Multilingual Semantics

Elias Stengel-Eskin

Advanced NLP: Multilingual Methods

Goals/questions

What is semantics?

What is semantic parsing?

Methods/data for semantic parsing

Multilingual and cross-lingual semantics

1. Logic-based semantic representations
    1. Review of logic
    2. Challenges
    3. Thematic roles
    4. Neo-Davidsonian Event Semantics
2. Abstract Meaning Representation (AMR)
    1. What is it?
    2. Problems it solves
    3. Problems it has
    4. Multilingual resources
3. Universal Dependencies and Semantic Dependencies (detour)
    1. Syntax, abstraction vs. lexicalization
    2. Multilingual resources
4. Universal Decompositional Semantics (UDS)
    1. Goals and overview
    2. Protoroles
5. Executable parsing (detour)
    1. Goals, datasets
6. Question break
7. Semantic parsing
    1. Alignment-based
    2. Seq2seq
    3. Seq2graph
8. Speculative/future work

What is semantics

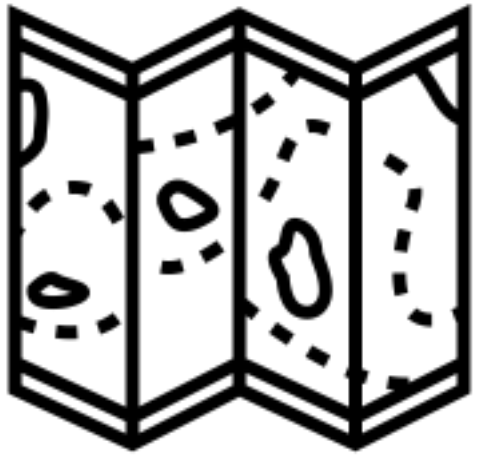
Roughly: relationship between words and meaning

What does a given sentence mean?

How to represent meaning

Abstraction

King and the mapmaker



Abstraction

Level of abstraction depends on goals

Want to capture important regularities in language

Quick logic review

Predicates

For a domain D

$p : D \to \{T, F\}$

$p(x) = T$ iff p is true of x

Connectives

$\wedge$:

$T \wedge T = T$

$T \wedge F = F$

$F \wedge T = F$

$F \wedge F = F$

Quick logic review

Quantifiers

$\exists x =$ There is some x in the domain D ...
$\forall x =$ For every x in the domain D ...

$\exists x.p(x) =$ There exists some x in the domain D for which p is true.

Language and Logic

One proposal: represent the semantics of language with logic

Seems reasonable…

"I walked the dog and fed him" = $\text{walked}(\text{I}, \text{dog}) \wedge \text{fed}(\text{I}, \text{dog})$

Language and Logic

Not so fast!

How do we know who “I” and “dog” are?

Fine if we have 1 dog

But what if we have 2?

$$\text{walked}(\text{I}, \text{dog}) \wedge \text{fed}(\text{I}, \text{dog})$$

$$\text{walked}(\text{I}, \text{dog}) \wedge \text{fed}(\text{I}, \text{dog})$$

Language and Logic

We need an abstraction with variables

$$\exists x, y.\mathrm{me}(x) \wedge \mathrm{dog}(y) \wedge \mathrm{walked}(x, y) \wedge \mathrm{fed}(x, y)$$

$$\exists x, y, z.\mathrm{me}(x) \wedge \mathrm{dog}(y) \wedge \mathrm{dog}(z) \wedge y! = z \wedge \mathrm{walked}(x, y) \wedge \mathrm{fed}(x, z)$$

Language and Logic

One more problem…

"The boy hit the ball" $\exists x, y.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{hit}(x, y)$

"The boy hit the ball with a bat" $\exists x, y, z.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{bat}(z) \wedge \text{hit}(x, y, z)$

"The boy hit the ball with a bat in the park"

$$\exists x, y, z, w.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{bat}(z) \wedge \text{park}(w) \wedge \text{hit}(x, y, z, w)$$

"The boy hit the ball in the park" $\exists x, y, w.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{park}(w) \wedge \text{hit}(x, y, w)$

Language and Logic

Thematic roles

“The boy hit the ball”

[Agent] [Patient]

“The boy hit the ball with a bat”

[Agent] [Patient] [Instrument]

“The boy hit the ball with a bat in the park”

[Agent] [Patient] [Inst.] [Loc.]

“The boy hit the ball in the park”

[Agent] [Patient] [Loc.]

Role	Description
Agent	Person/thing doing an action
Patient	Person/thing action is being done to
Instrument	Thing being used for performing action
Location	Where the action is being performed
…	…

Language and Logic

Neo-Davidsonian Event Semantics

"The boy hit the ball"

$$\exists x, y, e.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{hit}(e) \wedge \text{AGENT}(e, x) \wedge \text{PATIENT}(e, y)$$

"The boy hit the ball with a bat"

$$\exists x, y, z.e.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{bat}(z) \wedge \text{hit}(e) \wedge \text{AGENT}(e, x) \wedge \text{PATIENT}(e, y) \wedge \text{INST.}(e, z)$$

"The boy hit the ball in the park"

$$\exists x, y, z.e.\text{boy}(x) \wedge \text{ball}(y) \wedge \text{park}(z) \wedge \text{hit}(e) \wedge \text{AGENT}(e, x) \wedge \text{PATIENT}(e, y) \wedge \text{LOC.}(e, z)$$

Other representations

Logic still has many shortcomings

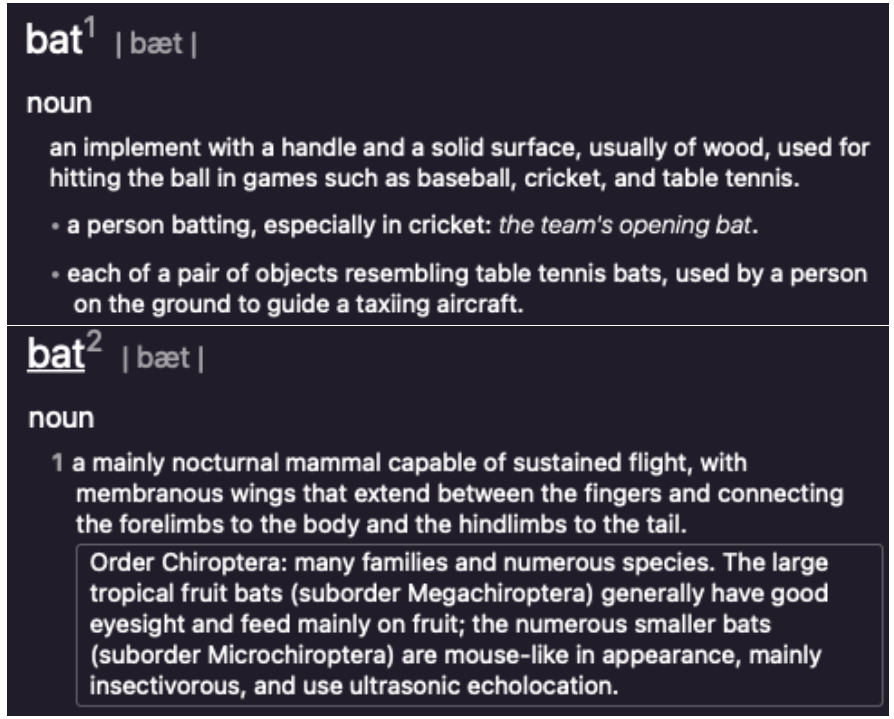
Especially at scale

Abstract Meaning Representation

One problem: wordsense

*"The boy hit the ball with a **bat**"*

*"The **bat** flew into the cave"*



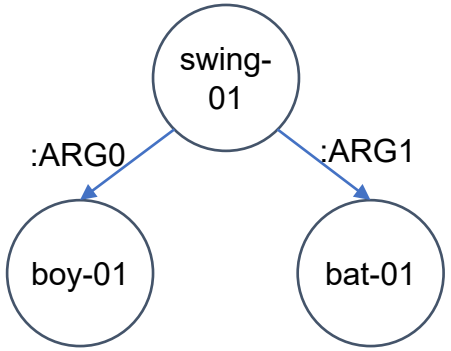
Abstract Meaning Representation

Goals:

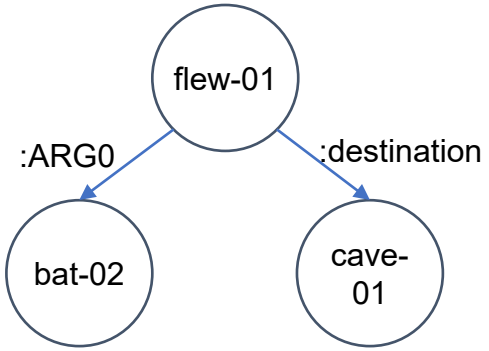
- **Large-scale**
- **Real language**
- **Abstracted from input and syntax**
- **Easy-to-process data structure**

Abstract Meaning Representation

The boy swung the bat



The bat flew into the cave



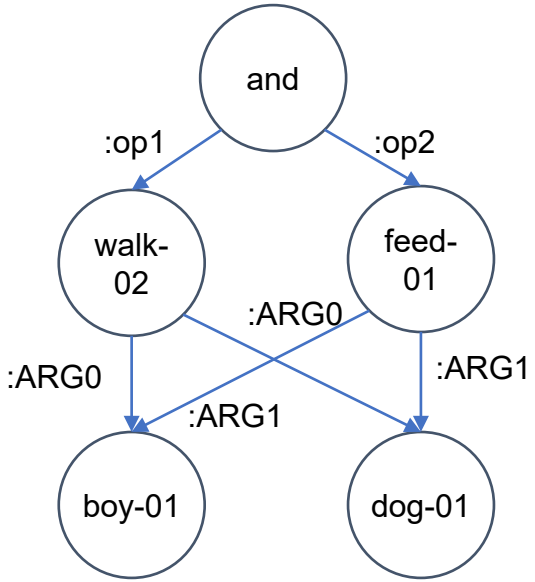
ARG0 ~= AGENT

ARG1 ~= PATIENT

[Banarescu et al., 2013]

Abstract Meaning Representation

The boy walked the dog and fed him



AMR Datasets

English

- **AMRv1, v2, v3 (LDC Corpora): ~60k sentences**
- **Little Prince: 1,562 sentences (1 book)**
- **BioAMR (PubMed) 6,952 sentences (3 papers)**

Chinese

- **Little Prince: 1,562 sentences**
- **CTBWEB: 5,015 sentences**

Spanish

- **Little Prince: 50 sentences**

Silver data

- **German, Spanish, Italian, Chinese**
- **AMR is in English!**

Why AMR? (for multilingual semantics)

Abstraction from syntax/lexicon

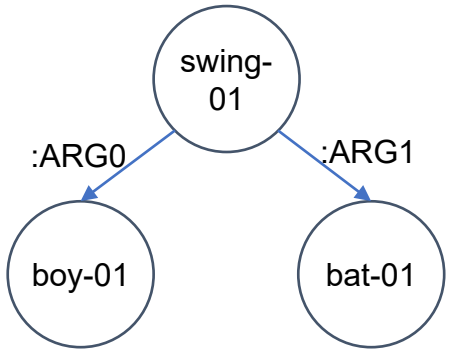
Event-centric representation

Data

The Little Prince (505 languages) 2nd only to the bible

Why not AMR? (general concerns)

The boy swung the bat



Agent, Patient

+ > 20 more roles

= very complicated guidelines

= hard to annotate (most data from LDC)

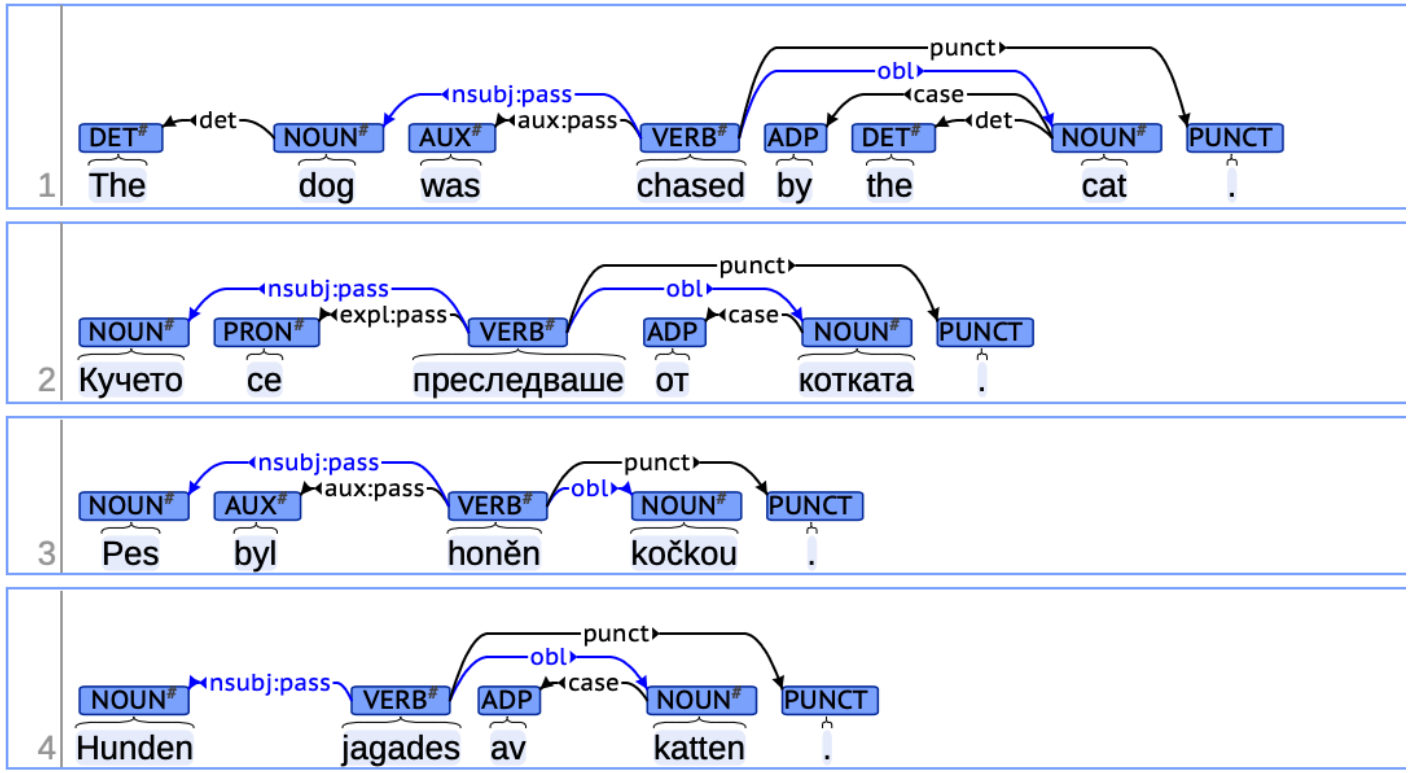
Also = brittle, hard to model

sparse data

nuanced/arbitrary decision boundaries

Universal Dependencies Detour

Universal Dependencies: Syntactic parsing



English

Bulgarian

Czech

Swedish

UD (Syntax)

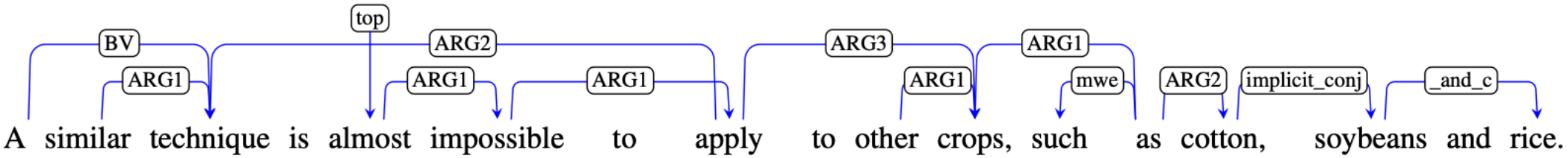
Over 100 languages

Lexicalized

Every token in input = 1 node in dependency parse tree

As a result: different per language

Universal Semantic Dependencies



(b) DELPH-IN Minimal Recursion Semantics–derived bi-lexical dependencies (DM).

[Oepen et al., 2014]

Why is this a problem?

Abstraction!

Modeling meaning

Meaning should be invariant to translation (?)

Universal Decompositional Semantics

Build an abstract graph from UD

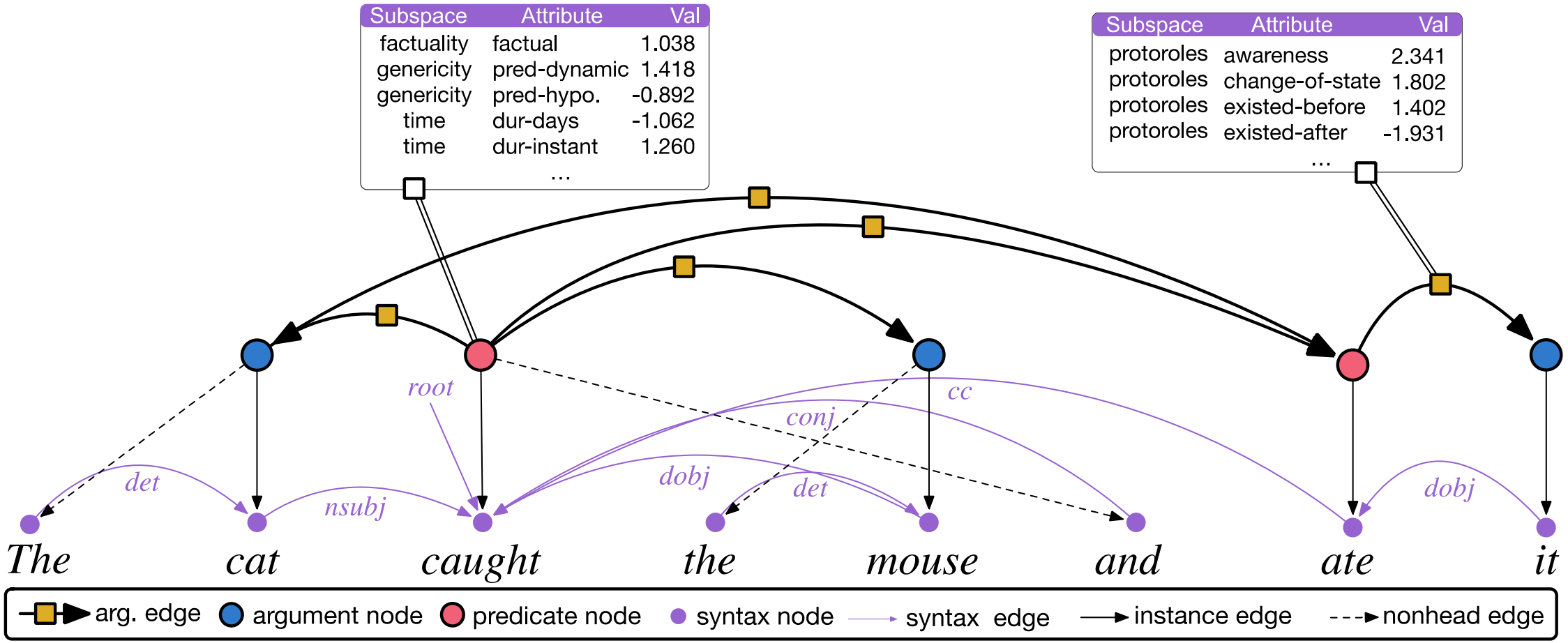
Tied to syntax, but not as closely as SDP

Remove complicated ontologies

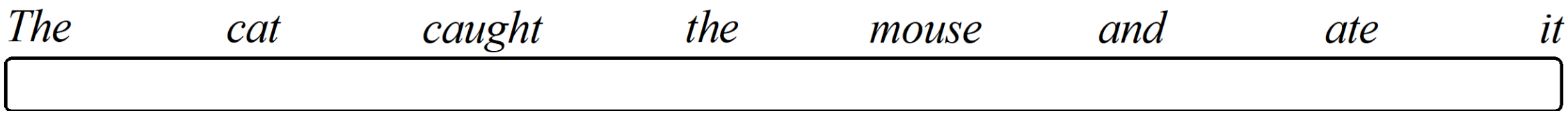
Collect graded judgements

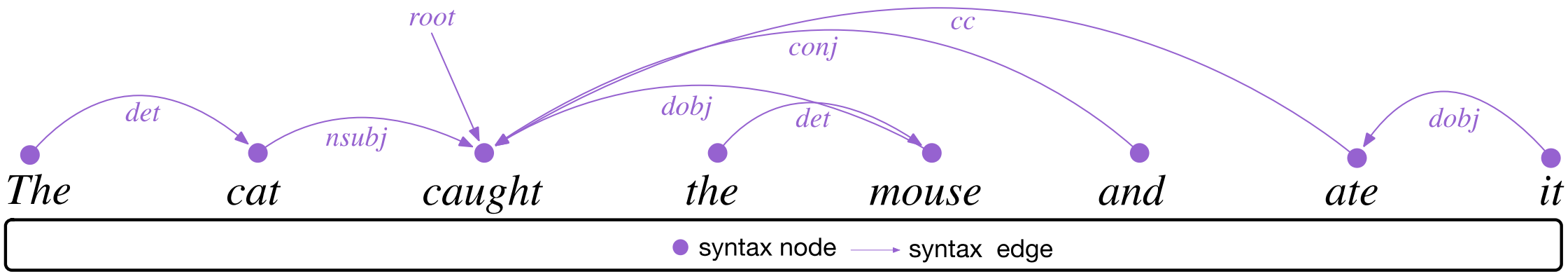
[White et al., 2016, 2020]

Universal Decompositional Semantics

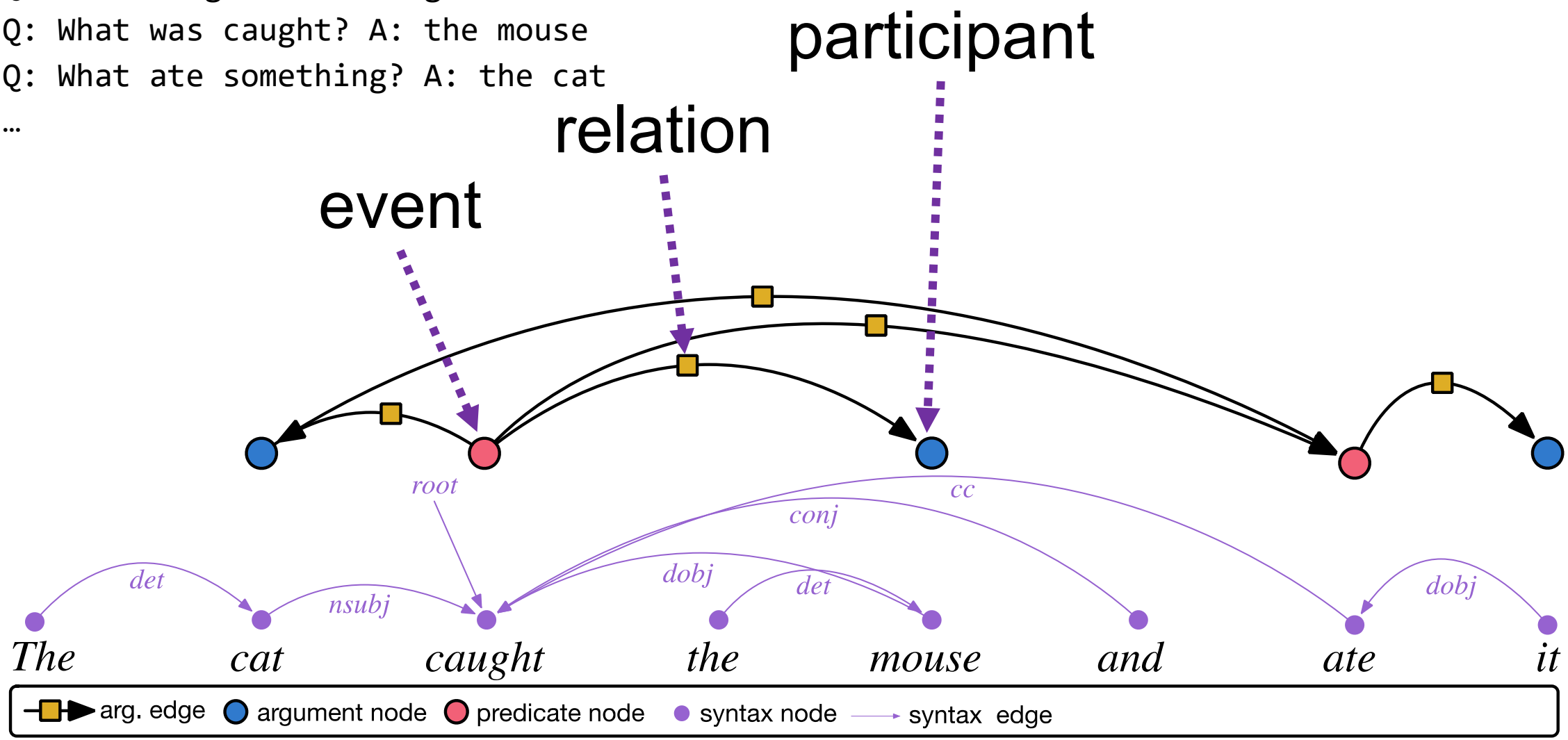


Universal Decompositional Semantics



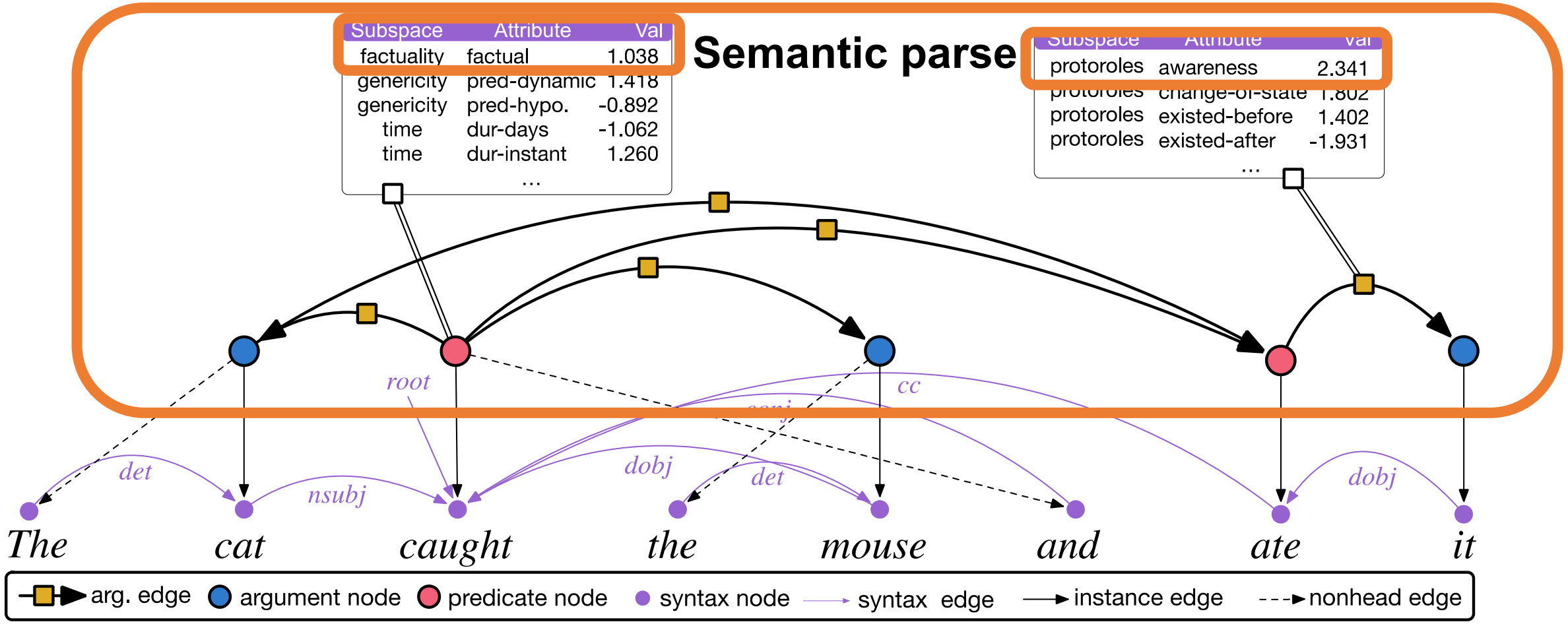

root
cc
conj
det
nsubj
dobj
det
dobj
The
cat
caught
the
mouse
and
ate
it
syntax node
syntax edge

Q: What events happened? A: Something caught something, something ate something
Q: What caught something? A: the cat
Q: What was caught? A: the mouse
Q: What ate something? A: the cat
…



Q: How likely is it the event “caught” really happened?

Q: was the mouse aware of the event “ate”?



Protoroles

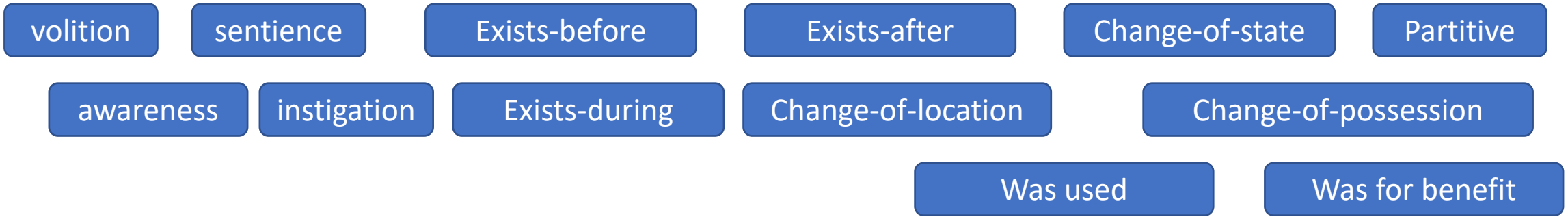
[Dowty, 1991]

Forget AGENT, PATIENT, THEME, INSTRUMENT

Forget ARG0, ARG1

Too many counter-examples

Instead, use a more expressive feature set



More agent-like (proto-agent)

More patient-like (proto-patient)

Protoroles

Annotated data from [Reisinger et al. 2015]

Data on a scale

> "***An assassin*** *in Colombia* ***killed*** *a federal judge on a Medellin street."*

```
Q: Was the assassin aware of the event?
Q: Did the assassin exist after the event?
Q: Did the assassin instigate the event?
Q: Was the assassin sentient during the event?
Q: Did the assassin act on purpose?
```

Protoroles

*"**She** was untrained and, in one botched job, **killed** a client"*

```
Q: Was she aware of the event?
Q: Did she exist after the event?
Q: Did she instigate the event?
Q: Was she sentient during the event?
Q: Did she act on purpose?
```

UDS Data

[White et al. 2020]

- **Variety of text**
- **~15k sentences**

English-only

Final MR detour: Executable parsing

Broadly: text-to-code

- **Goal is to execute program**

Common outputs/tasks

- **SQL**
- **Ordering food**
- **Scheduling**
- **Flights**

Executable parsing

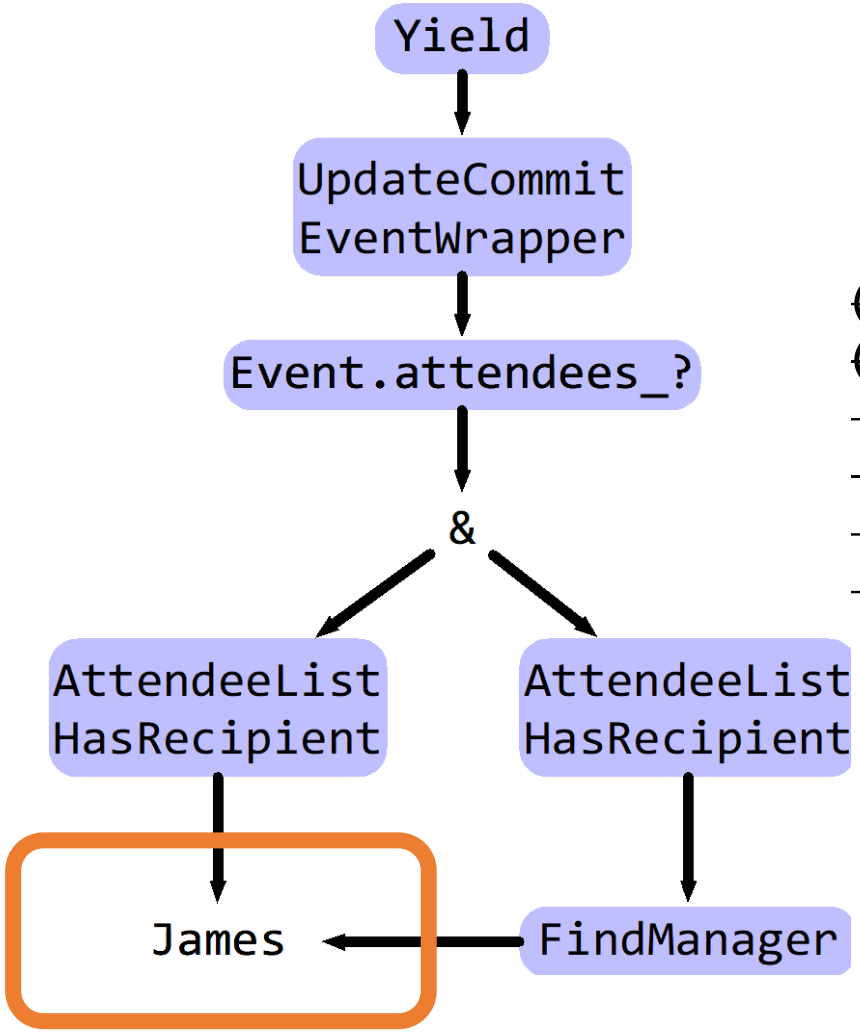
SMCalFlow Dataset

- Calendaring domain
- Lisp-like programs
- Modeling underlying DAG
- [Andreas et al. 2020]

Do I have anything going on tonight?


```
(Yield (> (size
   (QueryEventResponse.results
      (… (EventOnDateWithTimeRange
         (EventOnDate (Today)
            (^(Event) EmptyStructConstraint))
            (Night)))) 0L))
```

Execution graph



Make a meeting with James and his manager


```
(let (x0 … " James "))))
(Yield
  (UpdateCommitEventWrapper… (Event.attendees_?
      &
        (AttendeeListHasRecipient x0)
        (AttendeeListHasRecipient (FindManager x0))))))))
```

Question Break

Semantic parsing

Goal: learn to translate language →meaning representation

Approaches:

- **Alignment-based**
- **Seq2seq**
- **Seq2graph**

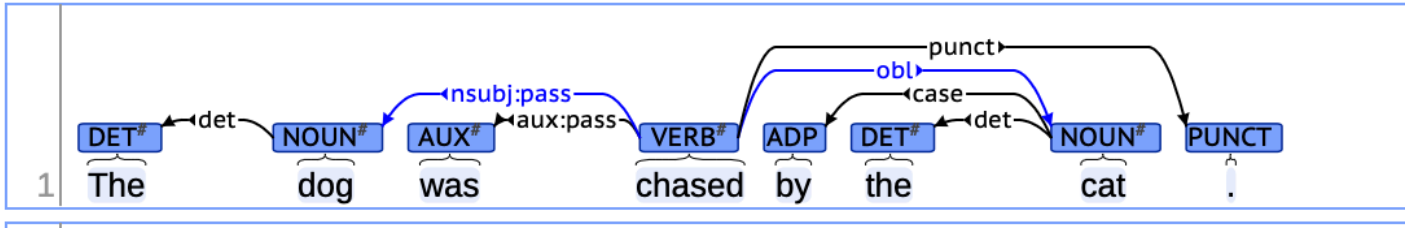
Alignment-based parsing

Older

Core idea:

Strong models for syntactic parsing

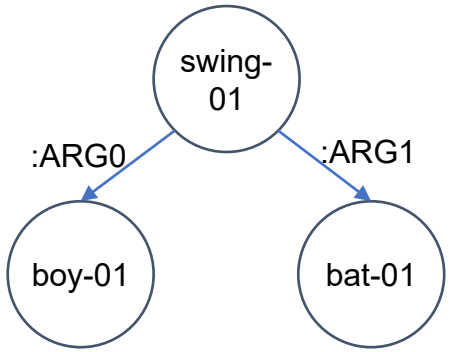
BUT syntactic parsing requires 1-1 node-token correspondence



Maybe we can align graph to input [Flanigan et al. 2014]

Alignment-based parsing

0	1	2	3	4
The	*boy*	*swung*	*the*	*bat*
[Null]	[boy-01]	[swing-01]	[null]	[bat-01]
[Null]	[ARG0-2]	[ROOT]	[null]	[ARG1-2]



Pros

- **Strong inductive bias**
- **Works with less data**

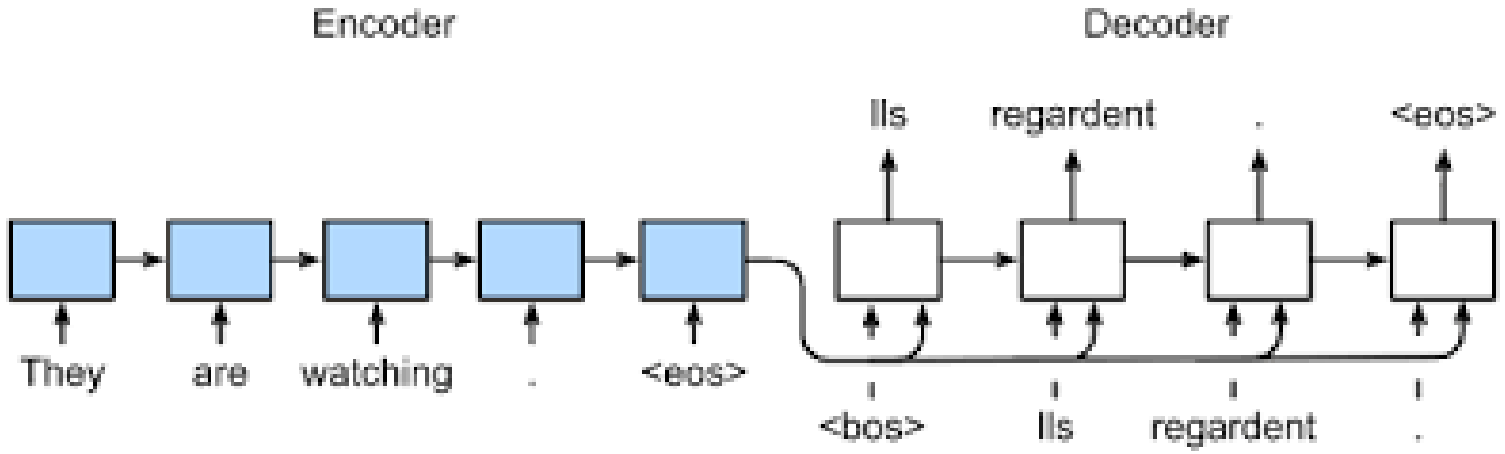
Cons

- **Requires aligner**
- **Aligner may be harder implement multilingually**

Seq2seq

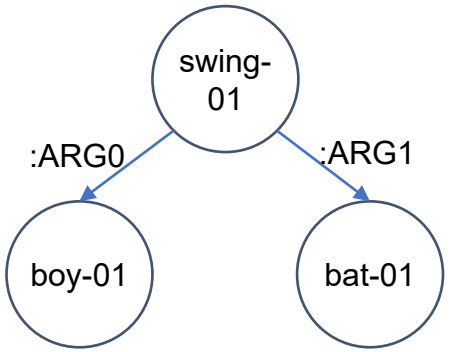
Instead of labelling node-by-node…

Encoder-decoder model



[Dong and Lapata, 2016]

Seq2seq



What's the target?

Linearize graph

```
(swing-01
      (:arg0 (boy-01)
      (:arg1 (bat-01))
```

Pros:

Seq2seq models work very well for other tasks, well-engineered

Flexible: target doesn't need to be English

No aligner needed, just one model

Cons:

Need to learn to formulate valid graphs

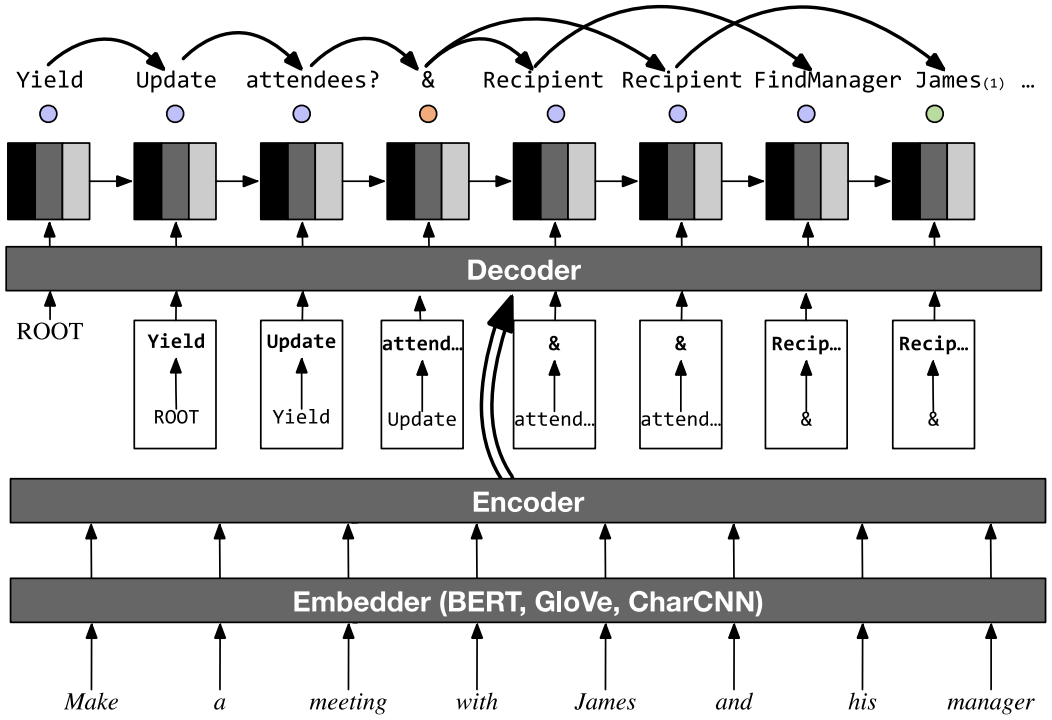
How to deal with re-entrancy?

Seq2graph

Seq2seq, except

Decode into a graph instead of sequence

[Zhang et al. 2019a]
[Zhang et al. 2019b]
[Stengel-Eskin et al. 2021]



Seq2graph

Pros:

- **Can handle re-entrant nodes**
- **Stronger inductive bias than seq2seq**
- **Easy to build in constraints**

Cons:

- **More engineering effort**
- **Harder to use pre-trained seq2seq models**

Speculative/
Future work

Important Questions

What do the semantics of different languages look like?

- **English-heavy, especially semantics**
- **Are current representations adequate for other languages?**
- **Can we use the same types of models?**

What divergences/differences exist between languages?

- **What can that tell us about semantics?**
- **How can we use that to improve models?**

Multi-view semantics

Languages differ in what they make explicit, what they leave unsaid

- **Example: Case**

English:

- **Word-order encodes semantic roles**

Latin:

- **Case-based**
- **Free word order**
- **Encodes roles in morphology**

Multi-view semantics

Brutus killed Caesar with a sword

[Agent] [Patient] [Instrument]

**With a sword killed Brutus Caesar*

**Killed Caesar Brutus with a sword*

**Caesar killed Brutus with a sword*

[Agent] and [Patient] encoded by position

[Instrument] encoded by position + “with a”

*Brut**us** gladi**o** Caesar**em** occidit*

Brut-NOM sword-ABL Caes.-ACC kill-PAST

Caesarem gladio Brutus occidit ☑

Brutus Caesarem gladio occidit ☑

Occidit Caesarem Brutus gladio ☑

[Agent] and [Patient] encoded by morphology

Brut**us** is the subject/agent

Caesar**em** is the object/patient

[Instrument] is encoded by morphology

gladi**o** is the instrument (gladius would be subject)

Multi-view semantics

Discuss:

- **Can you think of other examples?**
- **What can you do if you have access to multiple inputs in different languages?**

References

- [Banarescu et al. 2013] *Abstract meaning representation for sembanking,* Proceedings of the 7th linguistic annotation workshop and interoperability with discourse (2013)
- [Oepen et al. 2014] *SemEval 2014 Task 8: Broad-Coverage Semantic Dependency Parsing*, SemEval 2014 (2014)
- [White et al. 2016] *Universal decompositional semantics on universal dependencies.* EMNLP (2016)
- [White et al. 2020] *The Universal Decompositional Semantics Dataset and Decomp Toolkit.* LREC (2020)
- [Dowty, 1991] *Thematic Proto-Roles and Argument Selection,* Language (1991)
- [Reisinger et al. 2015], *Semantic Proto-Roles,* TACL (2015)
- [Flanigan et al. 2014], *A Discriminative Graph-Based Parser for the Abstract Meaning Representation,* ACL (2014)
- [Dong and Lapata, 2016], *Language to Logical Form with Neural Attention,* ACL (2016)
- [Zhang et al. 2019a], *AMR parsing as sequence-to-graph transduction,* ACL (2019)
- [Zhang et al. 2019b], *Broad-coverage Semantic Parsing as Transduction*, EMNLP (2019)
- [Stengel-Eskin et al. 2021], *Joint Universal Syntactic and Semantic Parsing*, TACL (2021)

References

https://github.com/nschneid/amr-tutorial

https://www.amazon.com/Petit-Prince-French-Antoine-Saint-Exupéry/dp/0152164154

https://universaldependencies.org

https://d2l.ai/chapter_recurrent-modern/seq2seq.html