Semantic Parsing

Claudio Delli Bovi
Tommaso Pasini
Roberto Navigli

DIPARTIMENTO
DI INFORMATICA

SAPIENZA
UNIVERSITÀ DI ROMA

> “*The fish trap exists because of the fish. Once you've gotten the fish you can forget the trap. The rabbit snare exists because of the rabbit. Once you've gotten the rabbit, you can forget the snare. Words exist because of meaning. Once you've gotten the meaning, you can forget the words. Where can I find a man who has forgotten words so I can talk with him?*

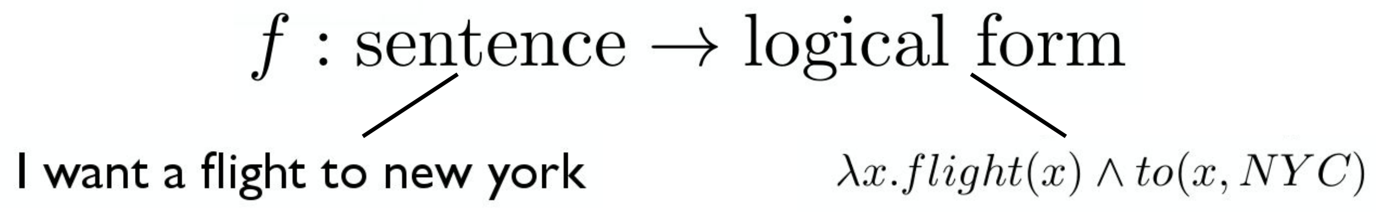
-- The Writings of Chuang Tzu, 4th century B.C.

荃者所以在魚，得魚而忘荃；蹄者所以在兔，得兔而忘蹄；言者所 以在意，得意而忘言。吾安得夫忘言之人而與之言哉！

《庄子・雜篇・外物第二十六》

Semantic Parsing: What is it exactly?

- Transforming Natural Language (NL) sentences into **computer executable** and **complete** Meaning Representations (MRs)



$$f : \text{sentence} \rightarrow \text{logical form}$$

I want a flight to new york

$$\lambda x.flight(x) \wedge to(x, NYC)$$

- MRs are fully formal languages that:
  - Have a rich **ontology** of types, properties, and relations
  - Support **automated reasoning** and/or **execution**

- Semantic parsers (especially earlier ones) are often designed with some **application domain** in mind

Semantic Parsing: Application Domains/Benchmarks

- **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?” ⇨

```
Show: (Flight-Number)
Origin: (City “Cleveland”)
Destination: (City
“Dallas”)
```

⇨ NA 1439, TQ 23, …

- **CLang: Robocup Coach Language**
  - Coaching instructions to simulated players are given in a language called CLang (Chen et al., 2003)

“If the ball is in our goal area then player 1 should intercept it” ⇨

```
(bpos (goal-area our) (do our {1} intercept))
```

- **GeoQuery: A Database Query Application**
  - Query application for U.S. geography database (Zelle and Mooney, 1006)

“Which rivers run through the states bordering Texas?” ⇨

```
answer(traverse(next_to(stateid(‘texas’))))
```

Differences with other NLP tasks

- “**Shallow**” semantic processing:
    - Information Extraction
    - Semantic Role Labeling

Semantic parsing involves **deeper** semantic analysis to understand the whole sentence

- **Intermediate** linguistic representations:
    - Part-of-speech Tagging
    - Syntactic Parsing
    - Semantic Role Labeling

Semantic parsing generates a “**final**” representation that assumes no further processing

- Output meant for **humans**:
    - Question Answering
    - Text Summarization
    - Machine Translation

Semantic parsing outputs a formal language **for computers to read**, with no room for implicit/incomplete output

Relations with other NLP tasks

- Tasks being performed **within** semantic parsing:
    - Word Sense Disambiguation
    - Syntactic Parsing

- Tasks **closely related** to semantic parsing:
    - Natural Language Generation
    - Machine Translation

Reversing a semantic parsing system yields a natural language generation system (Wong and Mooney, 2007)

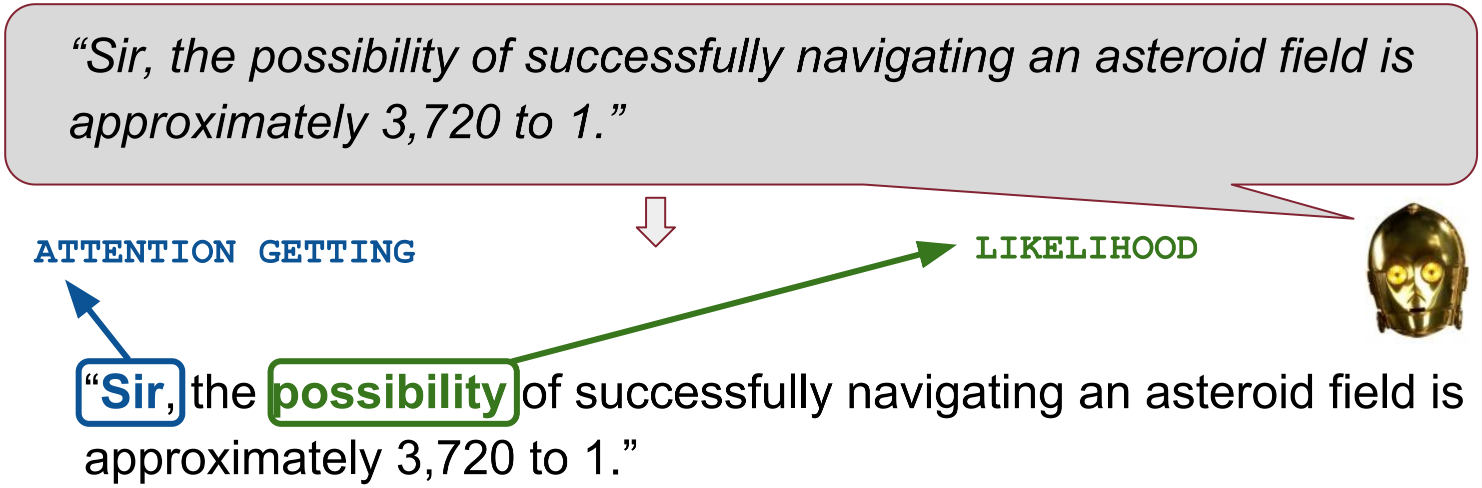
Any MR language can be looked upon as just another NL language! (**more on this later**)

Outline

- **Frame-semantic Parsing**
- **Supervised Semantic Parsing**
- **Semantic Parsing with CCG**
- **Unsupervised Semantic Parsing**
- **Semi-Supervised Semantic Parsing**
- **Learning from Q&A pairs**
- **Semantic Parsing with AMR** (latest trend!)

Frame-semantic Parsing

- Given a text/sentence, analyze its **frame semantics:**
  - Identify the **frame(s)** evoked by the sentence
  - Match **frame elements** with their realizations



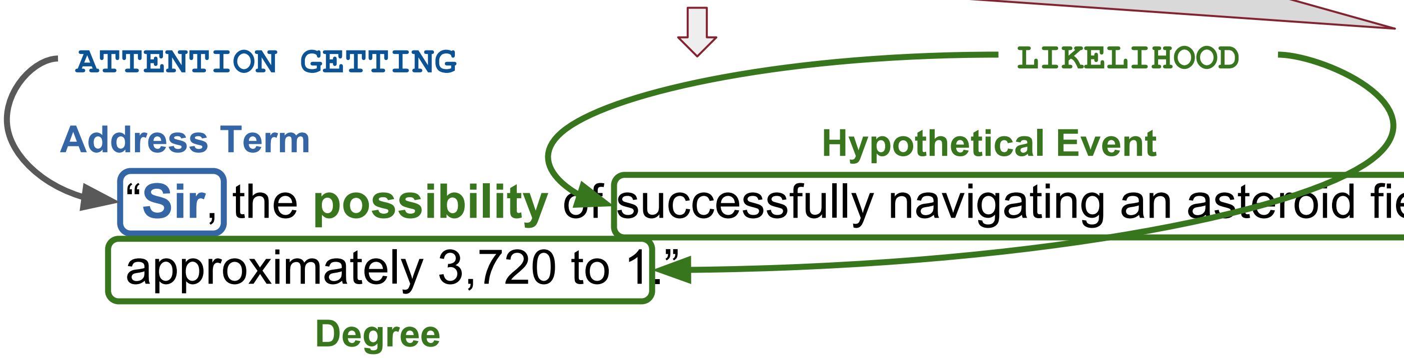
“Sir, the possibility of successfully navigating an asteroid field is approximately 3,720 to 1.”

“**Sir**, the **possibility** of successfully navigating an asteroid field is approximately 3,720 to 1.”

Frame-semantic Parsing

- Given a text/sentence, analyze its **frame semantics:**
  - Identify the **frame(s)** evoked by the sentence
  - Match **frame elements** with their realizations

“Sir, the possibility of successfully navigating an asteroid field is approximately 3,720 to 1.”



Frame-semantic Parsing

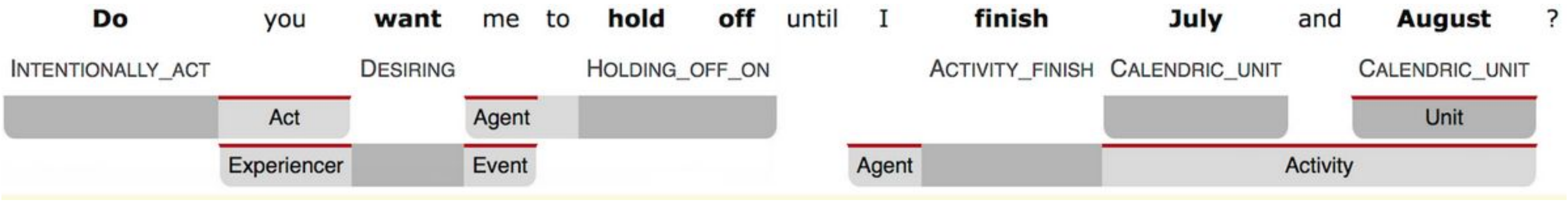
- Given a text/sentence, analyze its **frame semantics:**
  - Identify the **frame(s)** evoked by the sentence
  - Match **frame elements** with their realizations

- Strictly speaking, frame-semantic parsing is **not** semantic parsing!
  - Frame semantics does not yet provide a **formal meaning representation** (frames have human-level descriptions!)
  - On the practical ground, it can be seen as an “enhanced” Semantic Role Labeling task (Gildea and Jurafsky, 2002)

Frame-semantic Parsing: SEMAFOR (Das et al. 2012)

- **SEMantic Analyzer of Frame Representations:** State-of-the-art frame-semantics parser, discriminatively trained on the full-text annotated sentences in FrameNet

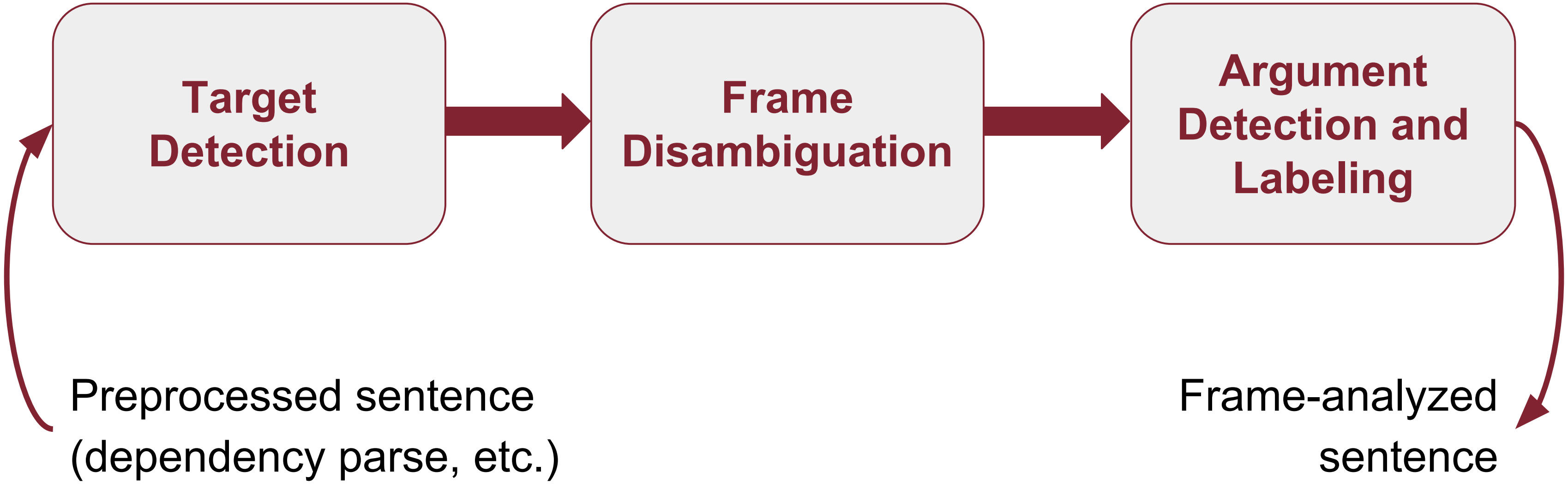
- **(Semi)supervised model** based on features over observable parts of the sentence (words, lemmas, POS-tags, dependency edges, …)



Open-source software available at: http://www.cs.cmu.edu/~ark/SEMAFOR/

Frame-semantic Parsing: SEMAFOR (Das et al. 2012)

- **Three-stages pipeline:**



Target Detection

Frame Disambiguation

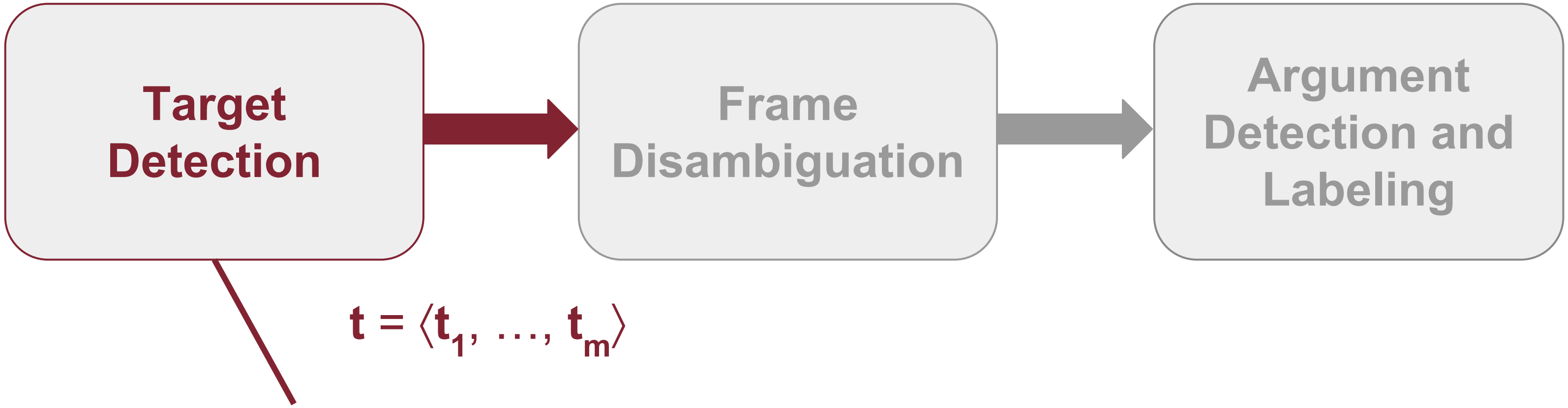
Argument Detection and Labeling

Preprocessed sentence (dependency parse, etc.)

Frame-analyzed sentence

Frame-semantic Parsing: SEMAFOR (Das et al. 2012)

- **Three-stages pipeline:**



Target Detection

Frame Disambiguation

Argument Detection and Labeling

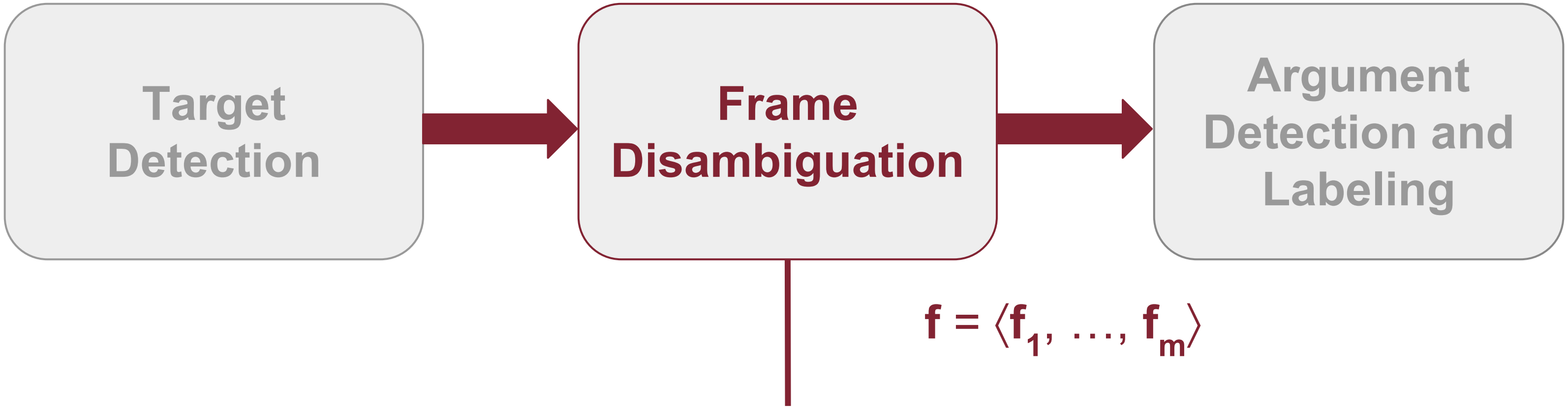
$\mathbf{t} = \langle t_1, \ldots, t_m \rangle$

Identify the target frame-evoking elements (**FEEs**):

- Whitelist + small set of rules based on POS criteria (Johansson and Nugues, 2007)

Frame-semantic Parsing: SEMAFOR (Das et al. 2012)

- **Three-stages pipeline:**



Target Detection

Frame Disambiguation

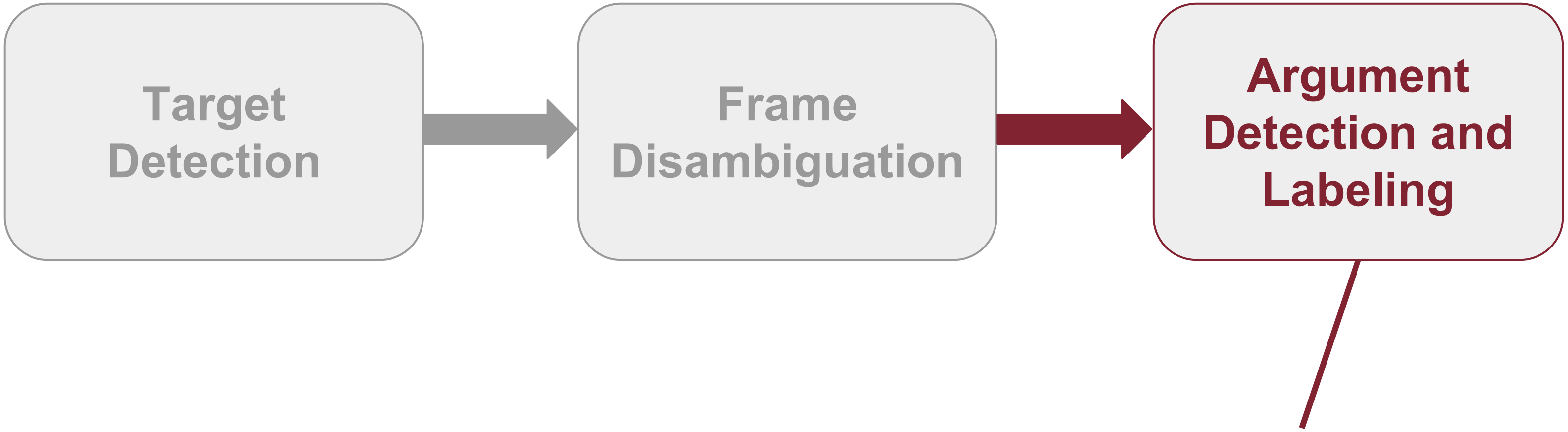
Argument Detection and Labeling

$\mathbf{f} = \langle \mathbf{f}_1, \ldots, \mathbf{f}_m \rangle$

Conditional log-linear model to identify the set of frames **$\mathbf{f} = \langle \mathbf{f}_1, \ldots, \mathbf{f}_m \rangle$** from the targets, trained to maximized trained data log-likelihood on a frame-annotated corpus (SemEval 2007, FrameNet 1.5 full-text annotations)

Frame-semantic Parsing: SEMAFOR (Das et al. 2012/2014)

- **Three-stages pipeline:**



Target Detection

Frame Disambiguation

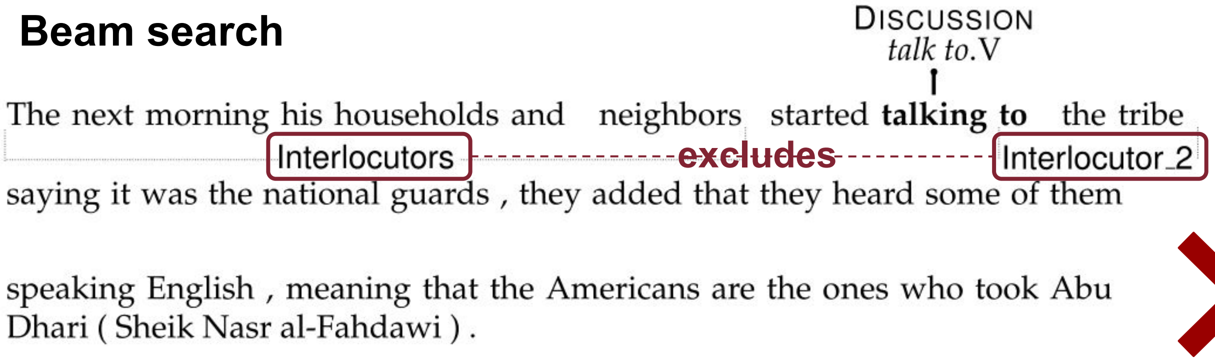
Argument Detection and Labeling

Conditional log-linear model to map, for each frame $\mathbf{f_i}$, a subset of the set of roles $\mathbf{R_i}$ = { $\mathbf{r_1}$, …, $\mathbf{r_{|fi|}}$ } to spans of the input sentence + **joint decoding/inference using beam search** (approximate) or **dual decomposition** (exact)

Frame-semantic Parsing: SEMAFOR (Das et al. 2012/2014)

- **Frame Element constraints (example):**

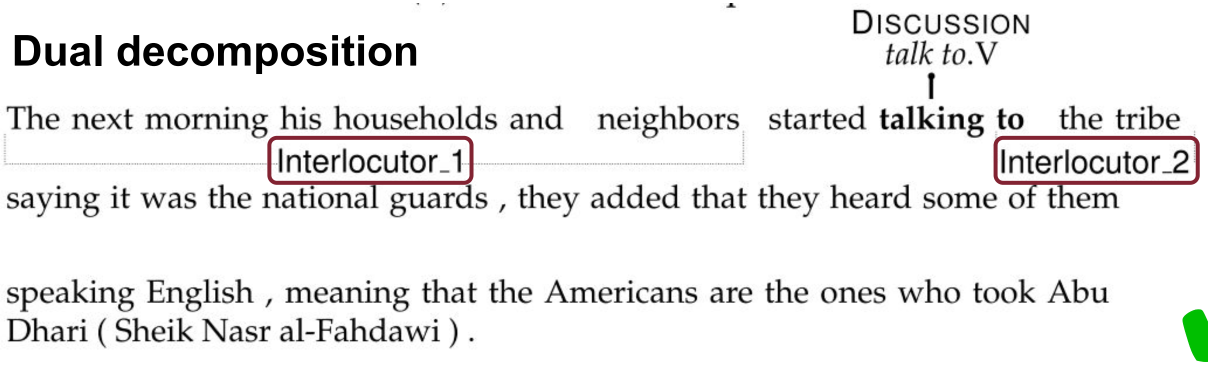
Beam search



The next morning his households and neighbors started **talking to** the tribe saying it was the national guards , they added that they heard some of them speaking English , meaning that the Americans are the ones who took Abu Dhari (Sheik Nasr al-Fahdawi) .

✗

Dual decomposition



The next morning his households and neighbors started **talking to** the tribe saying it was the national guards , they added that they heard some of them speaking English , meaning that the Americans are the ones who took Abu Dhari (Sheik Nasr al-Fahdawi) .

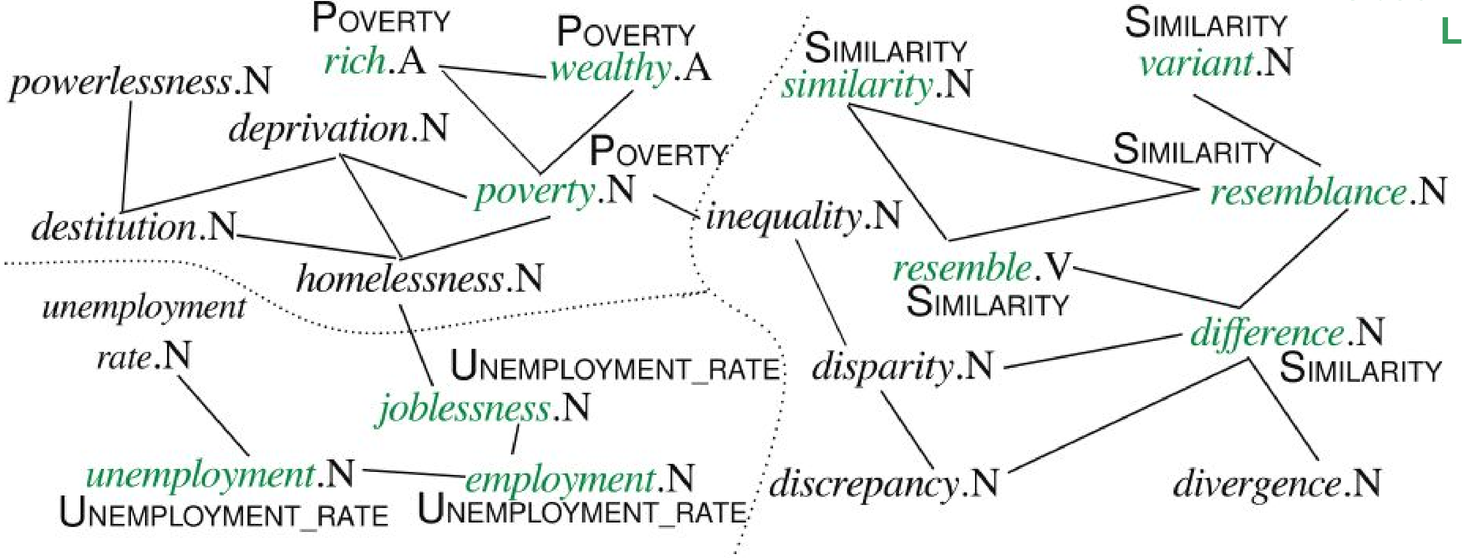
✓

Frame-semantic Parsing: Unknown Predicates

- **Problem: sparseness/lack of labeled data**
  Many frame-evoking predicates are seen neither in **lexicon** nor in **training data**!

Frame-semantic Parsing: Unknown Predicates

- **Solution:**
  Propagate frame labels from known predicates to unknown predicates in a **similarity graph** (Das et al., 2014)



Observed LUs

Observed LUs

Frame-semantic Parsing: Recent Advances

- **Frame Embeddings (Hermann et al., 2014):**
  Starting from frame-annotated data, learn an embedding model that projects the set of word representations for the **syntactic context** around a predicate to a low-dimensional representation

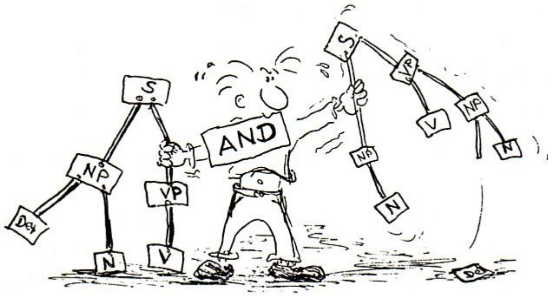
⇨ **Frame identification in the vector space (e.g. cosine similarity)**

- **Dynamic Programming (Täckström et al., 2015):**
  A frame's arguments should not overlap, but this means **classification decisions are not independent**: use a dynamic program to label Frame Elements (Google's variant of SEMAFOR)

Outline

- Frame-semantic Parsing
- Supervised Semantic Parsing
- Semantic Parsing with CCG
- Unsupervised Semantic Parsing
- Semi-Supervised Semantic Parsing
- Learning from Q&A pairs
- Semantic Parsing with AMR (latest trend!)

Supervised Semantic Parsing: Syntax-based Approaches



Intuition:

- Semantic parsing is a compositional process
- Sentence structures are key in building MRs

⇩

Syntax-based approach:

meaning composition follows the tree structure of a syntactic parse (meaning of a **constituent** from the meaning of its **sub-constituents**)

Hand-built approaches (Warren and Pereira, 1982)

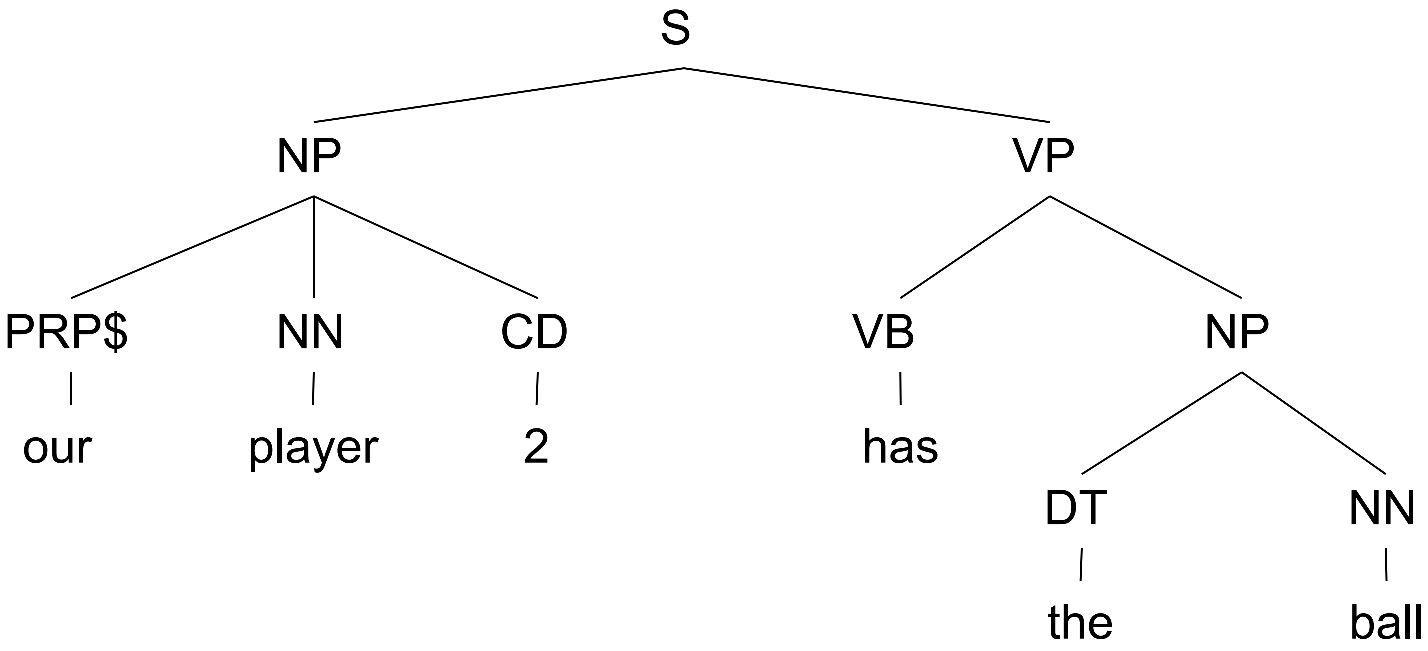
Learning approaches (Tang and Mooney, 2001; Kate and Mooney, 2006)

Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **SCISSOR (Semantic Composition that Integrates Syntax and Semantics to get Optimal Representations):**
    - Allows both **syntax** and **semantics** to be used simultaneously to obtain a syntactic-semantic analysis
    - Based on a statistical parser to generate a **semantically-augmented parse tree (SAPT)**

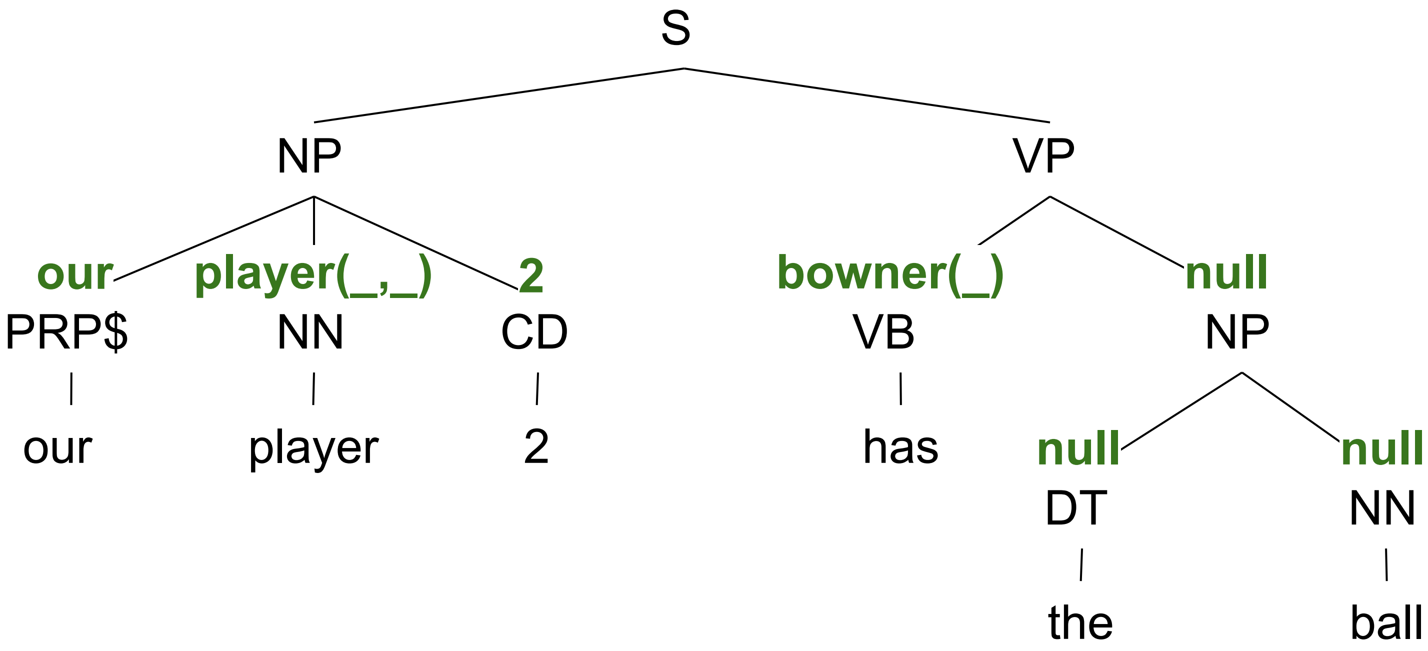
Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **SAPT example:**



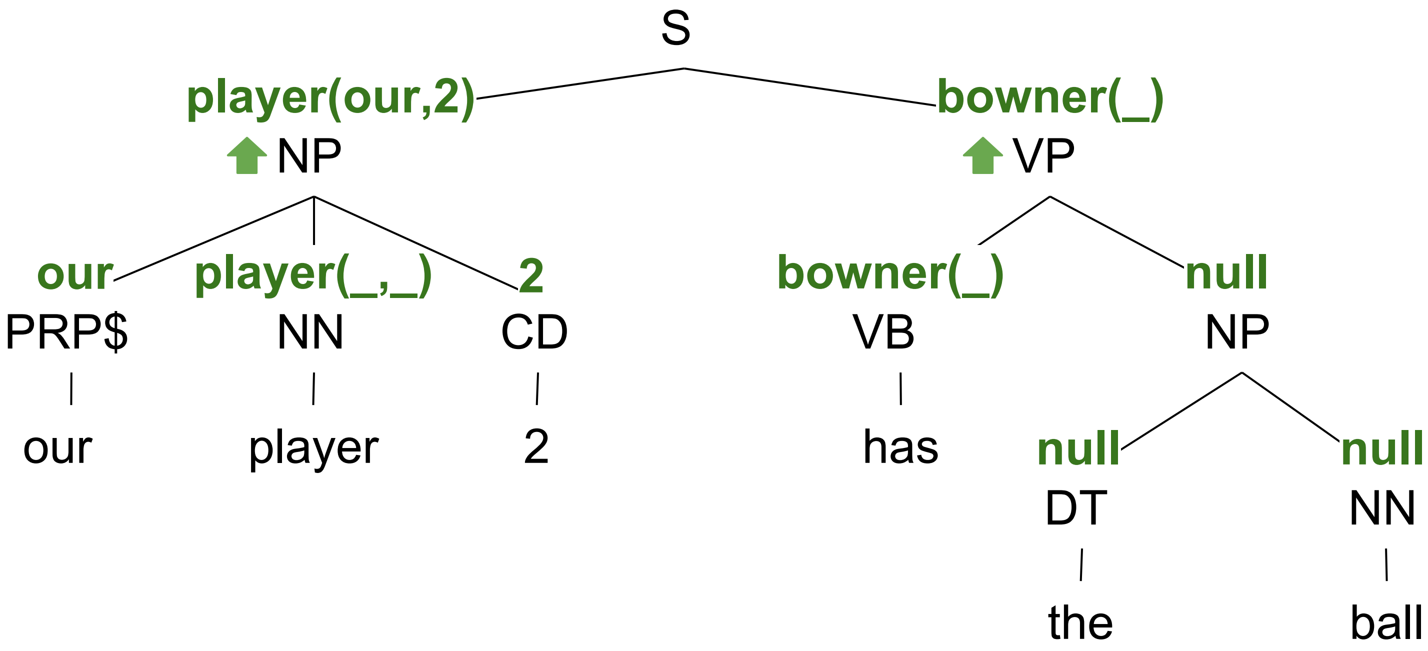
Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **SAPT example:**



Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **SAPT example:**



S

player(our,2) NP

bowner(_) VP

our PRP$ our

player(_,_) NN player

2 CD 2

bowner(_) VB has

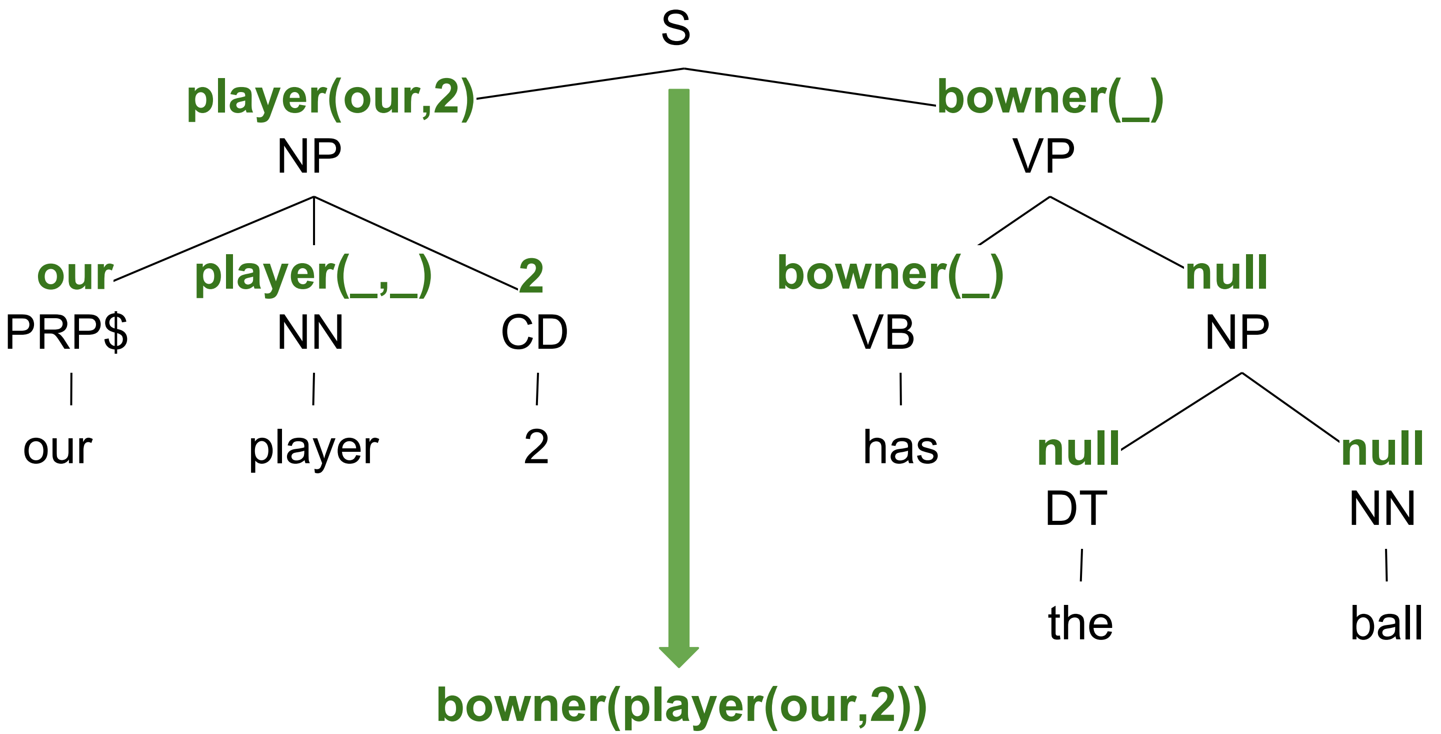
null NP

null DT the

null NN ball

Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **SAPT example:**



S

player(our,2) NP

bowner(_) VP

our PRP$ our

player(_,_) NN player

2 CD 2

bowner(_) VB has

null NP

null DT the

null NN ball

bowner(player(our,2))

Supervised Semantic Parsing: SCISSOR (Ge and Mooney, 2005)

- **Limitations:**
  - Knowledge of syntax (vs. flexibility loss) provides a limited gain for **short sentences**
  - Requires **MR annotation** + extra **SAPT annotation** for training
  - Must learn **both syntax and semantics from the same training corpus** (while high-performance syntactic parsers trained on larger corpora are available)

⇨ **SYNSEM** (Ge and Mooney, 2009): syntactic and semantic parsers trained separately

Supervised Semantic Parsing: Machine Translation Approaches

Intuition:

- MR languages can be statistically modeled exactly as **human languages**!
- In this perspective, MR-annotated corpora becomes **parallel corpora** (EN-MR)

Semantic Parsing as Machine Translation:

Syntax-based statistical machine translation (Chiang, 2005) can be used to learn **semantic grammars** as synchronous context-free grammars (Aho and Ullman, 1972)

Supervised Semantic Parsing: WASP (Wong and Mooney, 2006; 2007)

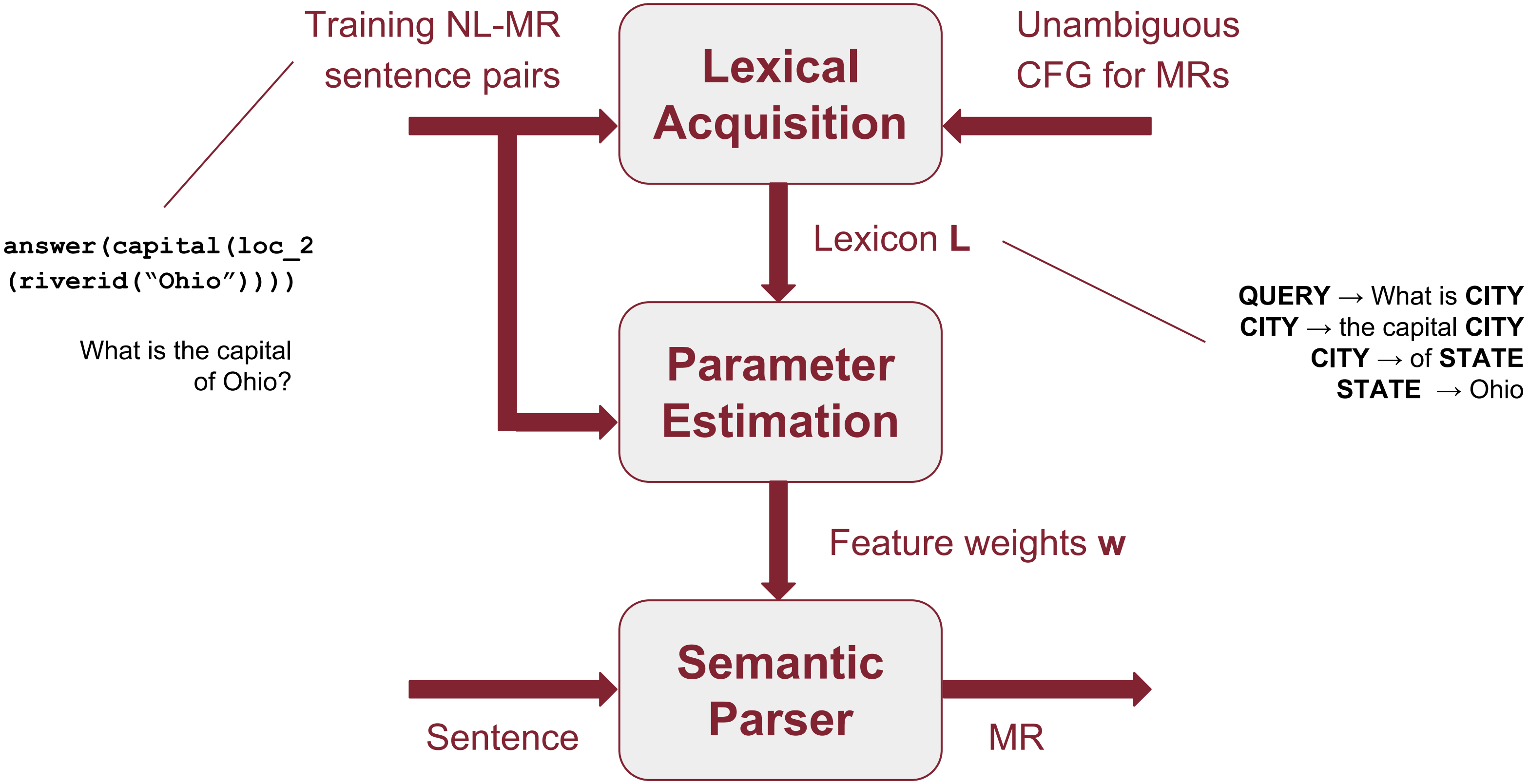
- **WASP (Word Alignment-based Semantic Parsing):**
  - A word alignment model is used to acquire a **bilingual lexicon** consisting of NL substrings coupled with their translations in the target MR language (CLang, GeoQuery, etc.)
  - Complete MRs are formed by combining NL-MRL substring pairs using **SCFG parsing** as in syntax-based translation models (Yamada and Knight, 2001; Chiang, 2005)

```
( (bowner our {4} )
( do our {6} (pos (left (half
our)))) )
```

If our player 4 has the ball, then our player 6 should stay in the left side of our half.

Target **Source**

Supervised Semantic Parsing: WASP (Wong and Mooney, 2006; 2007)

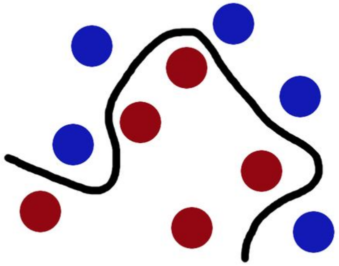


`answer(capital(loc_2`
`(riverid(“Ohio”))))`

What is the capital
of Ohio?

QUERY → What is **CITY**
CITY → the capital **CITY**
CITY → of **STATE**
STATE → Ohio

Supervised Semantic Parsing: Kernel-based Classification



Intuition:

- Statistical feature-based methods struggle to capture the full variety of natural language by only **enumerating** all the possible contexts in which a NL-MR mapping occurs!
- **Kernel methods** implicitly work with a potentially infinite number of features in order to deal with sparseness and noise

⇩

Semantic Parsing as Kernel-based Classification:

For each production of a MR grammar, a classifier based on **string kernels** estimates its probability over different substrings of the input sentence

Supervised Semantic Parsing: KRISP (Kate and Mooney, 2006; Kate, 2008)

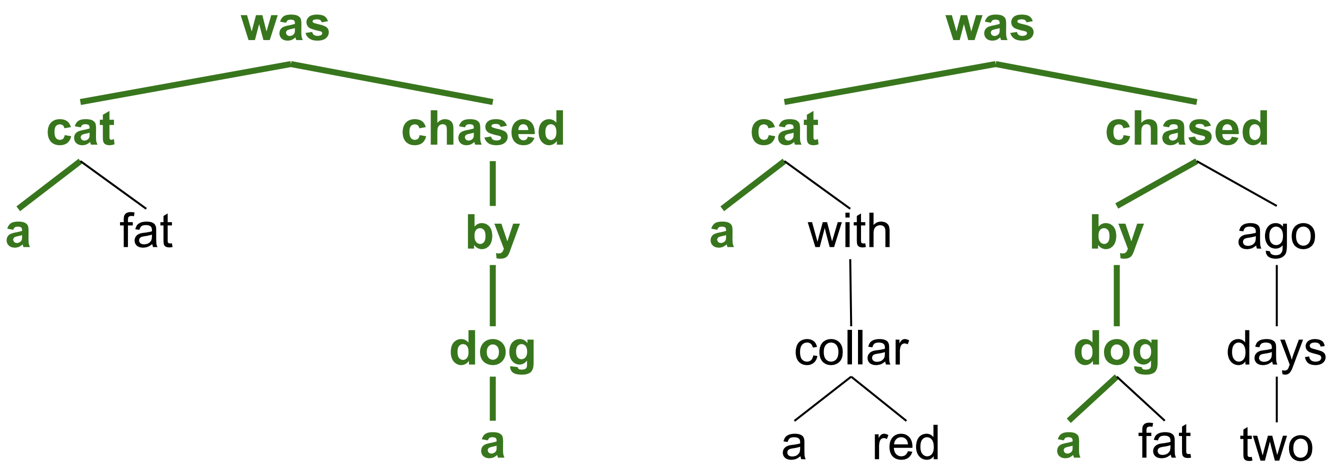
- **KRISP (Kernel-based Robust Interpretation of Semantic Parsing):**
  - Semantic parsing means finding the **most probable derivation** of an input sentence

    ⇨ Dynamic programming algorithm with beam search (Kate & Mooney, 2006)

  - Takes pairs of NL sentences and their respective MRs and induces the semantic parser through an **iterative process** of labeling **positive** and **negative samples** based on a SVM classifier with string-subsequence kernel (Lodhi et al., 2002)

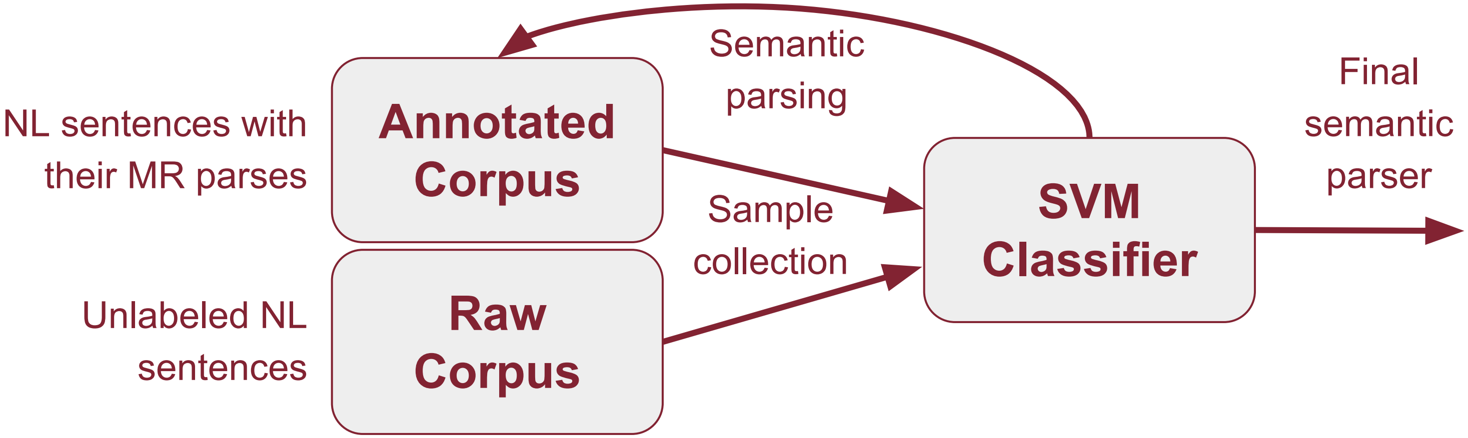
Supervised Semantic Parsing: KRISP (Kate and Mooney, 2006; Kate, 2008)

- **Further improvements:**
  - **Dependency-based word subsequence kernel** (Kate, 2008) to count the number of common paths in the dependency tree



Supervised Semantic Parsing: KRISP (Kate and Mooney, 2006; Kate, 2008)

- **Further improvements:**
  - **SEMISUP-KRISP** (Kate and Mooney, 2007) adopts a semi-supervised learning approach where **unlabeled examples** are considered in the iterative training algorithm



NL sentences with their MR parses

Annotated Corpus

Unlabeled NL sentences

Raw Corpus

Semantic parsing

Sample collection

SVM Classifier

Final semantic parser

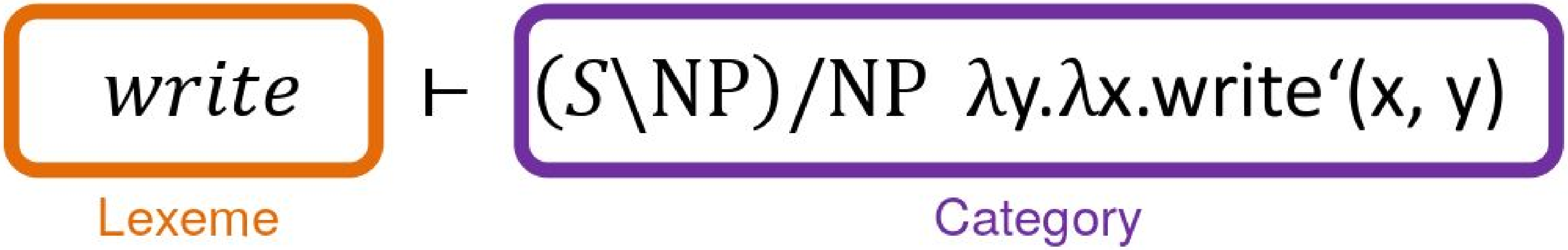
Outline

- Frame-semantic Parsing
- Supervised Semantic Parsing
- Semantic Parsing with CCG
- Unsupervised Semantic Parsing
- Semi-Supervised Semantic Parsing
- Learning from Q&A pairs
- Semantic Parsing with AMR (latest trend!)

Semantic Parsing using CCG

Wait! What is CCG?

- **Combinatory Categorial Grammar (CCG)** is an alternative approach to syntax compared to CFG
    - Transparent interface between **syntax** and **semantics**;
    - Instead of rules and constituents, we have **categories** associated with each element in the lexicon:

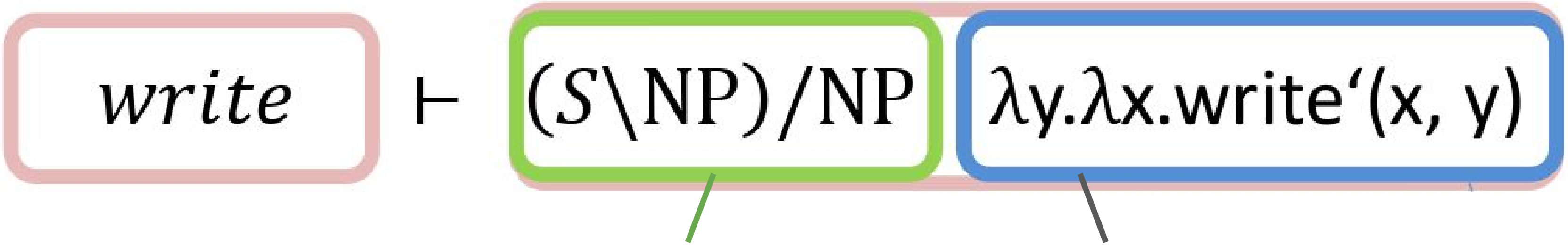


$$\textit{write} \vdash (S\backslash NP)/NP \ \ \lambda y.\lambda x.\mathrm{write}'(x, y)$$

Semantic Parsing using CCG

Wait! What is CCG?

- **Combinatory Categorial Grammar (CCG)** is an alternative approach to syntax compared to CFG
    - Transparent interface between **syntax** and **semantics**;
    - Instead of rules and constituents, we have **categories** associated with each element in the lexicon:



$$\textit{write} \vdash (S\backslash NP)/NP \quad \lambda y.\lambda x.\text{write}'(x, y)$$

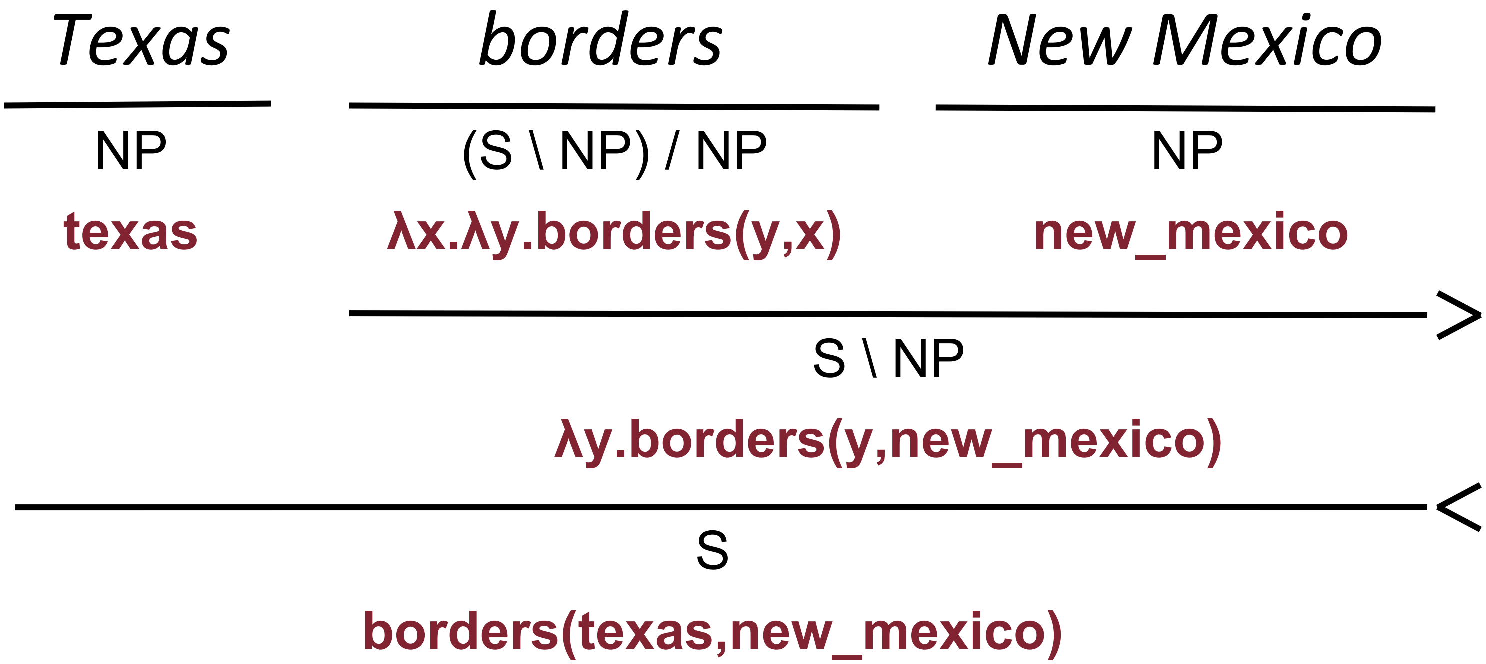
Syntax (Lambek notation):
reads as a function that takes as input a NP on the **left** ("\") and a NP on the **right** ("/") and output a sentence S

Semantics:
λ-calculus in this case, but can be any other formal language!

Semantic Parsing using CCG

Why CCG?

- Complex categories, but very few combination operations that are naturally based on **function composition**:



Texas

NP

texas

borders

(S \ NP) / NP

λx.λy.borders(y,x)

New Mexico

NP

new_mexico

>

S \ NP

λy.borders(y,new_mexico)

<

S

borders(texas,new_mexico)

Semantic Parsing using CCG

CCG is much more than this!

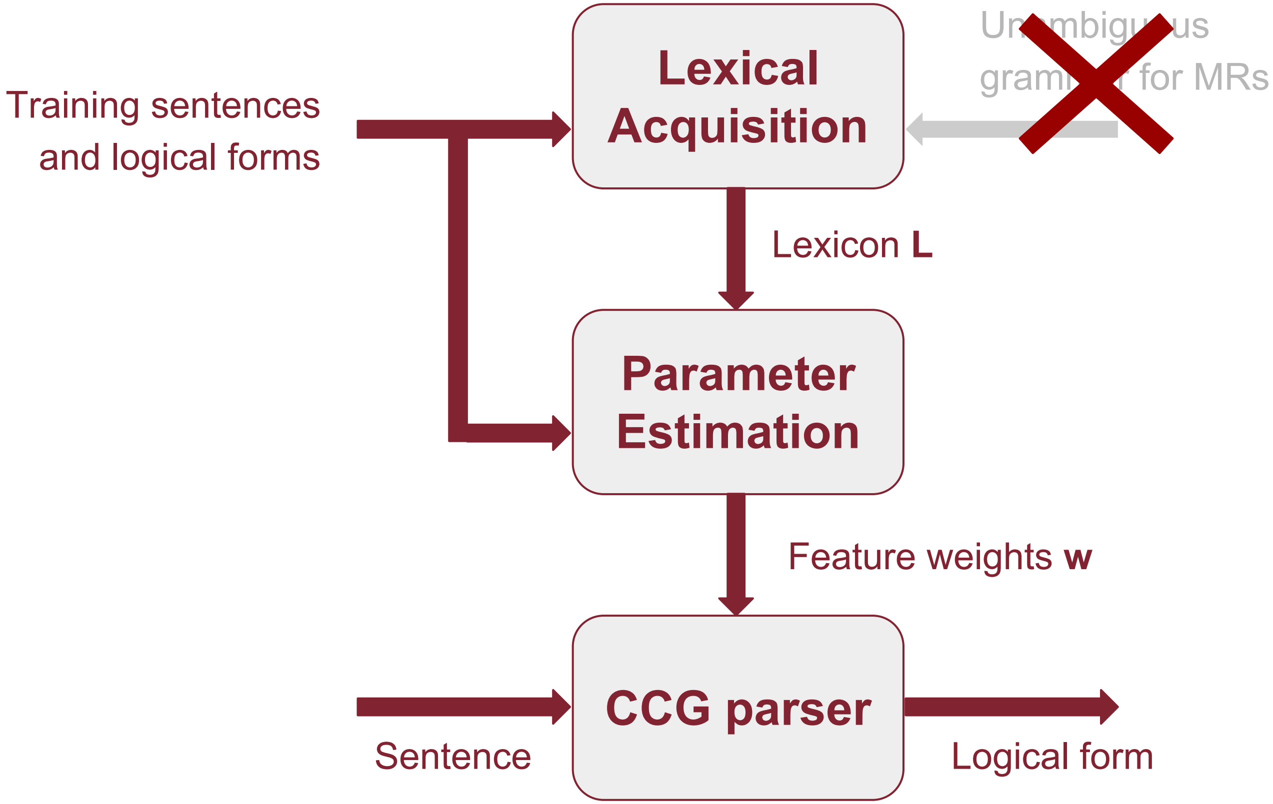
Have a look at my tutorial!

“CCG: a (gentle) introduction”

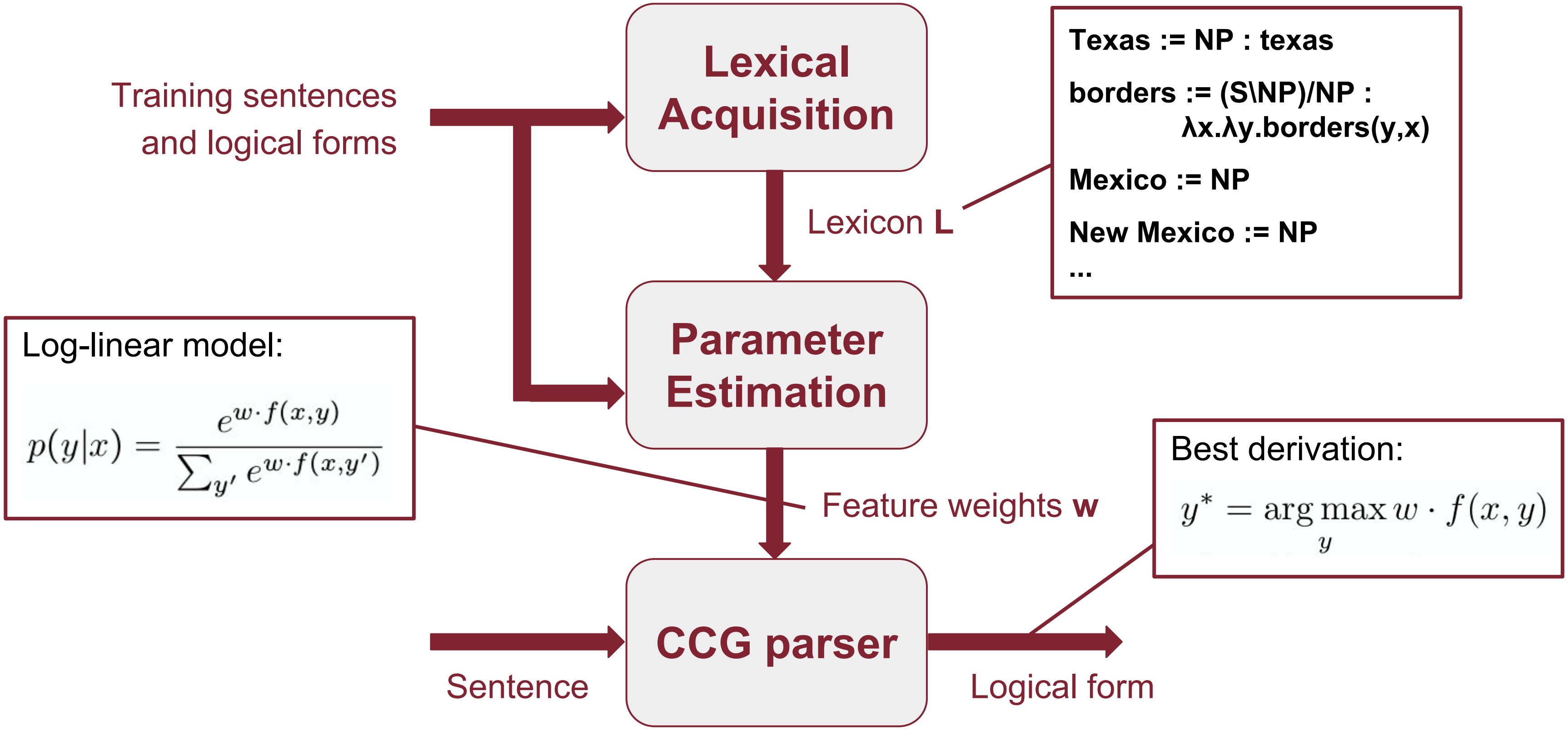
- Generalized **type-raising** operations
- **Cross composition** operations for cross serial dependencies
- Various associated **semantic theories**
- Part of a larger family (**Categorial Grammar**) which is in turn part of a class of “**mildly context-sensitive**” grammars

[Steedman 1996; 2000; 2011; Granroth and Steedman 2012]

Semantic Parsing using CCG (Zettlemoyer and Collins, 2005: 2007)



Semantic Parsing using CCG (Zettlemoyer and Collins, 2005: 2007)



Training sentences and logical forms

Lexical Acquisition

Lexicon **L**

Texas := NP : texas

borders := (S\NP)/NP : λx.λy.borders(y,x)

Mexico := NP

New Mexico := NP

...

Parameter Estimation

Log-linear model:

$$p(y|x) = \frac{e^{w \cdot f(x,y)}}{\sum_{y'} e^{w \cdot f(x,y')}}$$

Feature weights **w**

Best derivation:

$$y^* = \arg\max_y w \cdot f(x,y)$$

CCG parser

Sentence

Logical form

Supervised Semantic Parsing: Wrapping up!

Supervised Semantic Parsing: Wrapping up!

Commonalities:

- A model to connect **language** and **meaning**
- A mechanism for meaning **composition**
- **Parameters** to weight a given meaning representations
- An **iterative learning** algorithm
- A **generalization** mechanism

SCISSOR: Semantically Annotated Parse Trees (SAPT)

WASP: Synchronous CFG

KRISP: Probabilistic string classifiers

Zettlemoyer & Collins: CCG with semantic types

Supervised Semantic Parsing: Wrapping up!

Commonalities:

- A model to connect **language** and **meaning**
- A mechanism for meaning **composition**
- **Parameters** to weight a given meaning representations
- An **iterative learning** algorithm
- A **generalization** mechanism

SCISSOR: Semantically Annotated Parse Trees (SAPT)

WASP: MR grammar

KRISP: MR grammar

Zettlemoyer & Collins: CCG parsing rules

Supervised Semantic Parsing: Wrapping up!

Commonalities:

- A model to connect **language** and **meaning**
- A mechanism for meaning **composition**
- **Parameters** to weight a given meaning representations
- An **iterative learning** algorithm
- A **generalization** mechanism

SCISSOR: Parsing model weights

WASP: Grammar production weights

KRISP: SVM weights

Zettlemoyer & Collins: Weights for lexical items and parsing rules

Supervised Semantic Parsing: Wrapping up!

SCISSOR: No

WASP: Yes

KRISP: No

Zettlemoyer & Collins: Yes

Differences:

- Learn **lexicon** or not
- Exploit general **syntactic parsing** or not
- Use **matching patterns** or not

Supervised Semantic Parsing: Wrapping up!

SCISSOR: CFG

WASP: No

KRISP: No

Zettlemoyer & Collins: CCG

Differences:

- Learn **lexicon** or not
- Exploit general **syntactic parsing** or not
- Use **matching patterns** or not

✓ **Pro:** Leverage knowledge of natural language

✗ **Con:** Not immediately portable to other languages

Supervised Semantic Parsing: Wrapping up!

Differences:

- Learn **lexicon** or not
- Exploit general **syntactic parsing** or not
- Use **matching patterns** or not

SCISSOR: Yes

WASP: Yes

KRISP: No

Zettlemoyer & Collins: Yes

✓ **Pro:** The parser can be “inverted” to form a generation system

✗ **Con:** Affected by noise and data sparseness

Outline

- Frame-semantic Parsing
- Supervised Semantic Parsing
- Semantic Parsing with CCG
- Unsupervised Semantic Parsing
- **Semi-Supervised Semantic Parsing**
- **Learning from Q&A pairs**
- **Semantic Parsing with AMR** (latest trend!)

Unsupervised Systems

- Unsupervised Semantic Parsing - **Poon & Domingos** (2009)
    - Based on **Markov Logic** (Richardson & Domingos 2006)
    - Can be used in general domains

Unsupervised Semantic Parsing - USP

- **Key idea #1**:
    - **Clusters** of **syntactic** or **lexical variations** with **same meaning**
      Buy = {buy, acquire, purchase, ...}
      MICROSOFT = {Microsoft, Bill Gates' company, …}
    - Target **predicates** and **objects** can be **learned**

Unsupervised Semantic Parsing - USP

- **Key idea #2**:
  - **Relational clustering:** cluster relations with similar subexpressions

Example:

- Microsoft buys Powerset
- Microsoft acquires semantic search engine Powerset
- Powerset is acquired by Microsoft Corporation
- The Redmond software giant buys Powerset
- Microsoft’s purchase of Powerset, …

Unsupervised Semantic Parsing - USP

- **Key idea #2**:
    - **Relational clustering:** cluster relations with similar subexpressions

Example:

- Microsoft buys Powerset
- Microsoft acquires semantic search engine Powerset
- Powerset is acquired by Microsoft Corporation
- The Redmond software giant buys Powerset
- Microsoft's purchase of Powerset, …

Unsupervised Semantic Parsing - USP

- **Key idea #2**:
    - **Relational clustering:** cluster relations with similar subexpressions

Example:

- Microsoft buys Powerset
- Microsoft acquires semantic search engine Powerset
- Powerset is acquired by Microsoft Corporation
- The Redmond software giant buys Powerset
- Microsoft's purchase of Powerset, …

Unsupervised Semantic Parsing - USP

- **Key idea #2**:
  - **Relational clustering:** cluster relations with similar subexpressions

Example:

- Microsoft buys Powerset
- Microsoft acquires semantic search engine Powerset
- Powerset is acquired by Microsoft Corporation
- The Redmond software giant buys Powerset
- Microsoft's purchase of Powerset, …

Unsupervised Semantic Parsing - USP

- **Key idea #2**:
  - **Relational clustering:** cluster relations with similar subexpressions

Example:

- Microsoft buys Powerset
- Microsoft acquires semantic search engine Powerset
- Powerset is acquired by Microsoft Corporation
- The Redmond software giant buys Powerset
- Microsoft’s purchase of Powerset, …

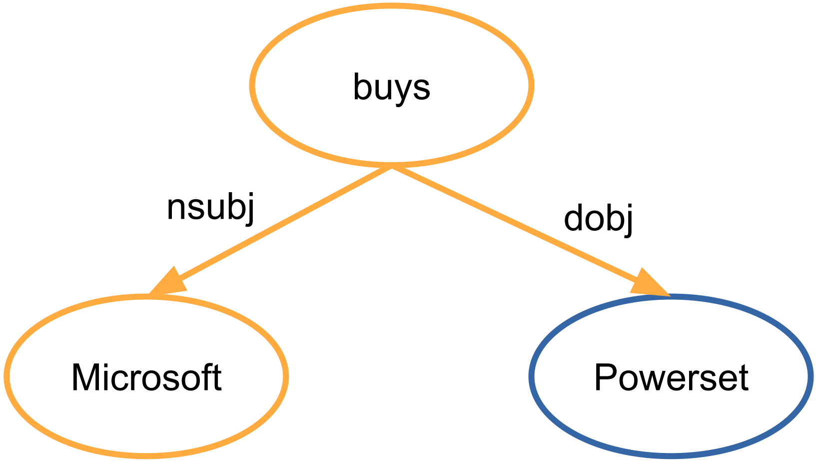
Unsupervised Semantic Parsing - USP

- **Key idea #3**:
  - Starts directly from **syntactic analyses**
  - Focus on translating **syntax trees** in **semantic trees**
  - Leverage rapid **progress** in **syntactic parsing**

Unsupervised Semantic Parsing - USP

Practical example

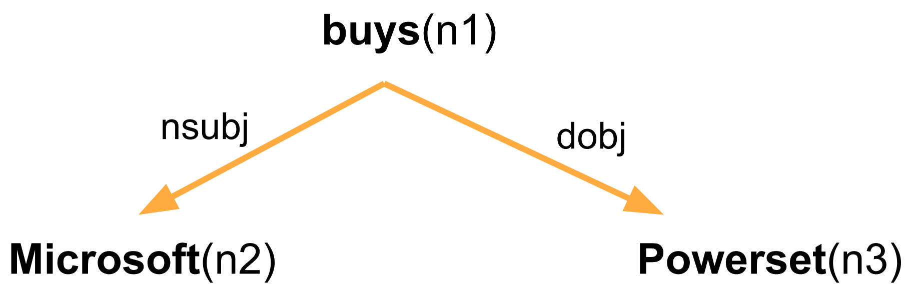
- Syntax tree



Unsupervised Semantic Parsing - USP

Practical example

- Vertices conversion into unary atoms



buys(n1)

nsubj

dobj

Microsoft(n2)

Powerset(n3)

n1, n2 and n3 are Skolem constants

Unsupervised Semantic Parsing - USP

Practical example

- Edge conversion into binary atoms

buys(n1)

nsubj(n1, n2)

dobj(n1, n3)

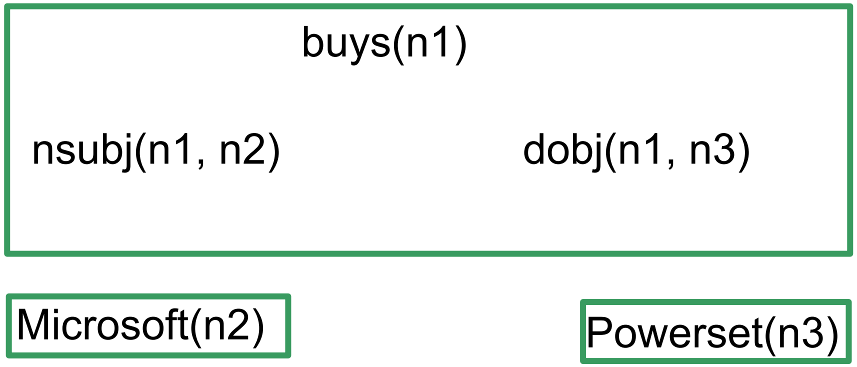
Microsoft(n2)

Powerset(n3)

Unsupervised Semantic Parsing - USP

Practical example

- Partitioning of Quasi-Logical Forms into sub-formulas



buys(n1)

nsubj(n1, n2)

dobj(n1, n3)

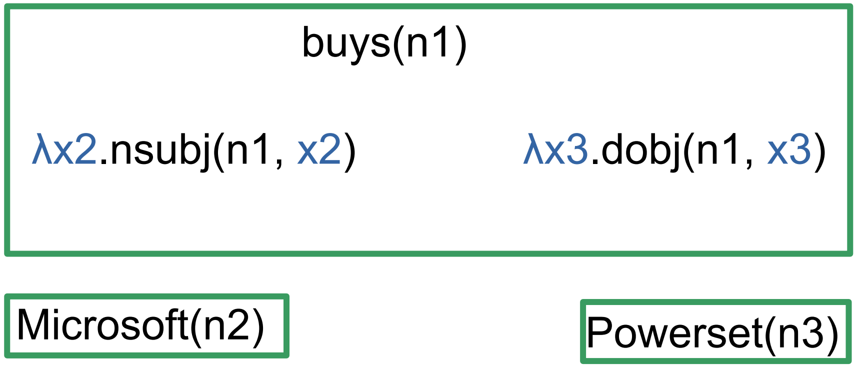
Microsoft(n2)

Powerset(n3)

Unsupervised Semantic Parsing - USP

Practical example

- Partitioning of Quasi-Logical Forms into sub-formulas



buys(n1)

λx2.nsubj(n1, x2)

λx3.dobj(n1, x3)

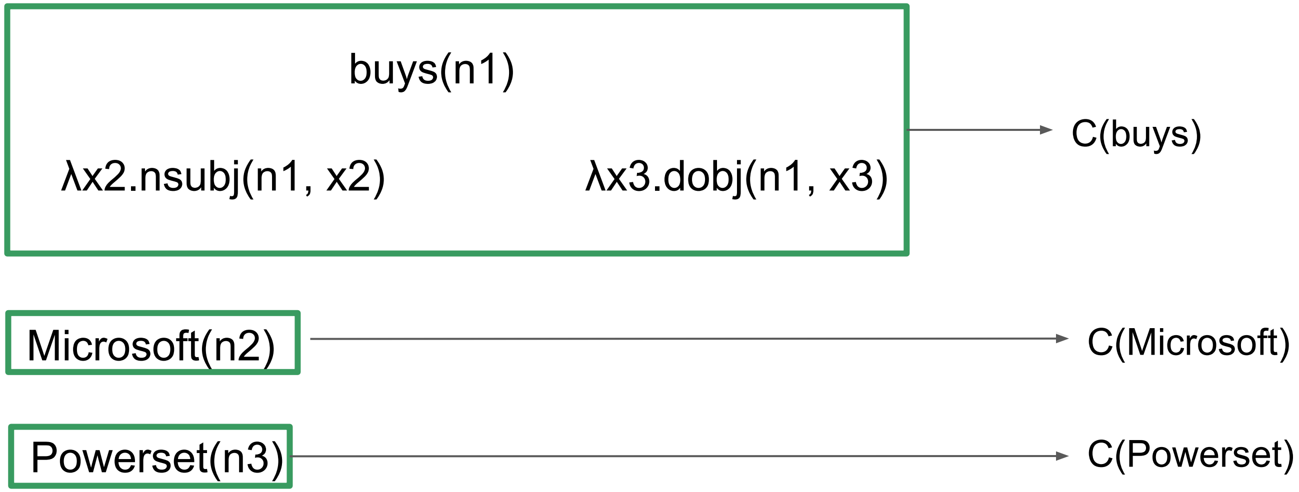
Microsoft(n2)

Powerset(n3)

Unsupervised Semantic Parsing - USP

Practical example

- Assign subformula to lambda-form clusters



buys(n1)

λx2.nsubj(n1, x2)

λx3.dobj(n1, x3)

C(buys)

Microsoft(n2)

C(Microsoft)

Powerset(n3)

C(Powerset)

Unsupervised Semantic Parsing - USP

Practical example

- Abstract lambda formula

buys(n1) ^ λx2.nsubj(n1, x2) ^ λx3.dobj(n1, x3)

↓

C(buys)(n1) ^ λx2.A(buyer)(n1, x2) ^ λx3.A(bought)(n1, x3)

Outline

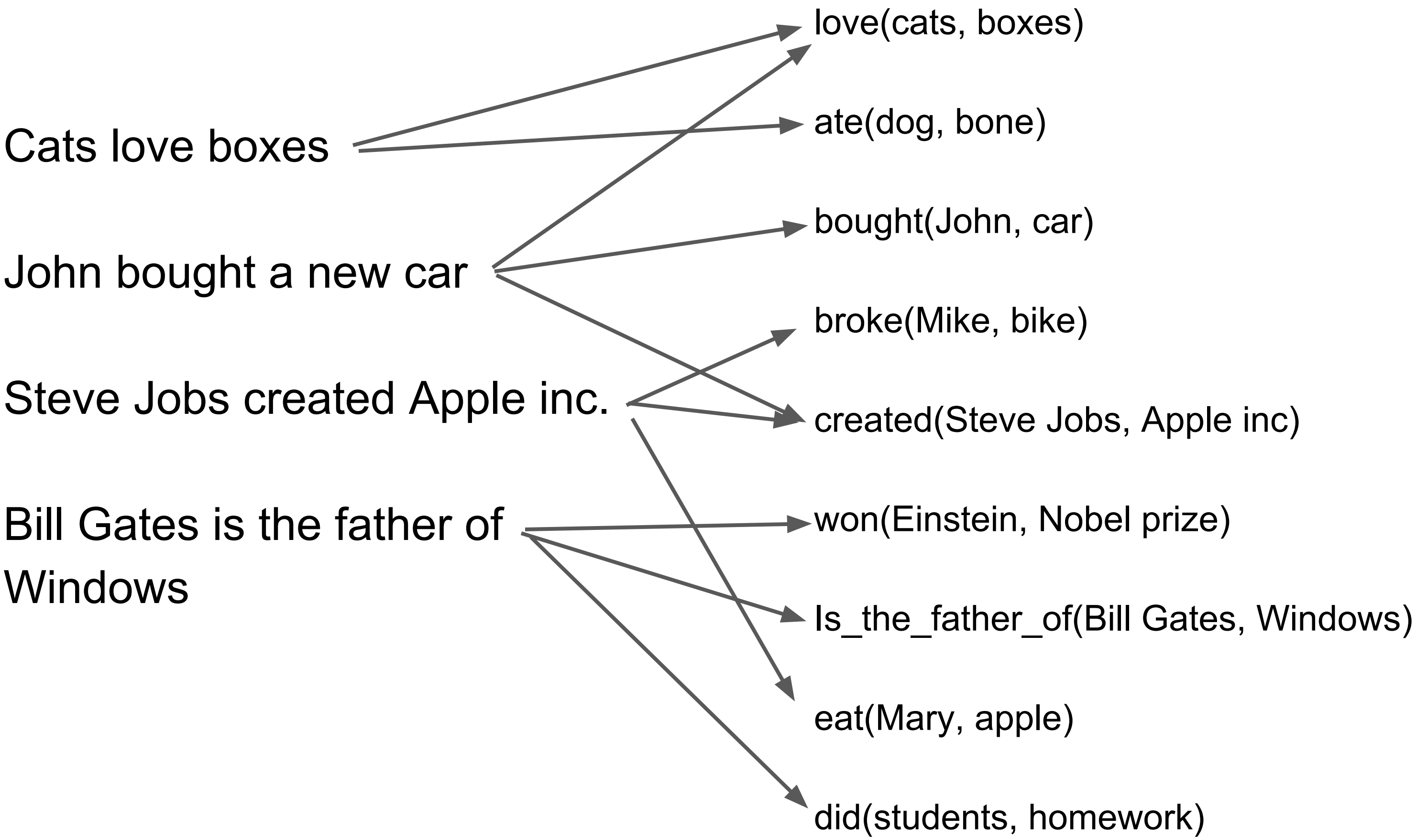
- Frame-semantic Parsing
- Supervised Semantic Parsing
- Semantic Parsing with CCG
- Unsupervised Semantic Parsing
- Semi-Supervised Semantic Parsing
- **Learning from Q&A pairs**
- **Semantic Parsing with AMR** (latest trend!)

Semi-Supervised Semantic Parsing

- **SEMISUP-KRISP (unambiguous supervision)**

- **KRISPER (ambiguous supervision)**
  (Learning Language Semantics from Ambiguous Supervision, Kate, Mooney 2007)

Semi-Supervised Semantic Parsing: Ambiguous Supervision



Cats love boxes

John bought a new car

Steve Jobs created Apple inc.

Bill Gates is the father of Windows

love(cats, boxes)

ate(dog, bone)

bought(John, car)

broke(Mike, bike)

created(Steve Jobs, Apple inc)

won(Einstein, Nobel prize)

Is_the_father_of(Bill Gates, Windows)

eat(Mary, apple)

did(students, homework)

Semi-Supervised Semantic Parsing:KRISPER

- Extends KRISP to handle **ambiguous training set**:
    - **Assigns weights** to each pair (NL-sentence, MR) equals to 1 over the number of MR for NL-sentence in the dataset.
    - During the SVM iterations the “**penalization score**” **is increased** in order to **allow incorrect classifications at the beginning**.
    - Once **all NL-sentences** have been paired **at most** with **1 MR**, then the original **KRISP algorithm is called** in order to learn a better semantic parser.

Outline

- Frame-semantic Parsing
- Supervised Semantic Parsing
- Semantic Parsing with CCG
- Unsupervised Semantic Parsing
- Semi-Supervised Semantic Parsing
- Learning from Q&A pairs
- **Semantic Parsing with AMR** (latest trend!)

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)

- Aviailable at http://nlp.stanford.edu/software/sempre/
- Exploits available **dataset of question-answer pairs** to automatically train a semantic parser.
- Does not need a dataset with Natural Language sentences and their associated Meaning Representations.
- Exploits **Freebase** (a huge knowledge base) in order to find the right Meaning Representations for a query, which, if used on the knowledge base, comes out with the right answer.

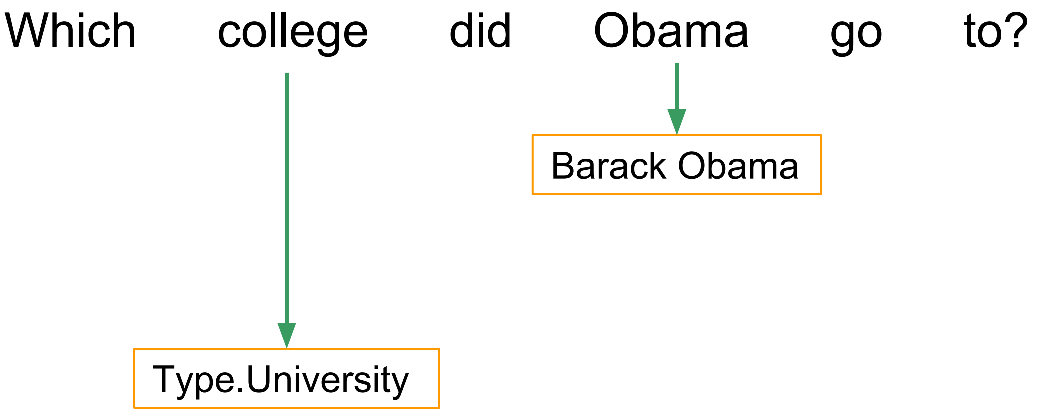
Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)

Which college did Obama go to?

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)



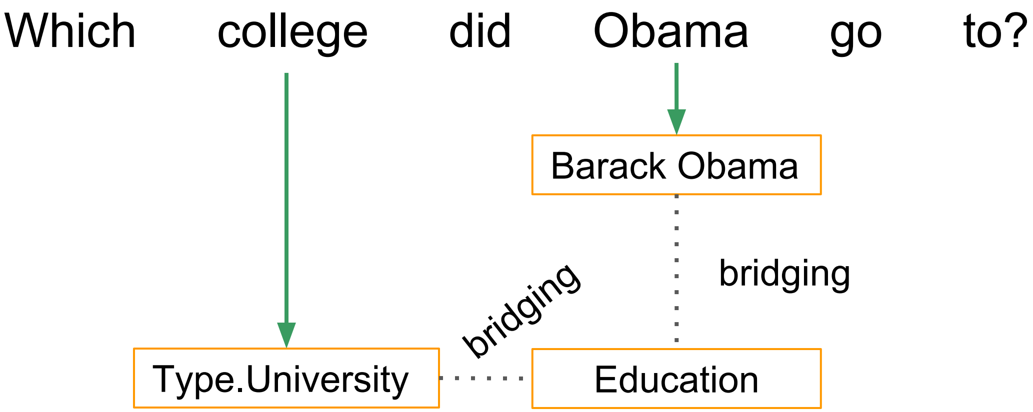
Which college did Obama go to?

Barack Obama

Type.University

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)



Which college did Obama go to?

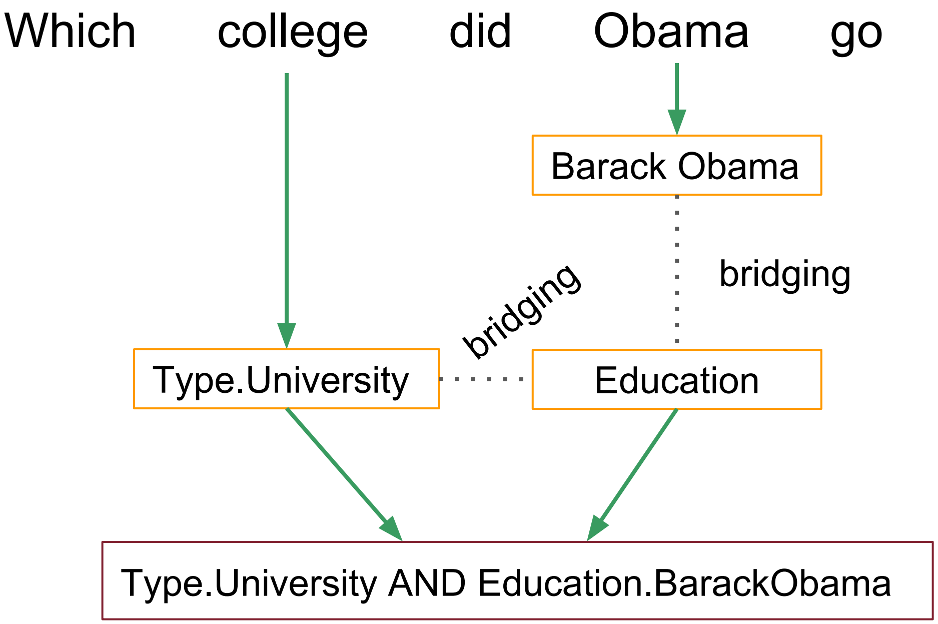
Barack Obama

Type.University

Education

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)



Which college did Obama go to?

Barack Obama

bridging

bridging

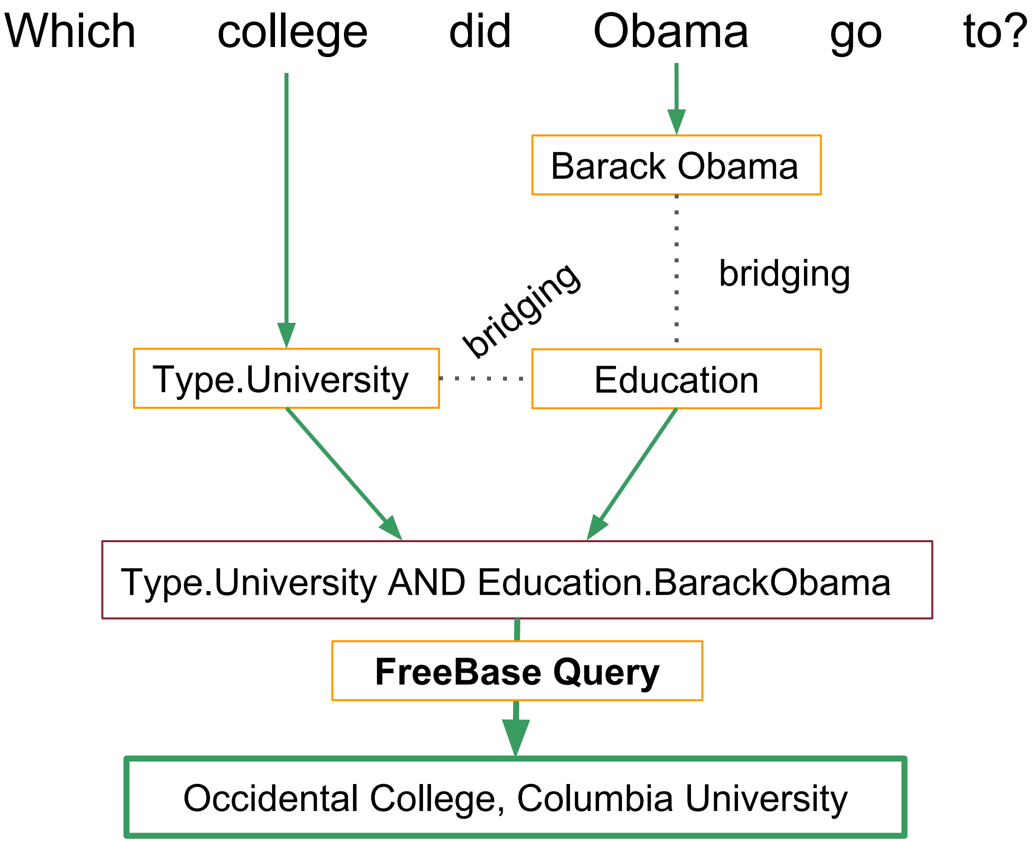
Type.University

Education

Type.University AND Education.BarackObama

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)



Which college did Obama go to?

Barack Obama

bridging

bridging

Type.University

Education

Type.University AND Education.BarackObama

FreeBase Query

Occidental College, Columbia University

Learning from Q&A pairs:

Semantic Parsing on Freebase from Question-Answer Pairs (Berant et al. 2013)

- **Bridging** aims at **establish novel relations** between distinct **parts of a sentence**.

- **Generates** a **binary predicate** based on **neighboring logical predicates** rather than on explicit lexical material.

- Given a pair of **unaries u1, u2** (e.g. BarackObama and Type. University) with **type t1,t2** (Person, University), given the binary operator **b(t1,t2)** (Education(Person, University)), then the formula **u1 AND b.u2** will be produced (e.g. Type.University AND Education. BarakObama).

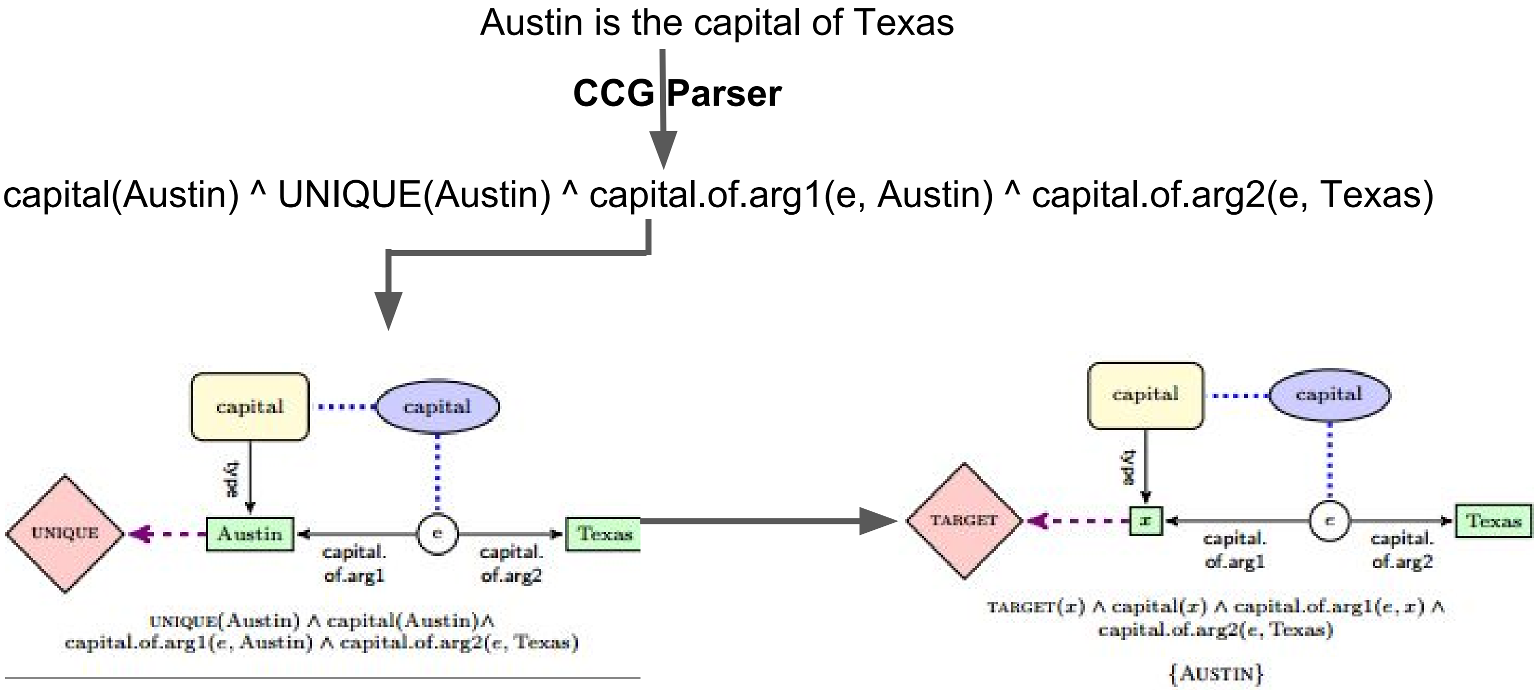
Learning ~~from~~ without Q&A pairs:

Large-scale Semantic Parsing without Question-Answer Pairs (Reddy, Lapata, Steedman 2014)

- Exploits existing **CCG syntactic parser** in order to build a **semantic parse**
- **Builds a graph for the semantic parse**
- **Grounds** the previously extracted **graph on FreeBase**
- **Matches the grounded graph on FreeBase** in order to retrieve the correct answer for the question expressed by the graph.

Learning ~~from~~ without Q&A pairs:

Large-scale Semantic Parsing without Question-Answer Pairs (Reddy, Lapata, Steedman 2014)



Austin is the capital of Texas

CCG Parser

capital(Austin) ^ UNIQUE(Austin) ^ capital.of.arg1(e, Austin) ^ capital.of.arg2(e, Texas)

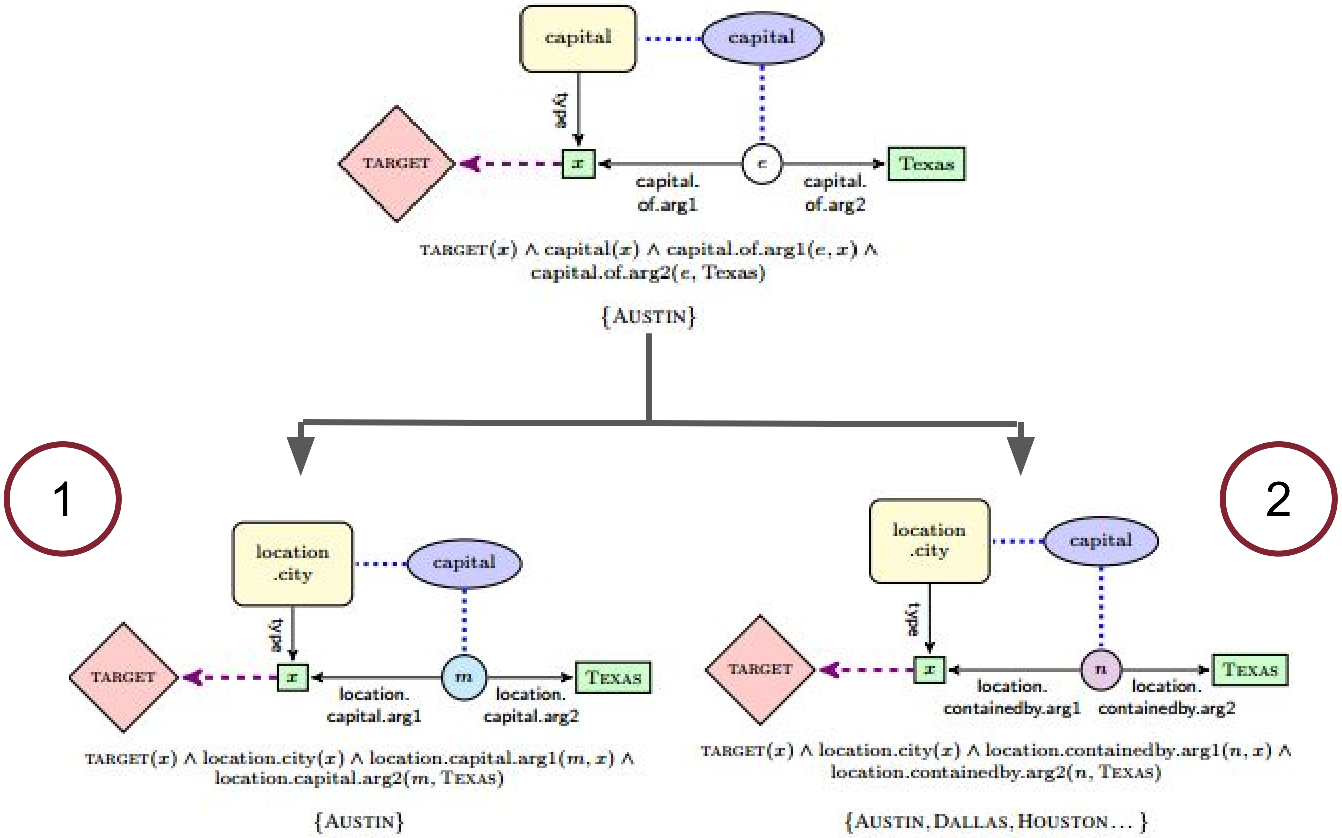
$\text{UNIQUE}(\text{Austin}) \wedge \text{capital}(\text{Austin}) \wedge$ $\text{capital.of.arg1}(e, \text{Austin}) \wedge \text{capital.of.arg2}(e, \text{Texas})$

$\text{TARGET}(x) \wedge \text{capital}(x) \wedge \text{capital.of.arg1}(e, x) \wedge$ $\text{capital.of.arg2}(e, \text{Texas})$

$\{\text{AUSTIN}\}$

Learning ~~from~~ without Q&A pairs:

Large-scale Semantic Parsing without Question-Answer Pairs (Reddy, Lapata, Steedman 2014)



$\text{TARGET}(x) \wedge \text{capital}(x) \wedge \text{capital.of.arg1}(e, x) \wedge \text{capital.of.arg2}(e, \text{Texas})$

{AUSTIN}

$\text{TARGET}(x) \wedge \text{location.city}(x) \wedge \text{location.capital.arg1}(m, x) \wedge \text{location.capital.arg2}(m, \text{TEXAS})$

{AUSTIN}

$\text{TARGET}(x) \wedge \text{location.city}(x) \wedge \text{location.containedby.arg1}(n, x) \wedge \text{location.containedby.arg2}(n, \text{TEXAS})$

{AUSTIN, DALLAS, HOUSTON … }

Learning ~~from~~ without Q&A pairs:

Large-scale Semantic Parsing without Question-Answer Pairs (Reddy, Lapata, Steedman 2014)

- Uses **features** (F) in order to find the **best grounded graph**
- Uses **Perceptron** in order to **learn weights** (W) for each features
- Chooses the **grounded graph** that **maximize** the **dot-product between F and W.**

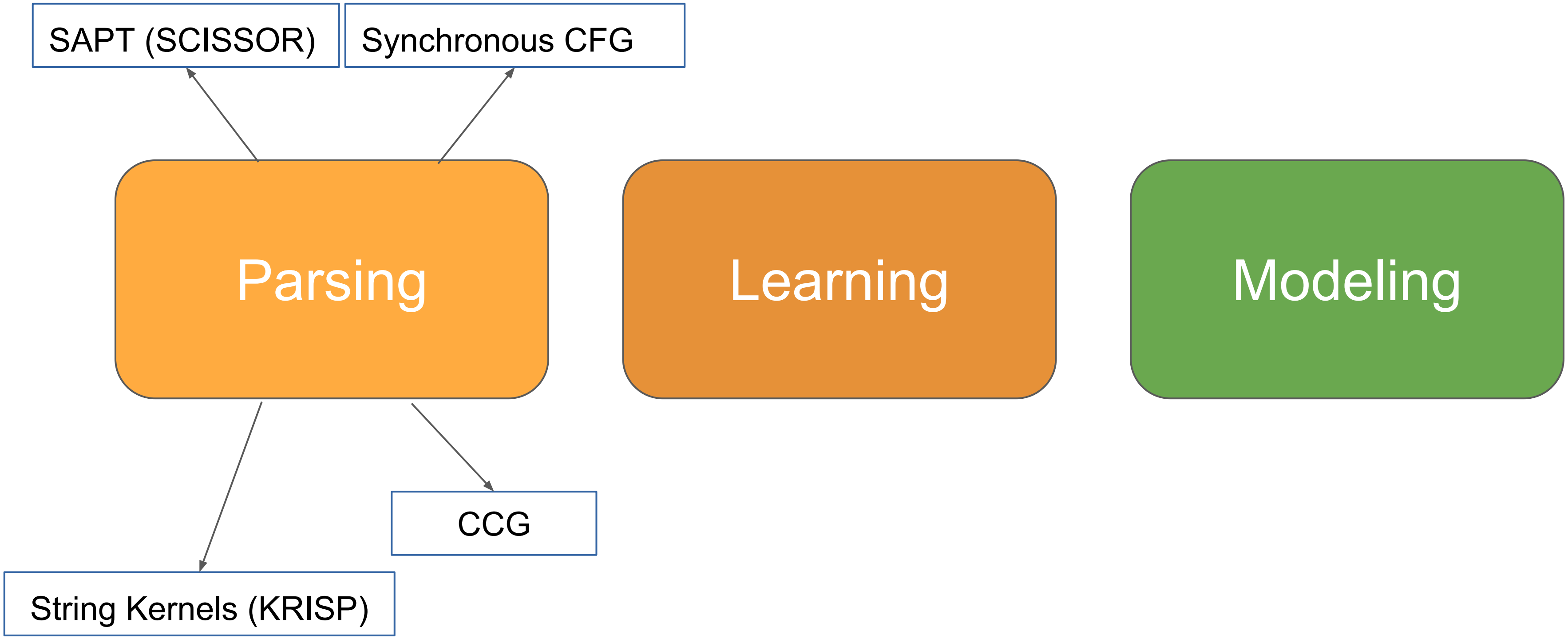
Recap

Parsing

Learning

Modeling

Recap



SAPT (SCISSOR)

Synchronous CFG

Parsing

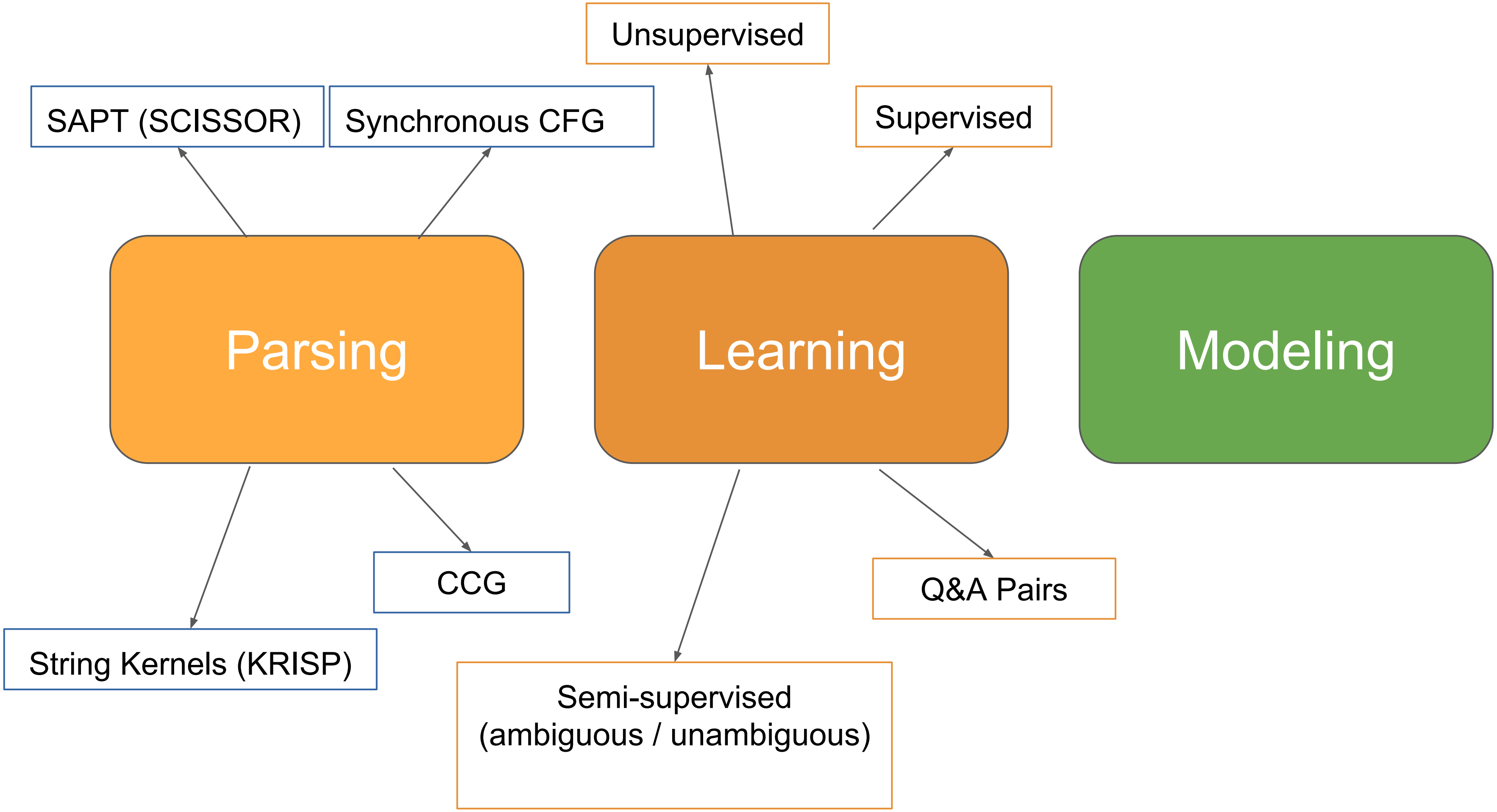
Learning

Modeling

CCG

String Kernels (KRISP)

Recap



Unsupervised

SAPT (SCISSOR)

Synchronous CFG

Supervised

Parsing

Learning

Modeling

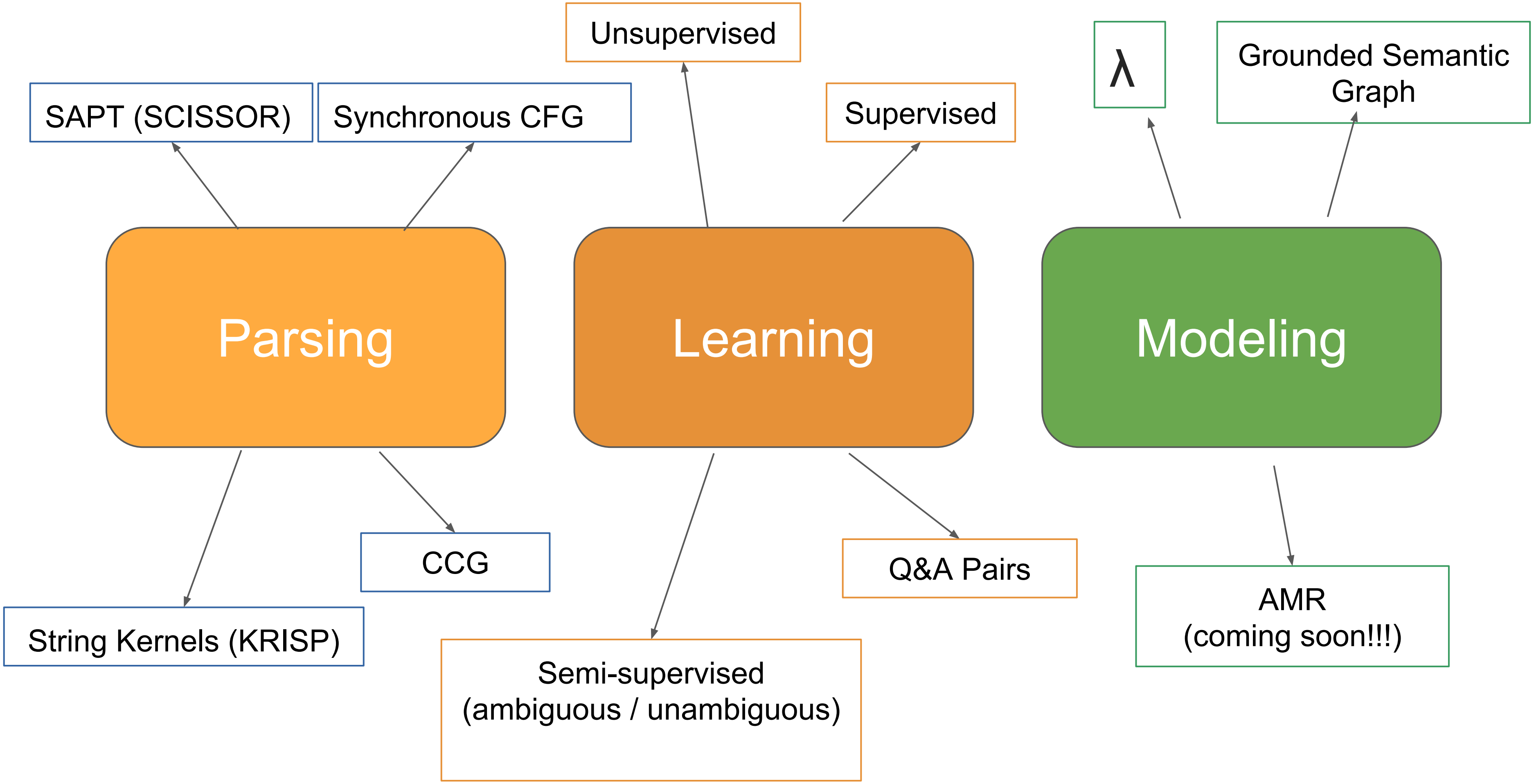
CCG

Q&A Pairs

String Kernels (KRISP)

Semi-supervised
(ambiguous / unambiguous)

Recap



Unsupervised

SAPT (SCISSOR)

Synchronous CFG

Supervised

λ

Grounded Semantic Graph

Parsing

Learning

Modeling

CCG

Q&A Pairs

AMR
(coming soon!!!)

String Kernels (KRISP)

Semi-supervised
(ambiguous / unambiguous)

References

- ## Frame-semantic Parsing

Das, D., Chen, D., Martins, A.F., Schneider, N. and Smith, N.A., 2014. **Frame-semantic Parsing.** *Computational Linguistics*, *40*(1), pp.9-56.

Hermann, K.M., Das, D., Weston, J. and Ganchev, K., 2014. **Semantic Frame Identification with Distributed Word Representations**. *Proceedings of ACL 2014*, pp. 1448-1458.

Täckström, O., Ganchev, K. and Das, D., 2015. **Efficient Inference and Structured Learning for Semantic Role Labeling**. *Transactions of the Association for Computational Linguistics*, vol 3, pp.29-41.

- ## Supervised Semantic Parsing

Ge, R. and Mooney, R.J., 2005. **A Statistical Semantic Parser that Integrates Syntax and Semantics**. *Proceedings of CoNLL 2005*, pp. 9-16.

Wong, Y.W. and Mooney, R.J., 2006. **Learning for Semantic Parsing with Statistical Machine Translation**. In *Proceedings of HLT-NAACL 2006*, pp. 439-446.

References

- ## Supervised Semantic Parsing (cont.)

Kate, R.J. and Mooney, R.J., 2006. **Using String-Kernels for Learning Semantic Parsers**. *Proceedings of ACL 2006,* pp. 913-920.

Kate, R.J. and Mooney, R.J., 2007. **Semi-Supervised Learning for Semantic Parsing using Support Vector Machines**. *Proceedings of HLT-NAACL 2007*, pp. 81-84.

Zettlemoyer L.S., Collins M., 2007. **Learning to Map Sentences to Logical Form: Structured Classification with Probabilistic Categorial Grammars.** *Proceedings of UAI 2005*, pp. 658-666.

- ## Unsupervised/Semi-Supervised Semantic Parsing

Poon, H. and Domingos, P., 2009. **Unsupervised Semantic Parsing**. *Proceedings of EMNLP 2009*, pp. 1-10.

Kate, R.J. and Mooney, R.J., 2007. **Learning Language Semantics from Ambiguous Supervision**. *Proceedings of AAAI 2007*, vol. 7, pp. 895-900.

References

- ## Semi-Supervised Semantic Parsing (cont.)

Berant, J., Chou, A., Frostig, R. and Liang, P., 2013 **Semantic Parsing on Freebase from Question-Answer Pairs**. *Proceedings of EMNLP 2013*, pp. 1533–1544.

Reddy, S., Lapata, M. and Steedman, M., 2014. **Large-Scale Semantic Parsing without Question-Answer Pairs**. *Transactions of the Association for Computational Linguistics*, *2*, pp. 377-392.

- ## Abstract Meaning Representation

Complete tutorial at: https://github.com/nschneid/amr-tutorial/tree/master/slides

Video of the talk also available at: http://techtalks.tv/talks/the-logic-of-amr-practical-unified-graph-based-sentence-semantics-for-nlp/61564/