
0.1
fine
How
are
you
EOS
I’m

Lexical semantics

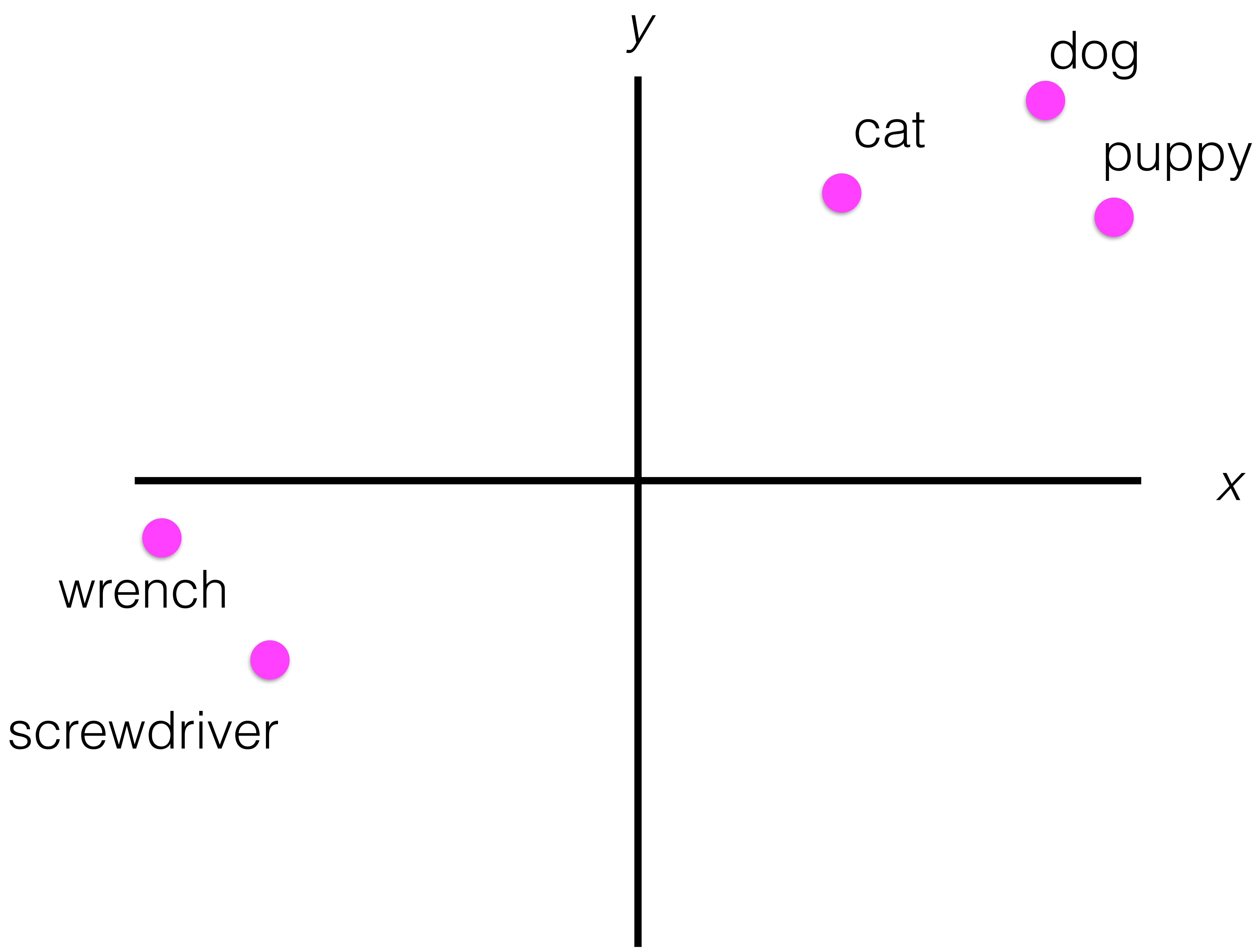
"You shall know a word by the company it keeps"

[Firth 1957]

Distributed representation

- Vector representation that encodes information about the distribution of contexts a word appears in

- Words that appear in similar contexts have similar representations (and similar meanings, by the distributional hypothesis).


y
dog
cat
puppy
x
wrench
screwdriver

Natural language inference

A man inspects the uniform of a figure in some East Asian country.	**contradiction** C C C C C	The man is sleeping
An older and younger man smiling.	**neutral** N N E N N	Two men are smiling and laughing at the cats playing on the floor.
A black race car starts up in front of a crowd of people.	**contradiction** C C C C C	A man is driving down a lonely road.
A soccer game with multiple males playing.	**entailment** E E E E E	Some men are playing a sport.
A smiling costumed woman is holding an umbrella.	**neutral** N N E C N	A happy woman in a fairy costume holds an umbrella.

Bowman et al. 2016

4. Evaluation is critical

Experiment design

	training	development	testing
size	80%	10%	10%
purpose	training models	model selection	evaluation; never look at it until the very end

Metrics

Metrics

- Perplexity

Metrics

- Perplexity
- Accuracy

Metrics

- Perplexity
- Accuracy
- Precision/recall/F1

Metrics

- Perplexity
- Accuracy
- Precision/recall/F1
- Parseval LAS/UAS (P/R/F1 over labeled constituents)

Metrics

- Perplexity
- Accuracy
- Precision/recall/F1
- Parseval LAS/UAS (P/R/F1 over labeled constituents)
- Correlation with human judgments

Metrics

- Perplexity
- Accuracy
- Precision/recall/F1
- Parseval LAS/UAS (P/R/F1 over labeled constituents)
- Correlation with human judgments
- B^3

Metrics

- Perplexity
- Accuracy
- Precision/recall/F1
- Parseval LAS/UAS (P/R/F1 over labeled constituents)
- Correlation with human judgments
- B^3
- BLEU

What's on the midterm

What’s on the midterm

- Anything from the beginning of the class to today is fair game topic-wise

What’s on the midterm

- Anything from the beginning of the class to today is fair game topic-wise
- Everything asked about can be learned from reviewing slides + assigned textbook readings

What’s on the midterm

- Anything from the beginning of the class to today is fair game topic-wise
- Everything asked about can be learned from reviewing slides + assigned textbook readings
- 4 open-ended questions (1 of which will relate to ethics)

What’s on the midterm

- Anything from the beginning of the class to today is fair game topic-wise
- Everything asked about can be learned from reviewing slides + assigned textbook readings
- 4 open-ended questions (1 of which will relate to ethics)
- ~3 short questions that will test your understanding

Strategies for the midterm

Strategies for the midterm

- Think about what concepts, methods, skills, and topics relate to the question

Strategies for the midterm

- Think about what concepts, methods, skills, and topics relate to the question
- If you have a clear area, go to work

Strategies for the midterm

- Think about what concepts, methods, skills, and topics relate to the question
- If you have a clear area, go to work
- Otherwise, review that week's readings for inspiration

Strategies for the midterm

- Think about what concepts, methods, skills, and topics relate to the question
- If you have a clear area, go to work
- Otherwise, review that week's readings for inspiration
- If you have no idea what to do with the question, skim some of the slides (or chapter introductions) to see how the content relates to the question


702
702

Actual questions from previous midterms

Break-out discussion

- Question is shared in Zoom chat
- Everyone opens question
- Breakout discussion for 5 minutes
- Class discussion based on groups
- Last year's midterm is posted to Canvas after