

# Ling/CSE 472: Introduction to Computational Linguistics

---

5/18/17

Meaning representation

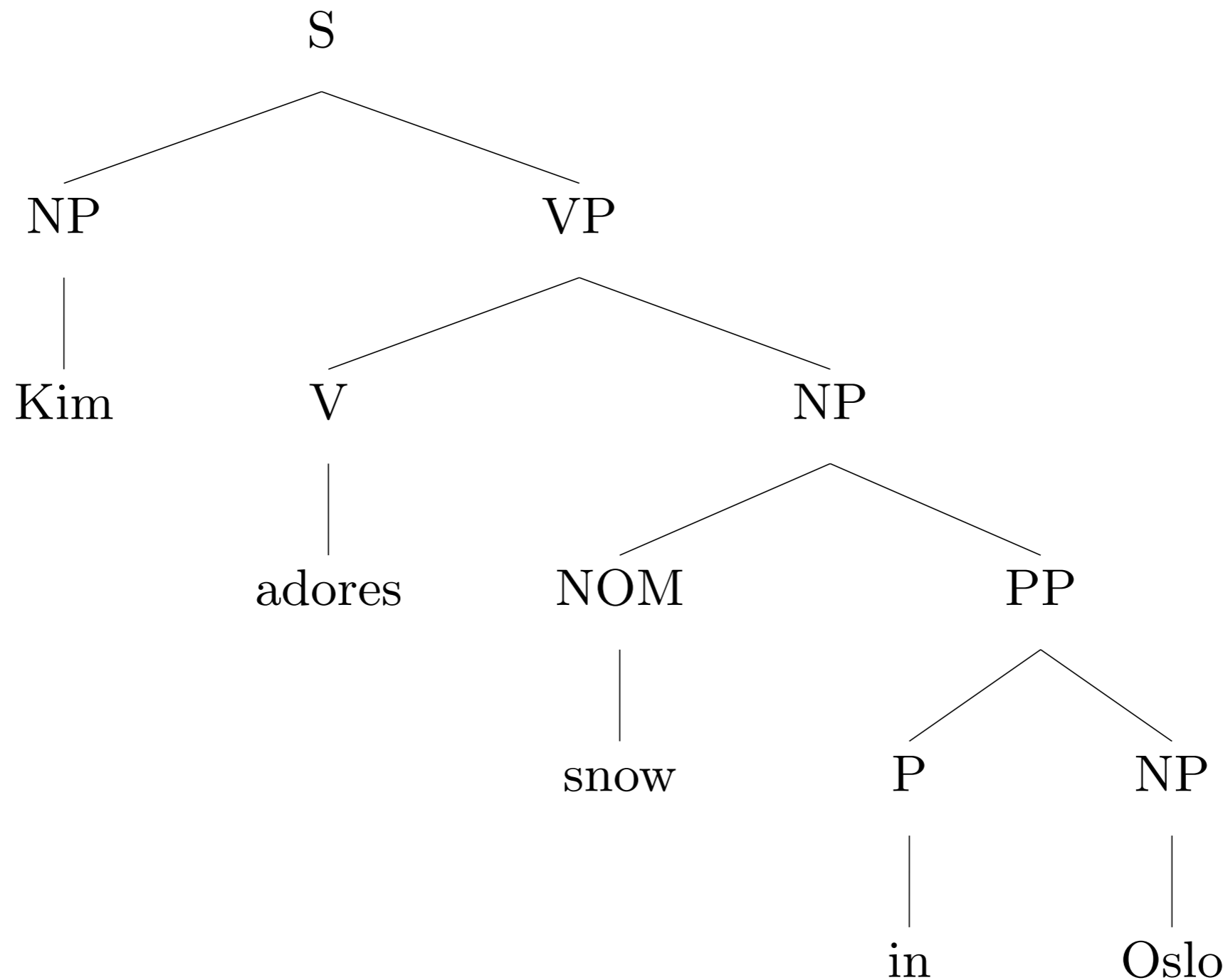
# Overview

---

- Semantics
- Semantics in NLP
- Scheduling term project presentations
- Reading questions

Parsing makes explicit inherent structure.  
So, does this tree represent meaning?

---



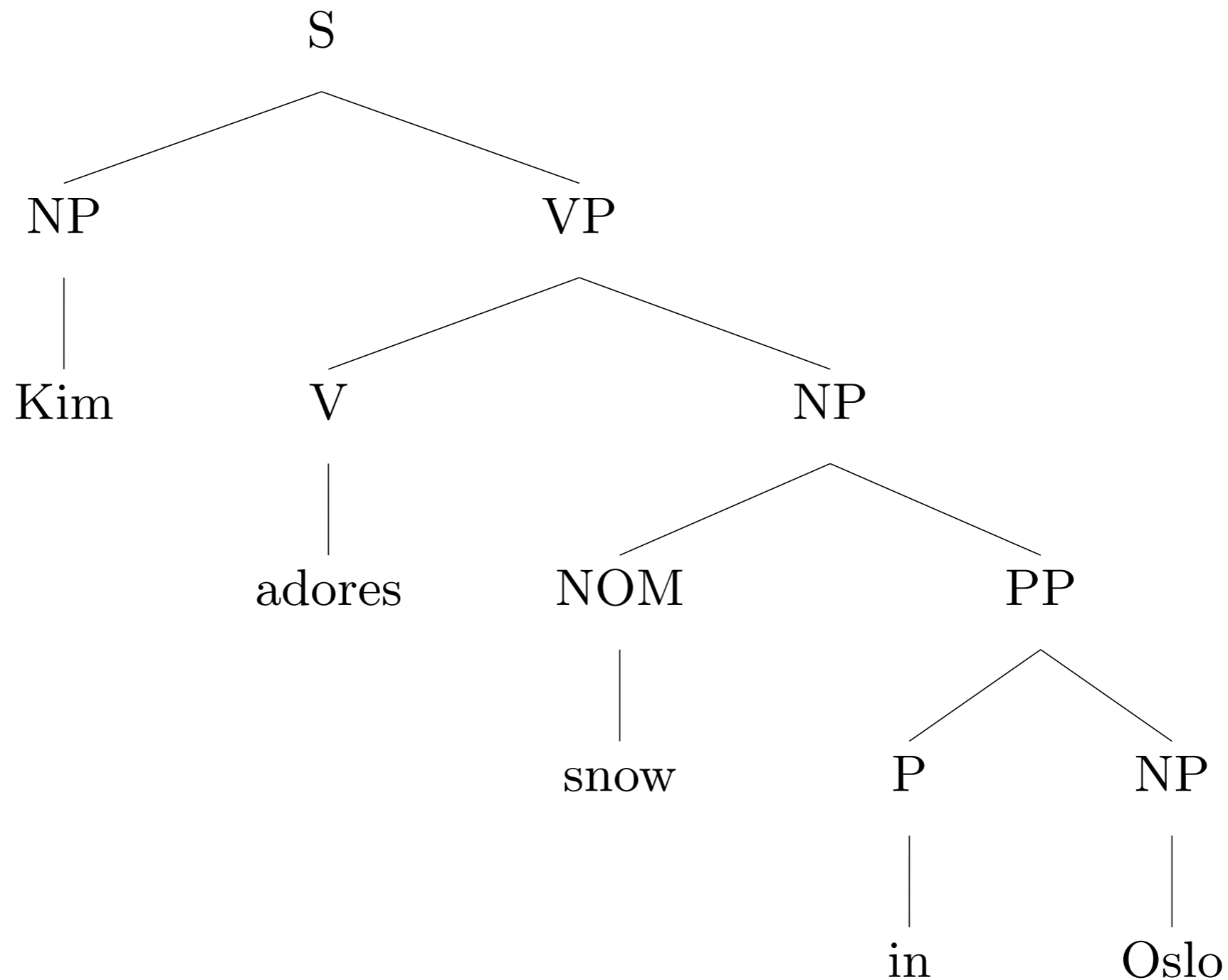
# Why represent semantics?

---

- When “earlier” levels aren’t enough
- Bridge between linguistics and real world items/models

How could we put this tree in correspondence to a model of the world?

---



# Semantics

---

- Create representations which can be put in correspondence with models of the world
- ... and which can be built compositionally via parsing

# Basic model-theoretic semantics

---

- Create a model of the world, consisting of elements, sets of elements and relations
- Create an interpretation function which maps linguistic elements (parts of the semantic structure) to parts of the model
- Simple propositions are interpreted by checking their truth in the model
- Define semantics for “logical vocabulary”: *and, or, not, if, every, some, ....*

# Model theoretic semantics example

---

• Entities: Joey:



Fluffy:



Tiger:



• Properties: calm: {



}; angry: {



}

• Relations: knows: {



> , <




> }



# Model theoretic semantics example: denotations

---

- $[[\text{Fluffy}]] =$  

- $[[\text{angry}]] = \{ x \mid x \text{ is angry} \} = \{$    $\}$

- $[[\text{Fluffy is angry}]] = \text{True}$  *iff* the entity denoted by *Fluffy* is in the set denoted by *angry*
- Compositionality: The process of determining the truth conditions of *Fluffy is angry* based on the denotations of its parts and its syntactic structure

# Logical vocabulary gets special treatment

---

- *Fluffy is angry and Joey is not angry.*
  - What does *and* mean? (How does it affect the truth conditions of the whole?)
  - What does *not* mean?
- *Every cat is angry.*
  - What does *cat* mean? (Is this a logical operator?)
  - What does *every* mean?
- Is the division into logical and non-logical vocabulary an inherent property of language or an artifact of the system of meaning representation?

# More on quantifiers

---

- The semantic type of a quantifier is a relation between sets, called the *restriction* and *body* (or *scope*) of the quantifier
  - $[[\text{every}]] \{ \langle P, Q \rangle \mid P \subseteq Q \}$
  - $[[\text{every cat is angry}]]$  is True *iff*  $\{ x \mid x \text{ is a cat} \} \subseteq \{ y \mid y \text{ is angry} \}$
  - $[[\text{some}]] \{ \langle P, Q \rangle \mid P \cap Q \neq \emptyset \}$
  - $[[\text{some cat is angry}]]$  is True *iff*  $\{ x \mid x \text{ is a cat} \} \cap \{ y \mid y \text{ is angry} \} \neq \emptyset$
- Where do those sets come from?

# Why represent semantics?

---

- When “earlier” levels aren’t enough
- Bridge between linguistics and real world items/models

# Semantics in NLP

---

- Construct knowledge base or model of the world
- Extract meaning representations from linguistic input
- Match input to world knowledge
- Produce replies/take action on the basis of the results
  
- In what other cases might semantic representations be useful?

# Semantics in NLP

---

- In what other cases might semantic representations be useful?
  - Transfer-based MT
  - Building a knowledge base by “reading” the web (or wikipedia or...)
  - Generation

# Semantic representations: Desiderata (Jurafsky & Martin)

---

- Verifiability: We must be able to compare the representation to a knowledge base
- Lack of ambiguity: A semantic representation should have just one interpretation
- Canonical form: A given interpretation should have just one representation
  - Does Maharani have vegetarian dishes?
  - Do they have vegetarian food at Maharani?
  - Are vegetarian dishes served at Maharani?
  - Does Maharani have vegetarian fare?
  - But not: Can vegetarians eat at Maharani?
- Expressiveness: Must be able to adequately represent a wide range of expressions

# Semantic Representations: Desiderata (Copestake et al 2005)

---

- Expressive Adequacy: The framework must allow linguistic meanings to be expressed correctly
- Grammatical Compatibility: Semantic representations must be linked clearly to other kinds of grammatical information (most notably syntax)
- Computational Tractability: It must be possible to process meanings and to check semantic equivalence and to express relationships between semantic representations straightforwardly
- Underspecifiability: Semantic representations should allow underspecification (leaving semantic distinctions unresolved), in such a way as to allow flexible, monotonic resolution of such partial semantic representations



# Evaluation slide

---

- How would we evaluate a system of semantic representations?
- How would we evaluate a parsing system which produces semantic representations from input?
  - What's the gold standard?
  - What's the baseline?
  - What are the metrics?
  - What else might we need?

# Overview

---

- Semantics
- Semantics in NLP
- Scheduling term project presentations
- Reading questions

# Reading questions

---

- For a meaning representation, what is the standard of being sufficiently expressive?
- "First-Order Logic...is expressive enough to handle quite a lot of what needs to be represented." What are ways in which first-order logic is insufficiently expressive?
- This chapter mainly covers FOL, but what about the higher order logic? Is higher order logic useful or applicable in NLP?
- How does the semantic formula for quantifiers solve the ambiguities like:"everyone gets his money." Where it can mean that everyone gets his own money or there's a specific person whose money was gotten by everyone.

# Reading questions

---

- As far as I know, the first order logic predicates rely heavily on the domain of discourse. However, productivity is an important attribute of language, which means people make up words that never occur before. Wouldn't it be problematic to use first order logic to model semantics if the domain of discourse is constantly changing?
- Are the logical and non-logical vocabularies related to content vs function words, or something different?
- What do they mean by saying that the 'or' operator is not disjunctive in the same way that the English word or is (p573)?

# Reading questions

---

- Can you please go over how the Davidsonian event representations make it so that we can add as many rules as we want?
- Also, what is neo-Davidsonian?
- Following the formatting of 17.50, say we have the sentence *I answered emails then ate lunch at my desk* (where *at my desk* refers to both *answering emails* and *eating*). I take it *answering emails* and *eating* should be separate events, but is it possible to specify that they are related since they occur at the same place on the same day? Like an event corresponding to *at my desk* contains the events *answering emails* and *eating*?

# Reading questions

---

- For the sentence *I only have five dollars and I don't have a lot of time*, the semantic formula is:

$\text{Have}(\text{Speaker}, \text{FiveDollars}) \wedge \neg \text{Have}(\text{Speaker}, \text{LotOfTime})$

- Does order matter in the formula? Does the semantic mapping represent part of speech as well?

# Reading questions

---

- How does subsumption as related to semantics differ/compare to subsumption in the context of CFG?
- How are word embeddings like Word2vec related and different from semantic networks?

# Reading questions

---

- Can we fire whoever named "accomplishment" and "achievement?"
- Seriously though, naming similar classes like this makes things really confusing. I think I get the general idea, but I have another question that might clarify things a bit: can verbs fit into multiple classes, depending on how they're used, or are the classes not mutually exclusive?
- The specific example I had in mind was *reading a book*. *I stopped reading a book* implies that I was reading a book, suggesting that *reading a book* is an activity and not an accomplishment, but *I read a book in a minute* also makes sense, which is characteristic of activities rather than accomplishments. So, is *read a book* both an activity and an accomplishment, or is it changing between the two sentences?



# Reading questions

---

- In figure 17.77, when discussing activity expressions, the example sentence is *I live in Brooklyn for a month*. To me, this sentence doesn't seem grammatical, I would only use this in the past tense of *I lived in Brooklyn for a month*. or the future tense *I will live in Brooklyn for a month*. but it doesn't seem right to be in present tense. If I was trying to say I was in the month I would shift to *I am living in Brooklyn for a month*. Thoughts? Is this grammatically correct and I'm out in left field? Or do others have this native speaker judgement?
- The aspect section just describes verb aspects, but doesn't really connect the concept to how it is used in FOL or another meaning representation language. It almost seems like a non-sequitur (literary) in a chapter about semantics!