Meaning Representation and Parsing

Daniel Hershcovich

DIKU Bits
February 18, 2020

Short Introduction

2005–2010

B.Sc. in Mathematics and Computer Science,

The Open University of Israel

Short Introduction

2005–2010
B.Sc. in Mathematics and Computer Science,
The Open University of Israel

2008–2019
Software Engineer
IBM Research

Short Introduction

2005–2010
B.Sc. in Mathematics and Computer Science,
The Open University of Israel

2008–2019
Software Engineer
IBM Research

2012–2019
Ph.D. in Computational Neuroscience
The Hebrew University of Jerusalem

Short Introduction

2005–2010 The Open University of Israel

2008–2019 IBM Research

2012–2019 The Hebrew University of Jerusalem

Since 2019
University of Copenhagen

IBM Project Debater (2012–2019)

AI system that can debate humans on complex topics
(e.g., **We should ban the sale of violent video games**)

5 research papers, e.g., *Context Dependent Claim Detection* (2014)

Argument Invention from First Principles (2019)

IBM Project Debater (2012–2019)

AI system that can debate humans on complex topics (e.g., **We should ban the sale of violent video games**)

> Because violence in video games is interactive and not passive, critics such as Dave Grossman and Jack Thompson argue that **violence in games hardens children to unethical acts**, calling first-person shooter games "murder simulators", although no conclusive evidence has supported this belief

5 research papers, e.g., *Context Dependent Claim Detection* (2014)

Argument Invention from First Principles (2019)

IBM Project Debater (2012–2019)

AI system that can debate humans on complex topics
(e.g., **We should ban the sale of violent video games**)

Freedom of choice → People have the right to make their own choices, including bad ones

Black market → Prohibiting products and activities makes them less visible and available, and thus less harmful

5 research papers, e.g., *Context Dependent Claim Detection* (2014)

Argument Invention from First Principles (2019)

What can we teach computers to do with language?

Translate:

Dave Grossman and Jack Thompson argue that violent games are harmful

↓

Dave Grossman og Jack Thompson hævder, at voldsomme spil er skadelige

Recognize entities:

Dave Grossman and Jack Thompson argue that violent games are harmful

Infer:

Violence in games hardens children to unethical acts

↓ entails

Violent games are harmful

What can we teach computers to do with language?

Translate:

Dave Grossman and Jack Thompson argue that violent games are harmful

↓

Dave Grossman og Jack Thompson hævder, at voldsomme spil er skadelige

Recognize entities:

Dave Grossman and Jack Thompson argue that violent games are harmful

Infer:

Violence in games hardens children to unethical acts

↓ entails

Violent games are harmful

What can we teach computers to do with language?

Translate:

Dave Grossman and Jack Thompson argue that violent games are harmful

↓

Dave Grossman og Jack Thompson hævder, at voldsomme spil er skadelige

Recognize entities:

<u>Dave Grossman</u> and <u>Jack Thompson</u> argue that violent games are harmful

Infer:

Violence in games hardens children to unethical acts

↓ entails

Violent games are harmful

What can we teach computers to do with language?

Translate:

Dave Grossman and Jack Thompson argue that violent games are harmful

$\downarrow$

Dave Grossman og Jack Thompson hævder, at voldsomme spil er skadelige

Recognize entities:

Dave Grossman and Jack Thompson argue that violent games are harmful

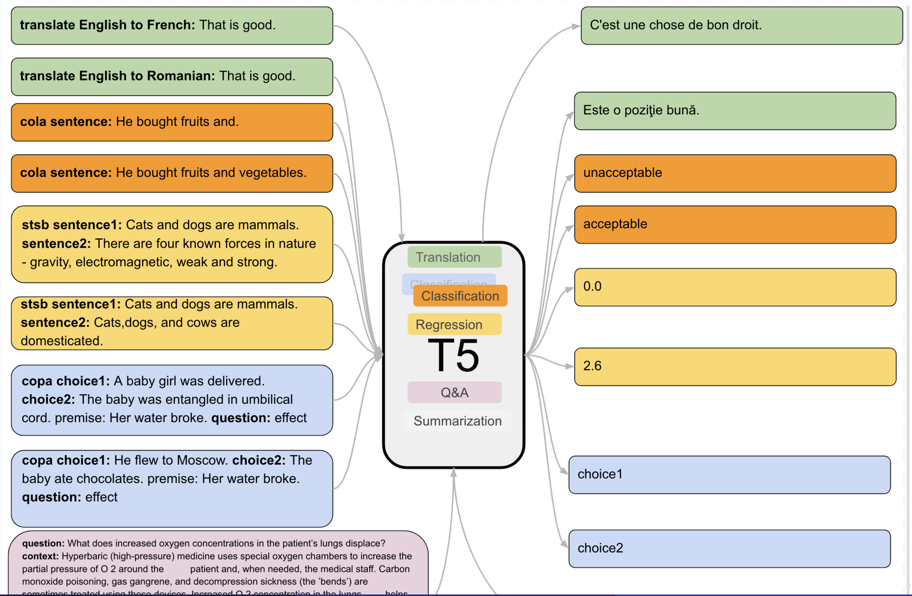
Infer:

Violence in games hardens children to unethical acts

$\downarrow$ entails

Violent games are harmful

What can we teach computers to do with language?



Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts
?
Violent games are harmful

$\downarrow$

$\downarrow$

`entailment`

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts
?
Violent games are harmful

entailment

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \rightarrow \mathbb{R}^n$$

2. Train classifiers:

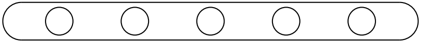
$$\mathbb{R}^n \rightarrow Y$$

3. Deploy:

$$\Sigma^* \rightarrow Y$$

Violence in games hardens children to unethical acts ? Violent games are harmful

$\downarrow$



$\downarrow$

`entailment`

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

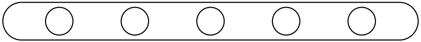
$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts
?
Violent games are harmful

$\downarrow$



$\downarrow$

`entailment`

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

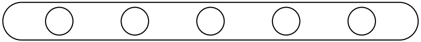
$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts
?
Violent games are harmful

$$\downarrow$$



$$\downarrow$$

`entailment`

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

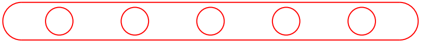
$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts ? Violent games are harmful

$\downarrow$



$\downarrow$

`entailment`

Natural Language Processing in 2020: The Basics

1. Pre-train *representations*:

$$\Sigma^* \to \mathbb{R}^n$$

2. Train classifiers:

$$\mathbb{R}^n \to Y$$

3. Deploy:

$$\Sigma^* \to Y$$

Violence in games hardens children to unethical acts
?
Violent games are harmful

Representation = vector of real numbers?

$\downarrow$

`entailment`

Learning from plain text: masked language modeling

Which Sesame Street $\boxed{?}$ is your favorite

Learning from plain text: masked language modeling

Which $\boxed{?}$ Street character is your favorite

Learning from plain text: masked language modeling

Which Sesame $\boxed{?}$ character is your favorite

Learning from plain text: masked language modeling

$\boxed{\color{red}?}$ Sesame Street character is your favorite

Learning from plain text: masked language modeling

Which Sesame Street character $\boxed{\textcolor{red}{?}}$ your favorite

Learning from plain text: masked language modeling

Which Sesame Street character is $\boxed{\textcolor{red}{?}}$ favorite

Learning from plain text: masked language modeling

Which Sesame Street character is your $\boxed{?}$

Learning from plain text: masked language modeling

Which Sesame Street character is your favorite

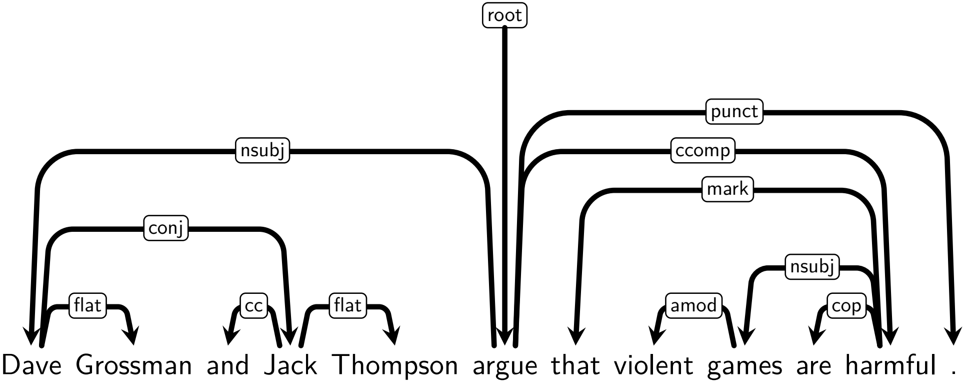
BERT (Bidirectional Encoder Representations from Transformers):

- Trained on 16GB of text.
- 16 TPU chips for 4 days.

https://demo.allennlp.org/masked-lm

What can we teach computers to do with language?

Identify relations between concepts (**parsing**, various frameworks):



Breaking it down

[Meaning], [Representation] and [Parsing]
1. What we mean, 2. How to represent (something), 3. How to parse (something)

or

[Meaning Representation] and [Parsing]
1. How to represent what we mean, 2. How to parse (something)

or

[Meaning [Representation and Parsing]]
1. How to represent what we mean, 2. How to parse what we mean

or

[Meaning Representation] and [Parsing *(to Meaning Representation)*]
1. How to represent what we mean, 2. How to parse (1)

Breaking it down

[Meaning], [Representation] and [Parsing]
1. What we mean, 2. How to represent (something), 3. How to parse (something)

or

[Meaning Representation] and [Parsing]
1. How to represent what we mean, 2. How to parse (something)

or

[Meaning [Representation and Parsing]]
1. How to represent what we mean, 2. How to parse what we mean

or

[Meaning Representation] and [Parsing *(to Meaning Representation)*]
1. How to represent what we mean, 2. How to parse (1)

Breaking it down

[Meaning], [Representation] and [Parsing]
1. What we mean, 2. How to represent (something), 3. How to parse (something)

or

[Meaning Representation] and [Parsing]
1. How to represent what we mean, 2. How to parse (something)

or

[Meaning [Representation and Parsing]]
1. How to represent what we mean, 2. How to parse what we mean

or

[Meaning Representation] and [Parsing *(to Meaning Representation)*]
1. How to represent what we mean, 2. How to parse (1)

Breaking it down

[Meaning], [Representation] and [Parsing]
1. What we mean, 2. How to represent (something), 3. How to parse (something)

or

[Meaning Representation] and [Parsing]
1. How to represent what we mean, 2. How to parse (something)

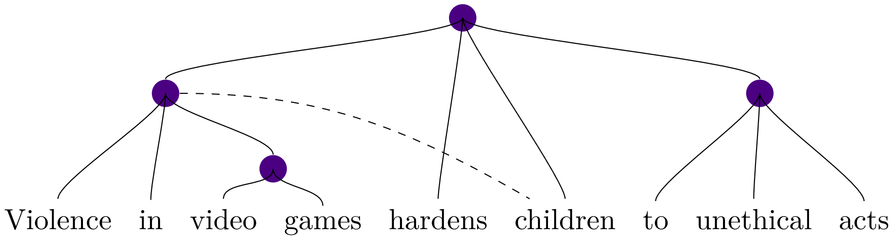
or

[Meaning [Representation and Parsing]]
1. How to represent what we mean, 2. How to parse what we mean

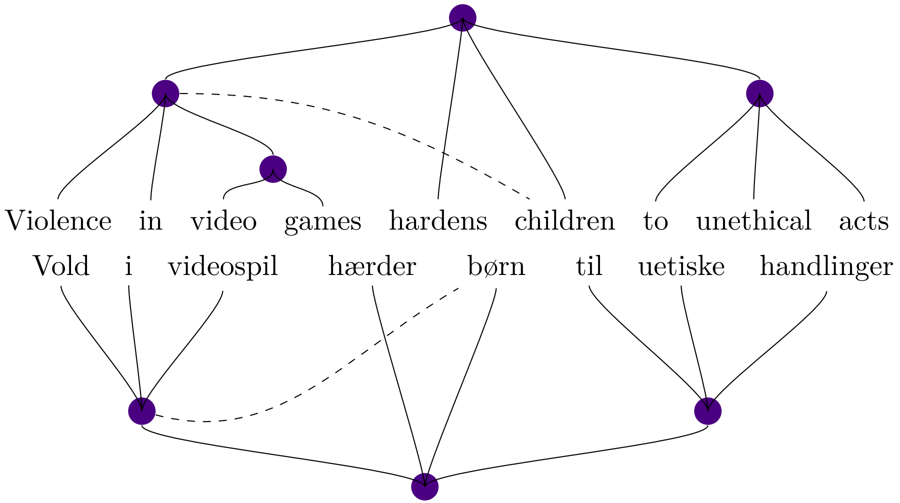
or

[Meaning Representation] and [Parsing *(to Meaning Representation)*]
1. How to represent what we mean, 2. How to parse (1)

Meaning Representation

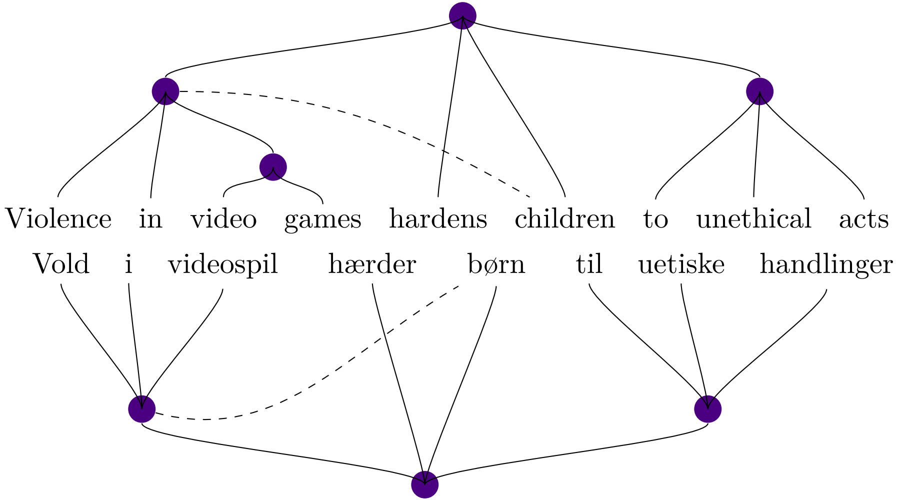


Meaning Representation

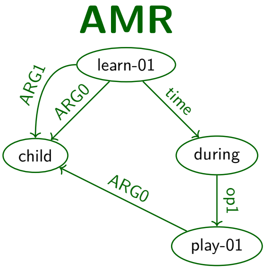


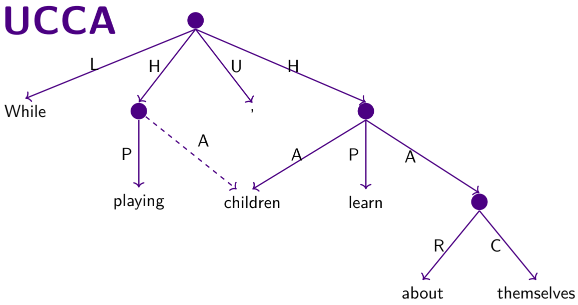
Meaning Representation

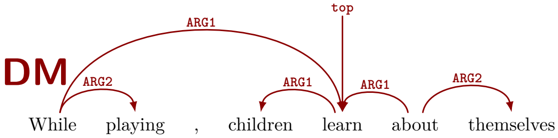
Universal Conceptual Cognitive Annotation (UCCA):



Many meaning representation frameworks exist







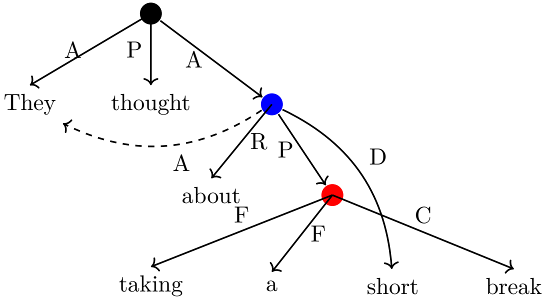
Parsing

A Transition-Based Directed Acyclic Graph Parser for UCCA (2017)

`http://bit.ly/tupademo`

Parsing

A Transition-Based Directed Acyclic Graph Parser for UCCA (2017)
`http://bit.ly/tupademo`



$\Updownarrow$

SHIFT, RIGHT-EDGE$_A$, SHIFT, SWAP, RIGHT-EDGE$_P$, REDUCE, SHIFT, SHIFT, NODE$_R$, REDUCE, LEFT-REMOTE$_A$, SHIFT, SHIFT, NODE$_C$, REDUCE, SHIFT, RIGHT-EDGE$_P$, SHIFT, RIGHT-EDGE$_F$, REDUCE, SHIFT, SWAP, RIGHT-EDGE$_D$, REDUCE, SWAP, RIGHT-EDGE$_A$, REDUCE, REDUCE, SHIFT, REDUCE, SHIFT, RIGHT-EDGE$_C$, FINISH

TUPA: Transition-based UCCA Parser

Parses text $w_1 \ldots w_n$ to graph G incrementally by applying transitions to the parser state, consisting of: stack, buffer and constructed graph.

TUPA: Transition-based UCCA Parser

Parses text $w_1 \dots w_n$ to graph G incrementally by applying transitions to the parser state, consisting of: stack, buffer and constructed graph.

Initial state:

stack

●

buffer

They	thought	about	taking	a	short	break

TUPA: Transition-based UCCA Parser

Parses text $w_1 \dots w_n$ to graph G incrementally by applying transitions to the parser state, consisting of: stack, buffer and constructed graph.

Initial state:

stack

●

buffer

They	thought	about	taking	a	short	break

Transitions:

{Shift, Reduce, Node$_X$, Left-Edge$_X$, Right-Edge$_X$, Left-Remote$_X$, Right-Remote$_X$, Swap, Finish}

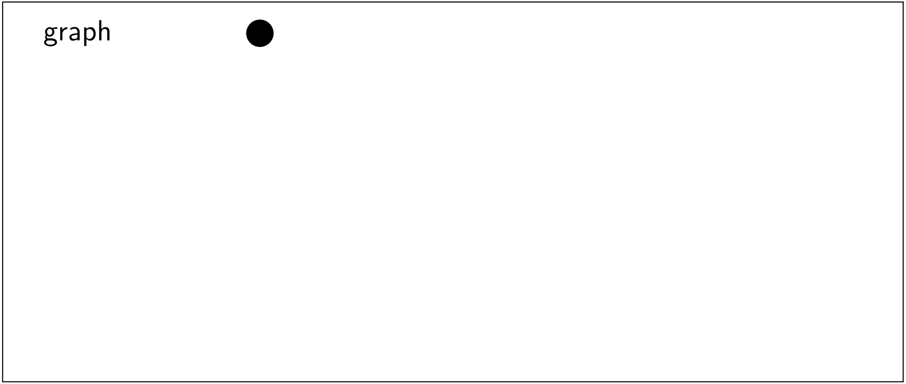
⇒ Shift

stack

●	They

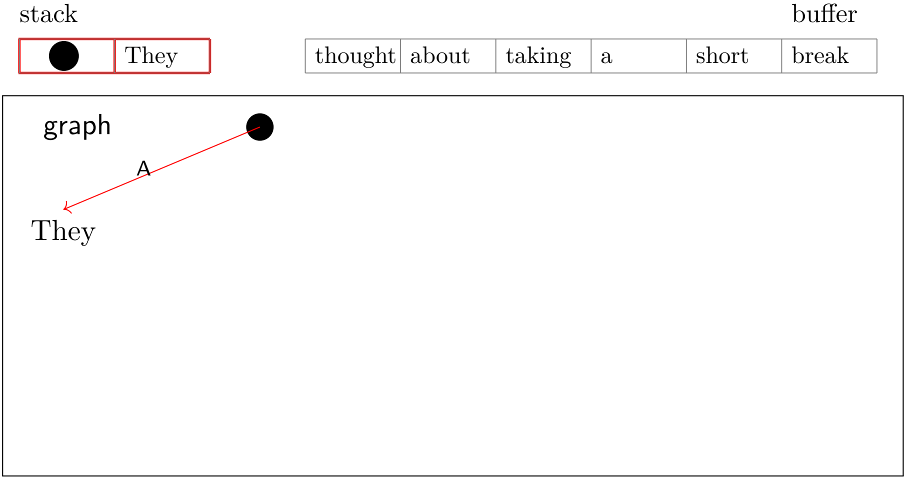
buffer

thought	about	taking	a	short	break



Example: TUPA Transition Sequence

⇒ RIGHT-EDGE$_A$

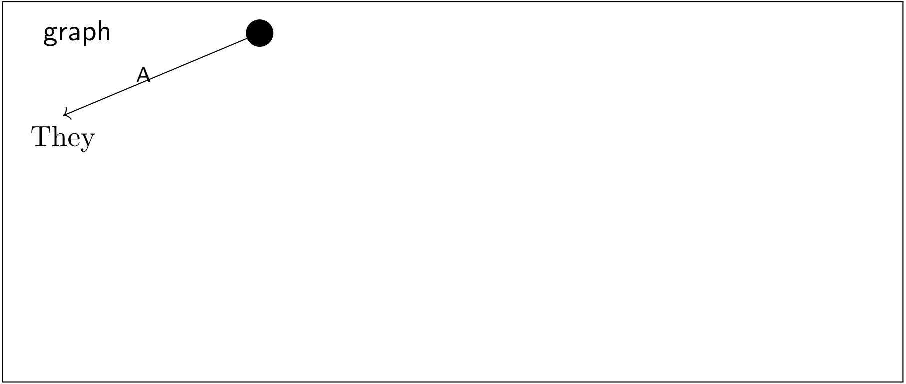


Example: TUPA Transition Sequence

⇒ SHIFT

stack: ● | They | thought

buffer: about | taking | a | short | break

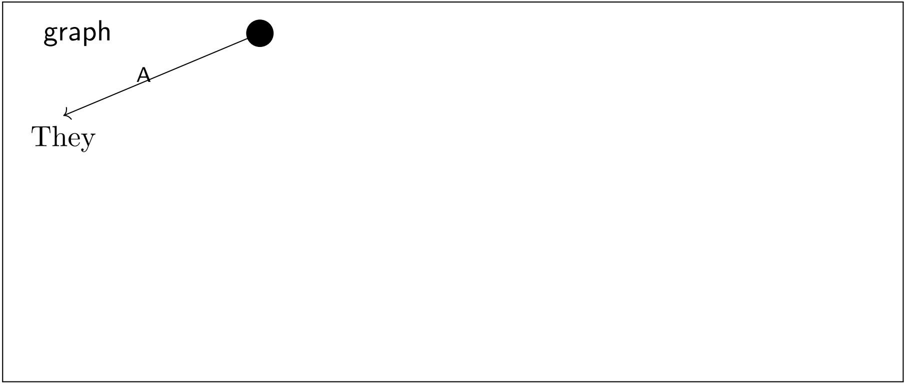


Example: TUPA Transition Sequence

⇒ Swap

stack	
●	thought

					buffer
They	about	taking	a	short	break



Example: TUPA Transition Sequence

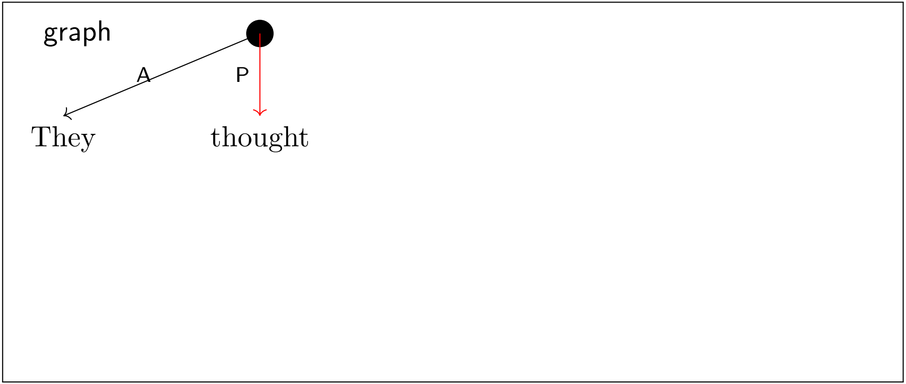
⇒ RIGHT-EDGE$_P$

stack

●	thought

buffer

They	about	taking	a	short	break



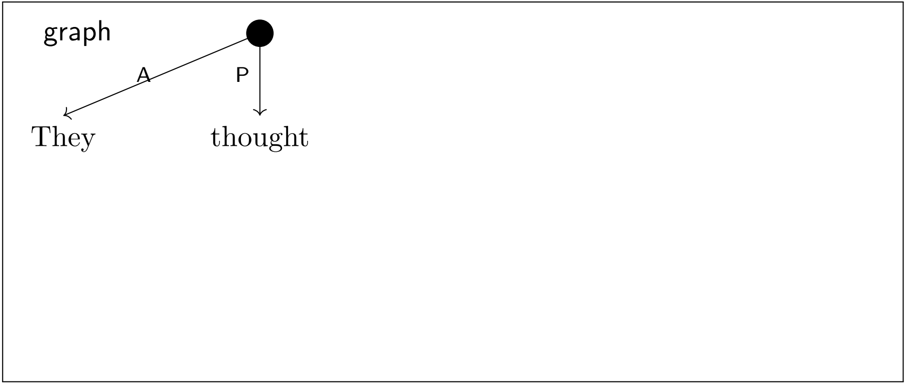
Example: TUPA Transition Sequence

⇒ REDUCE

stack

buffer

They	about	taking	a	short	break

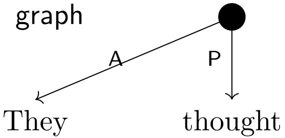


Example: TUPA Transition Sequence

⇒ SHIFT

stack							buffer
●	They	about	taking	a	short	break	

graph



Example: TUPA Transition Sequence

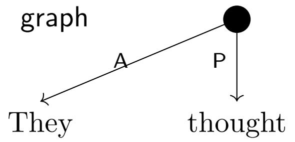
$\Rightarrow$ Shift

stack

●	They	$\boxed{\text{about}}$

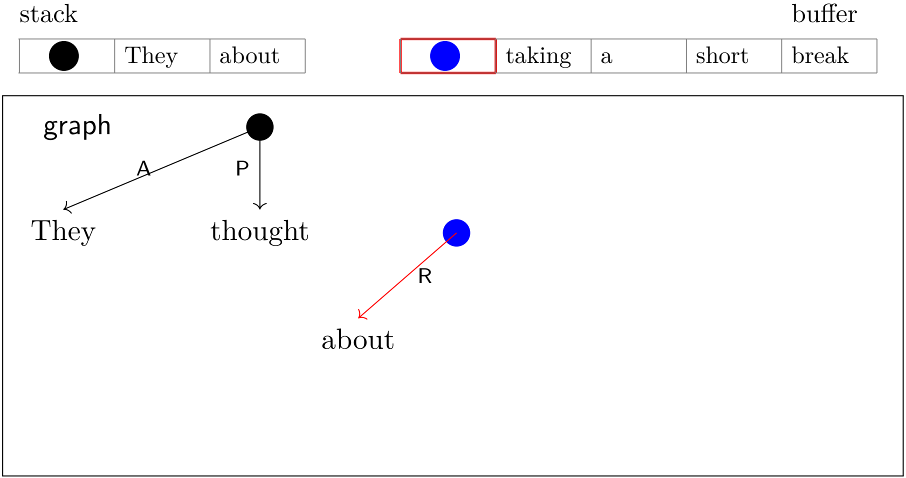
buffer

taking	a	short	break



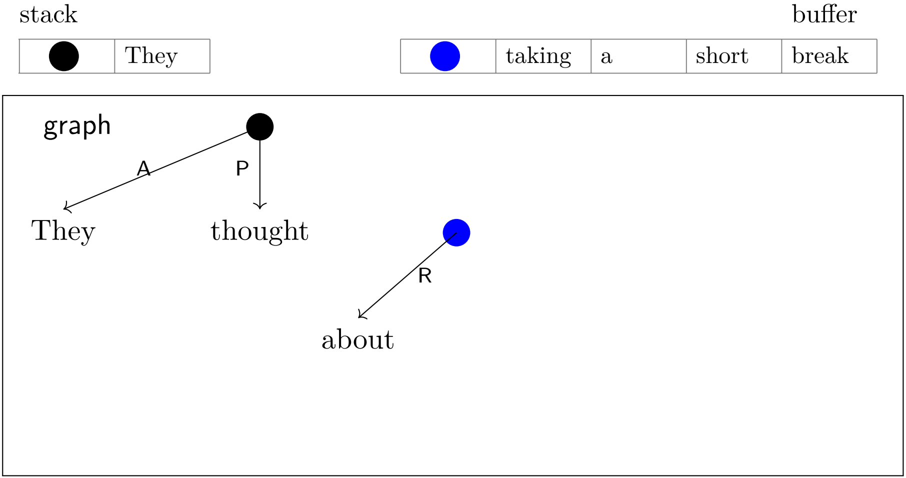
Example: TUPA Transition Sequence

$\Rightarrow$ NODE$_R$



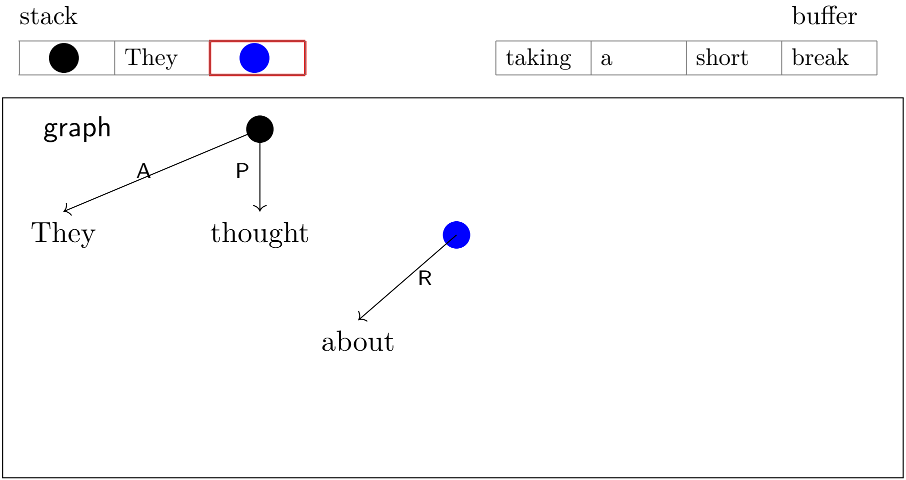
Example: TUPA Transition Sequence

⇒ REDUCE



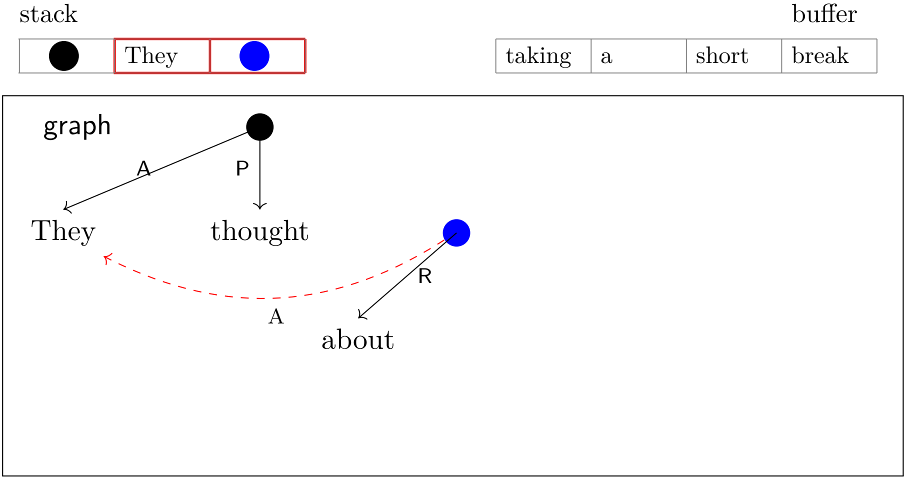
Example: TUPA Transition Sequence

⇒ Shift



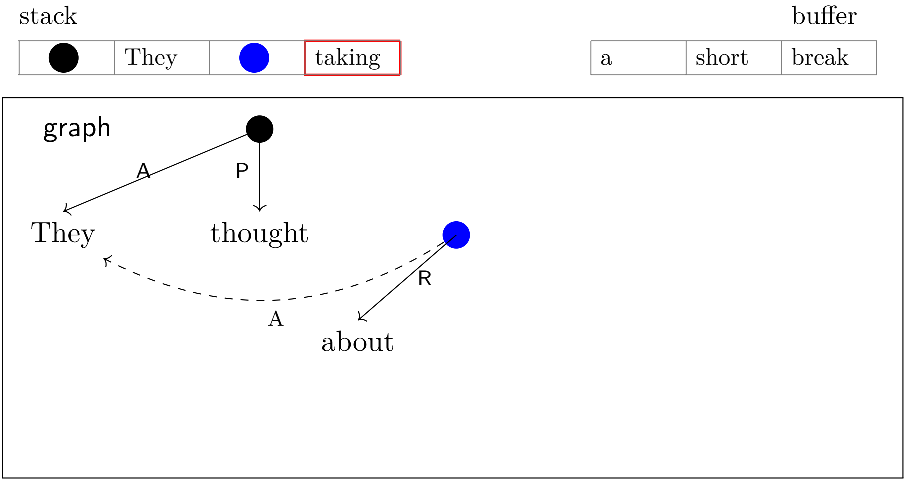
Example: TUPA Transition Sequence

⇒ LEFT-REMOTE$_A$



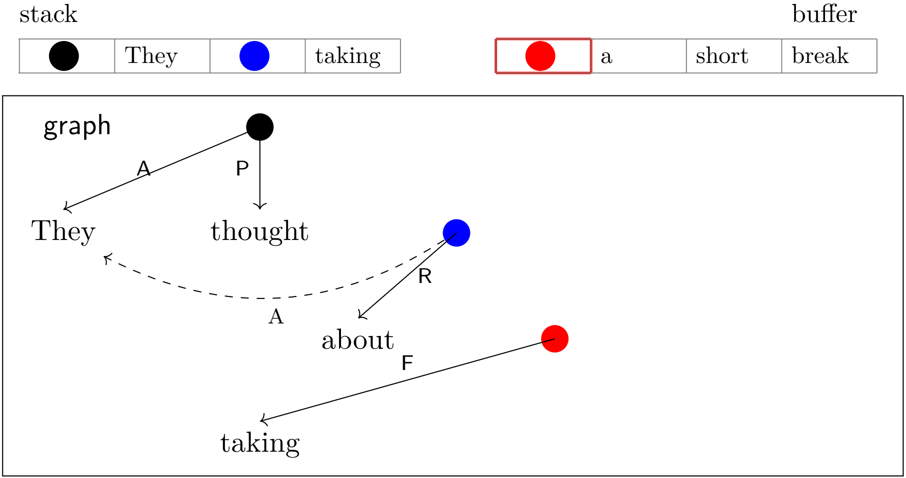
Example: TUPA Transition Sequence

⇒ Shift



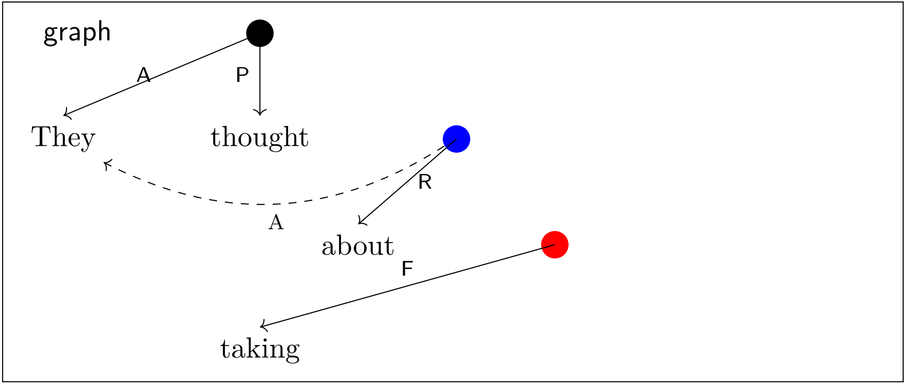
Example: TUPA Transition Sequence

$\Rightarrow$ NODE$_C$



Example: TUPA Transition Sequence

⇒ Reduce



Example: TUPA Transition Sequence

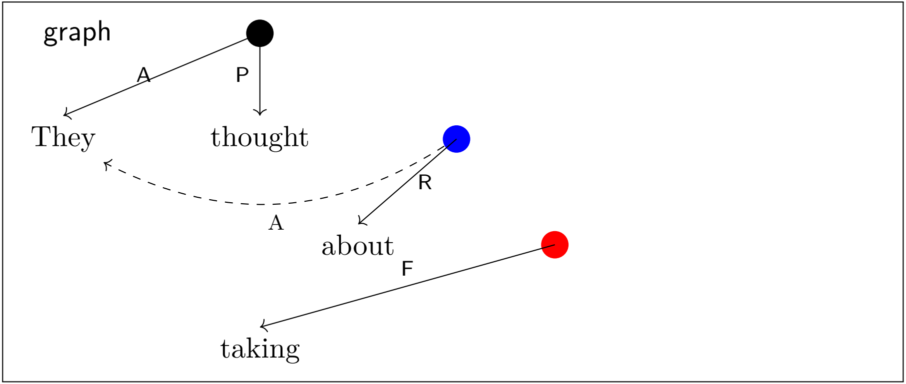
⇒ Shift

stack

●	They	●	●

buffer

a	short	break

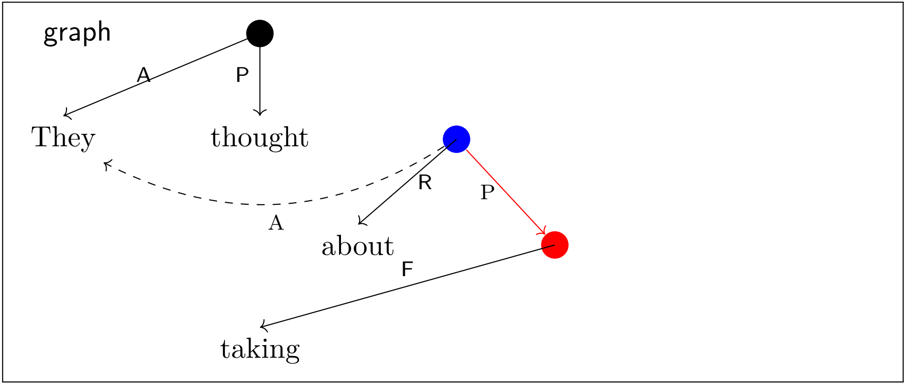


Example: TUPA Transition Sequence

⇒ RIGHT-EDGE$_P$

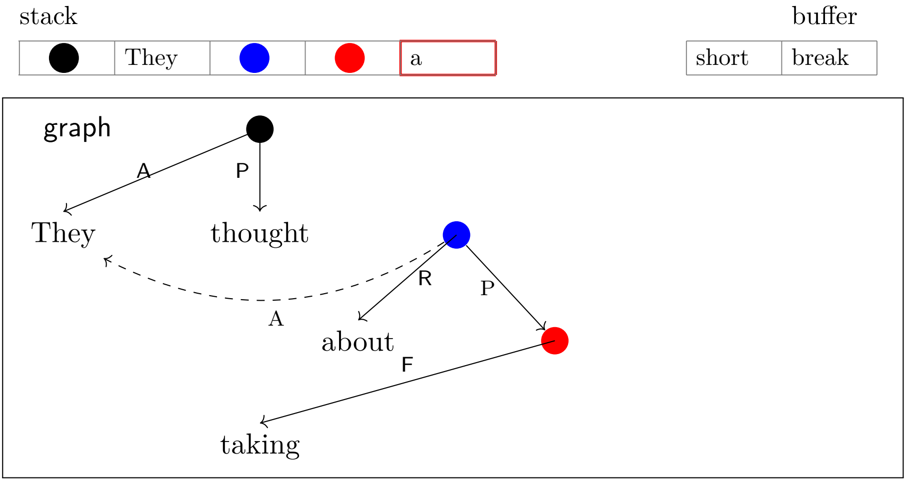
stack			
●	They	●	●

		buffer
a	short	break



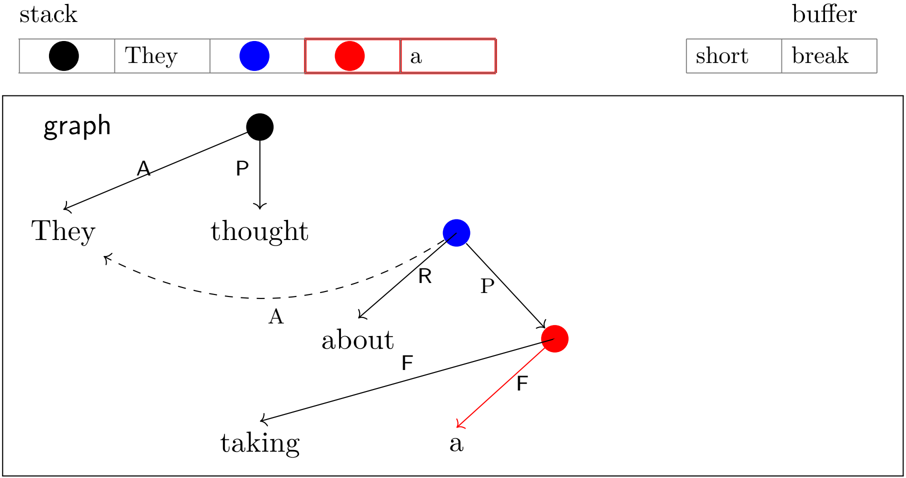
Example: TUPA Transition Sequence

$\Rightarrow$ SHIFT



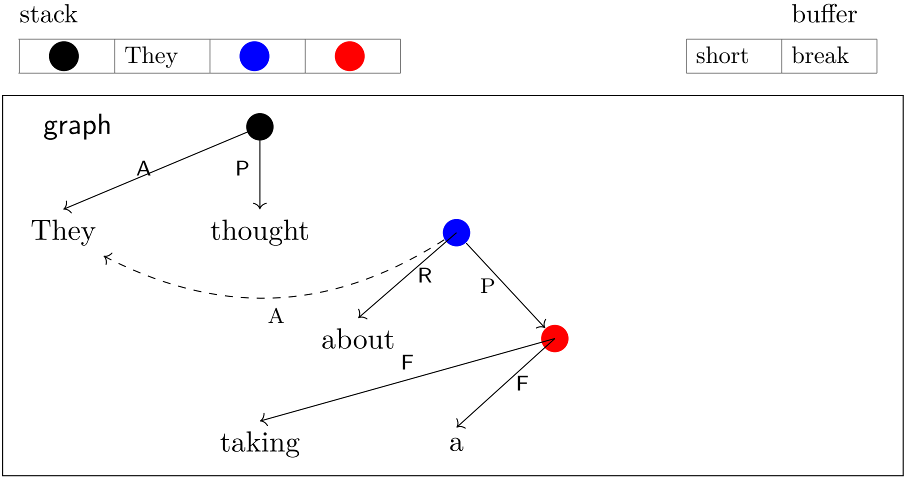
Example: TUPA Transition Sequence

⇒ RIGHT-EDGE$_F$



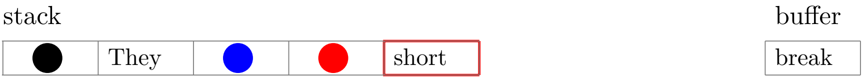
Example: TUPA Transition Sequence

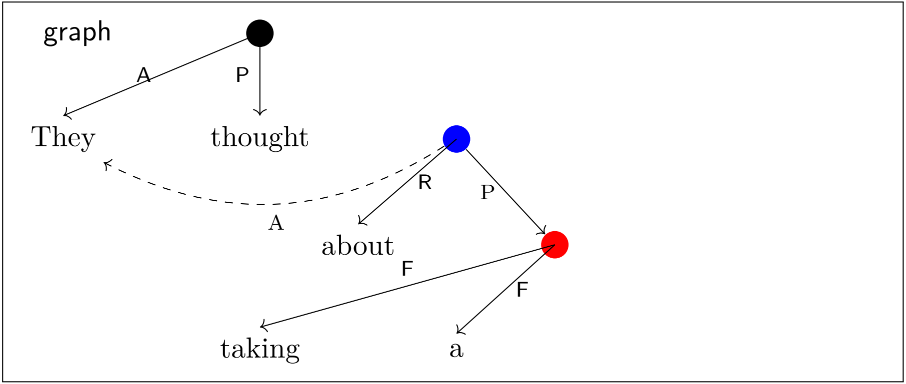
⇒ REDUCE



Example: TUPA Transition Sequence

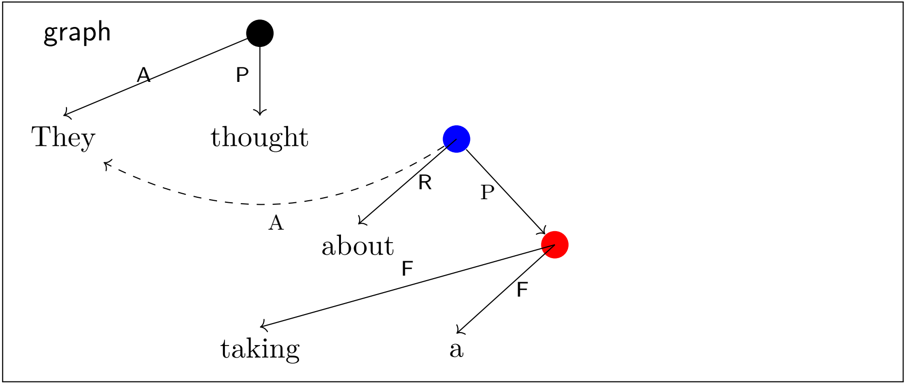
⇒ Shift





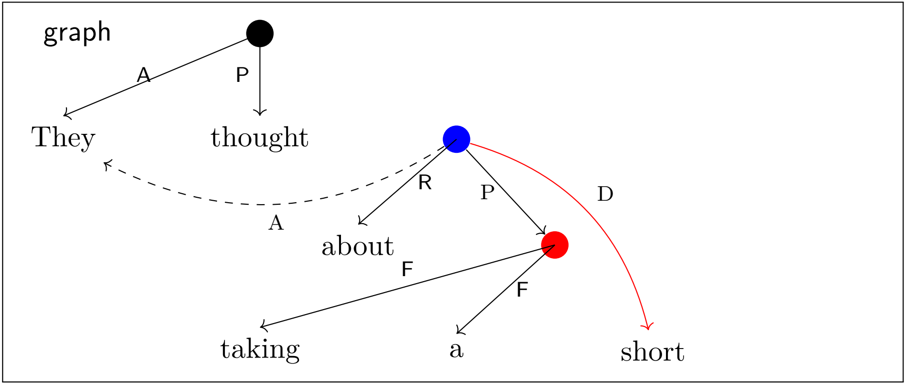
Example: TUPA Transition Sequence

⇒ SWAP



Example: TUPA Transition Sequence

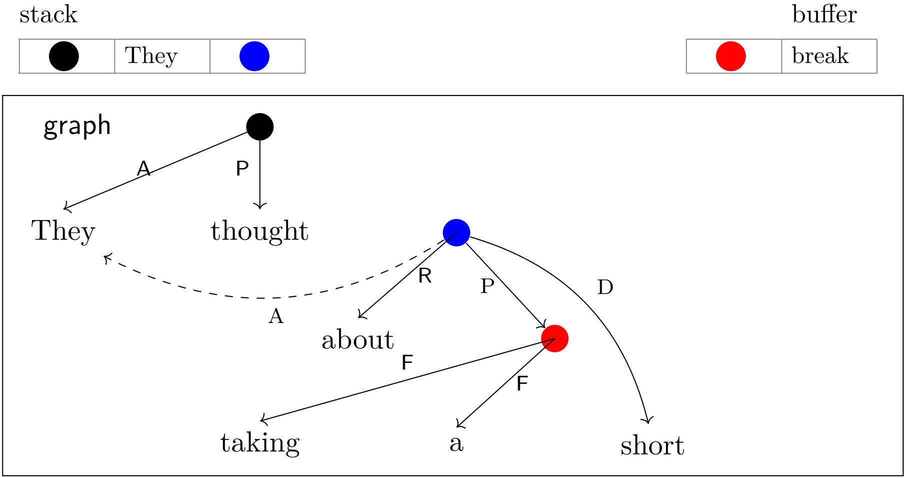
⇒ $\textsc{Right-Edge}_D$



Example: TUPA Transition Sequence

⇒ REDUCE

stack		buffer

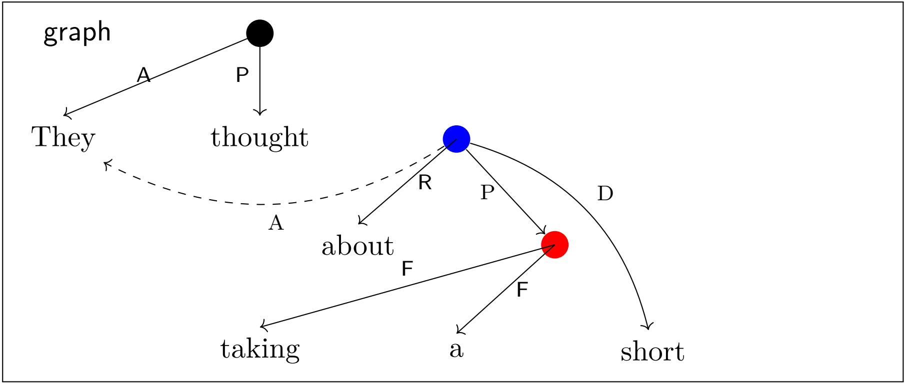


Example: TUPA Transition Sequence

$\Rightarrow$ SWAP

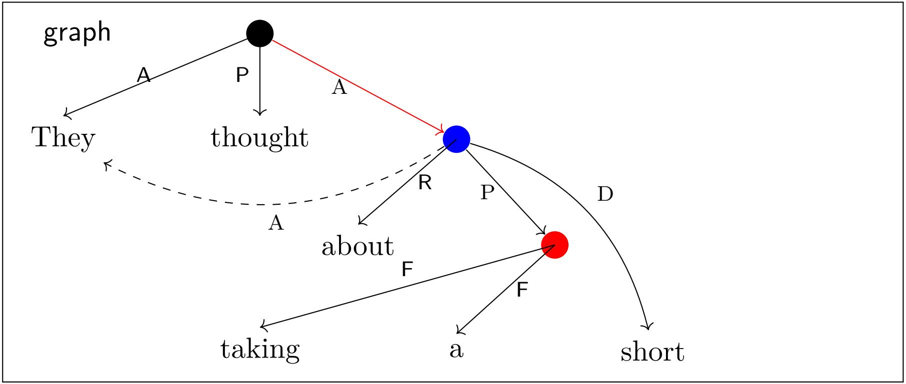
stack

buffer



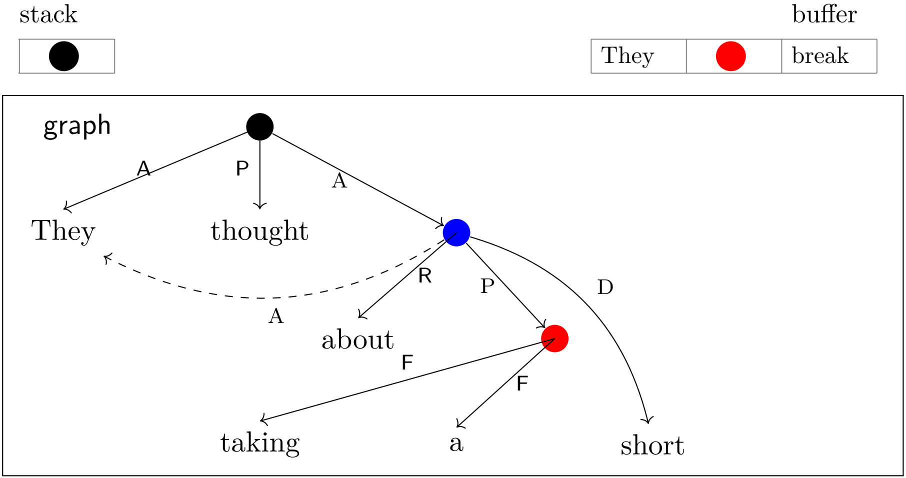
⇒ RIGHT-EDGE$_A$


stack
buffer
They
break


graph
A
P
A
They
thought
R
P
D
A
about
F
F
taking
a
short

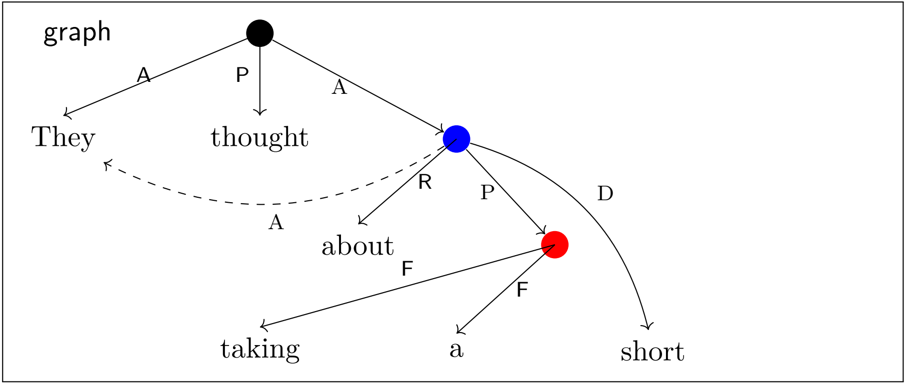
Example: TUPA Transition Sequence

⇒ REDUCE



Example: TUPA Transition Sequence

⇒ REDUCE



Example: TUPA Transition Sequence

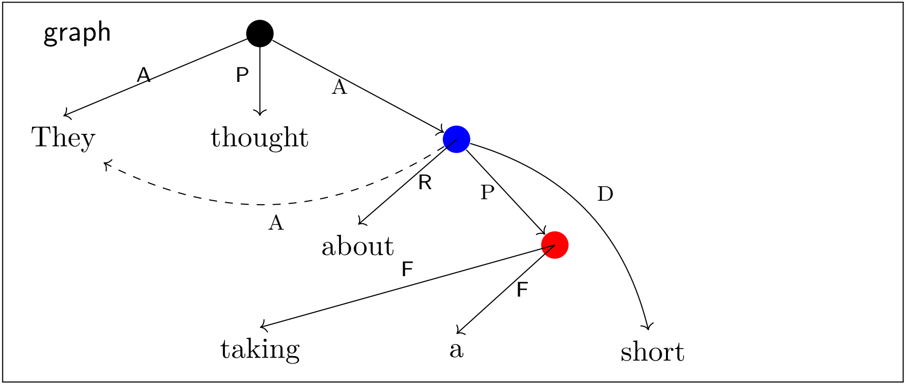
$\Rightarrow$ SHIFT

stack

They

buffer

●	break

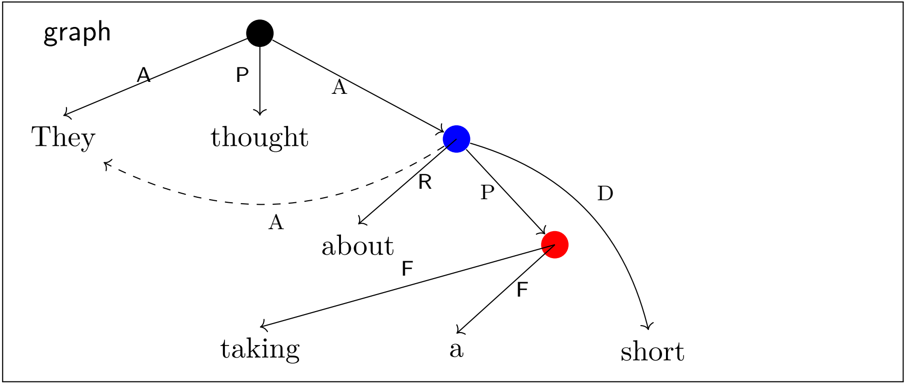


Example: TUPA Transition Sequence

⇒ REDUCE

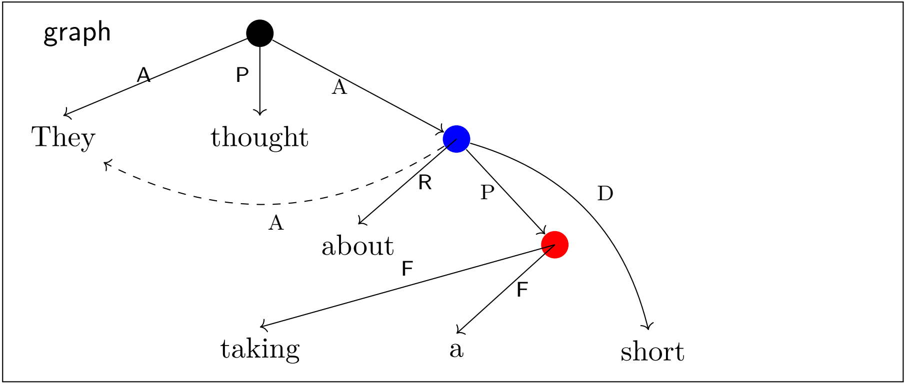
stack

buffer



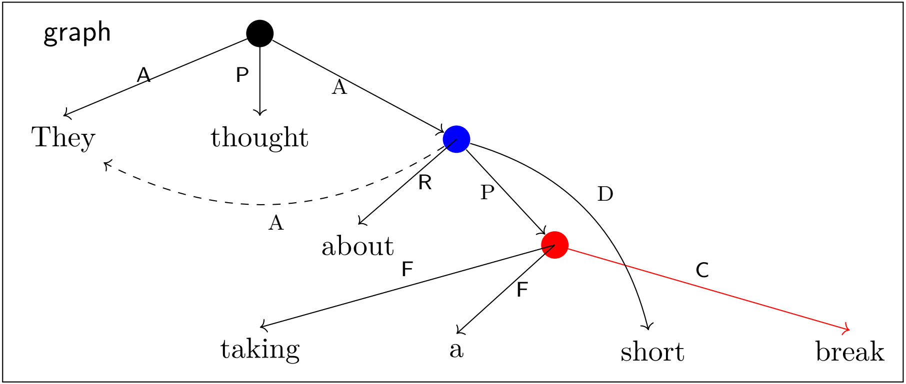
Example: TUPA Transition Sequence

⇒ SHIFT



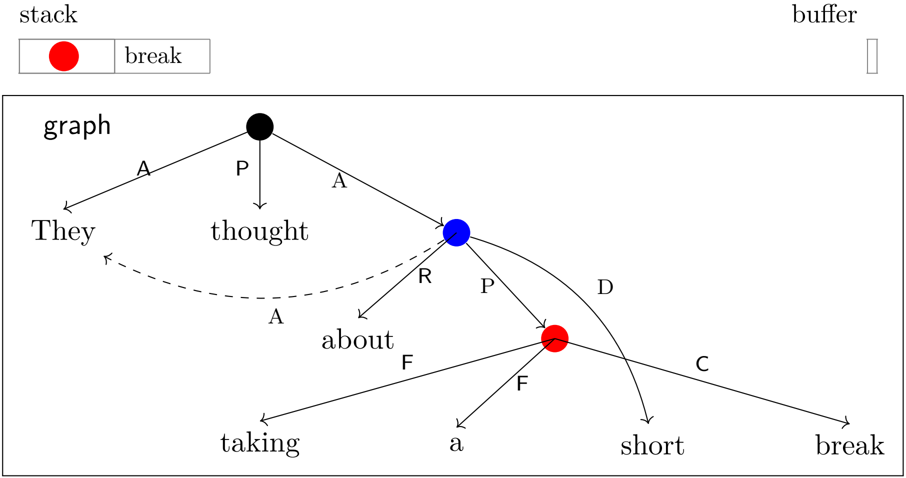
Example: TUPA Transition Sequence

⇒ RIGHT-EDGE$_C$



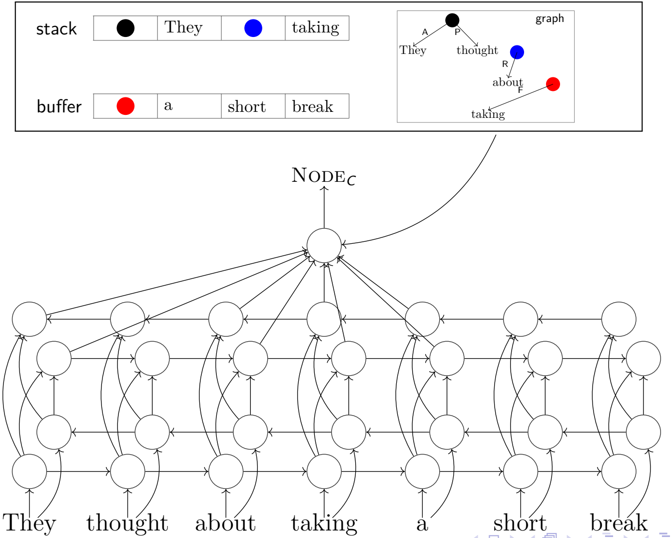
Example: TUPA Transition Sequence

⇒ FINISH



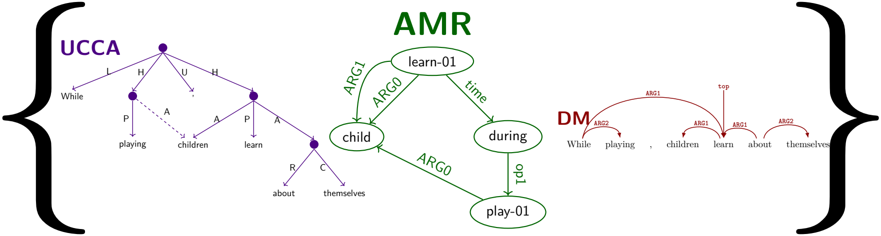
TUPA model

Learns to predict next transition based on current state.



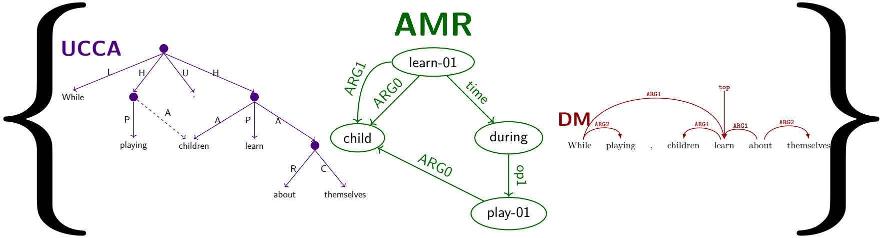
Sharing for better generalization

Multitask Parsing Across Semantic Representations (2018)



Sharing for better generalization

Multitask Parsing Across Semantic Representations (2018)



Improved UCCA parsing in English, French and German.

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA

- 3 languages.
- 8 teams from 6 countries.

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA

- 3 languages.
- 8 teams from 6 countries.

MRP 2019: Cross-Framework Meaning Representation Parsing

- 5 frameworks.
- 18 teams from 8 countries.

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA

- 3 languages.
- 8 teams from 6 countries.

MRP 2019: Cross-Framework Meaning Representation Parsing

- 5 frameworks.
- 18 teams from 8 countries.

soon...

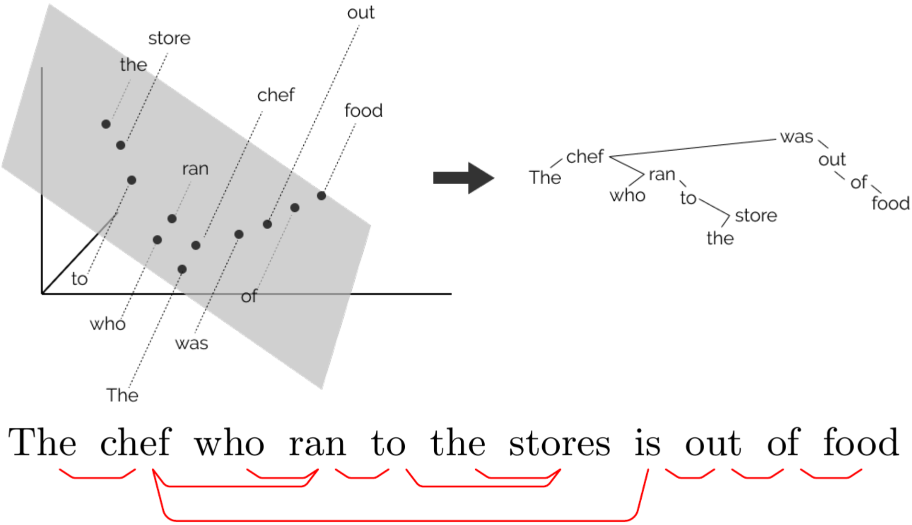
MRP 2020

- More frameworks.
- 5 languages.

What can meaning representation do for NLP?

- Probing for linguistic knowledge
- Querying knowledge bases
- Better machine translation

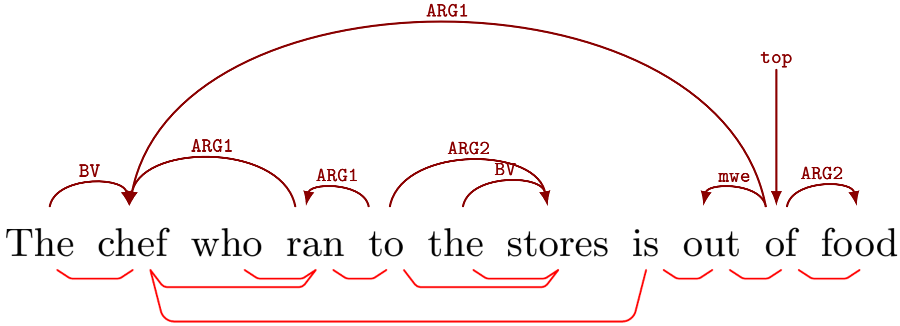
Probing for linguistic knowledge



https://nlp.stanford.edu/~johnhew/structural-probe.html

Probing for linguistic knowledge

Are meaning representations implicitly learned by pretrained encoders?



https://nlp.stanford.edu/~johnhew/structural-probe.html

Querying knowledge bases

Executable meaning representations: SQL, SPARQL

impressionist painters
18th-century American poets
birthplace of the writer of Tarzan
...

Querying knowledge bases

Executable meaning representations: SQL, SPARQL

impressionist painters
18th-century American poets
birthplace of the writer of Tarzan
...

WIKIDATA

Querying knowledge bases

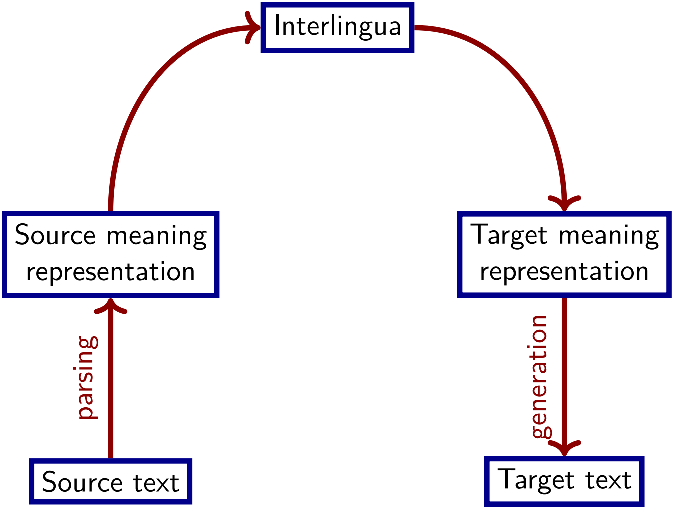
Executable meaning representations: SQL, SPARQL

```
SELECT DISTINCT ?painter ?painterLabel (count (DISTINCT ?exhibition) as ?
exhibition_count)
(group_concat(DISTINCT ?exhibitionLabel; separator=", ") as ?exhibitions)
WHERE {
  ?painter wdt:P106 wd:Q1028181 . #give me all people with occupation (P106) painter
(Q1028181)
  ?painter wdt:P135 wd:Q40415 . #who belonged to the impressionist (Q40415) movement
(P135)
  ?painting wdt:P170  ?painter . #the paintings created by (P170) the painter
  ?painting wdt:P608 ?exhibition . #have an exhibition history (P608) at an exhibition
  ?exhibition rdfs:label ?exhibitionLabel . #give me the english Labels of these
exhibitions, if possible
  FILTER (lang(?exhibitionLabel) = "en")

  SERVICE wikibase:label {bd:serviceParam wikibase:language "en".}
} GROUP BY ?painter ?painterLabel
```

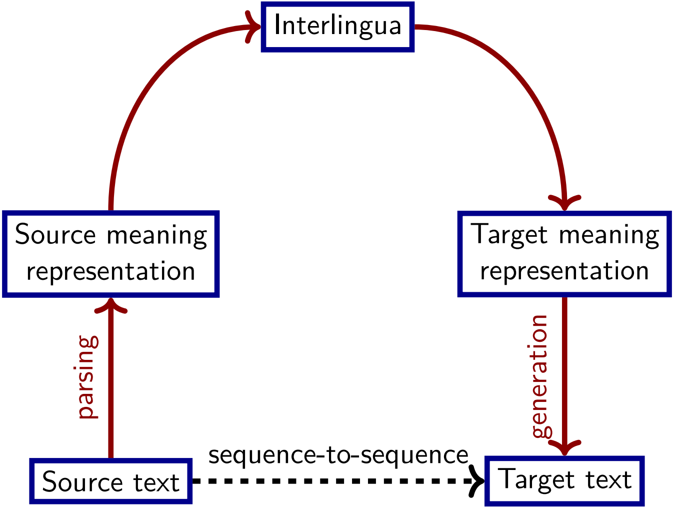
Better machine translation

(Simplified) *Vauquois triangle*:



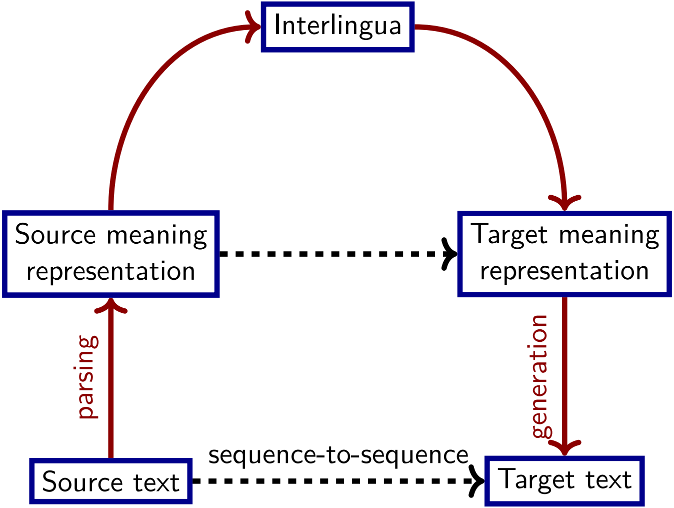
Better machine translation

(Simplified) *Vauquois triangle*:

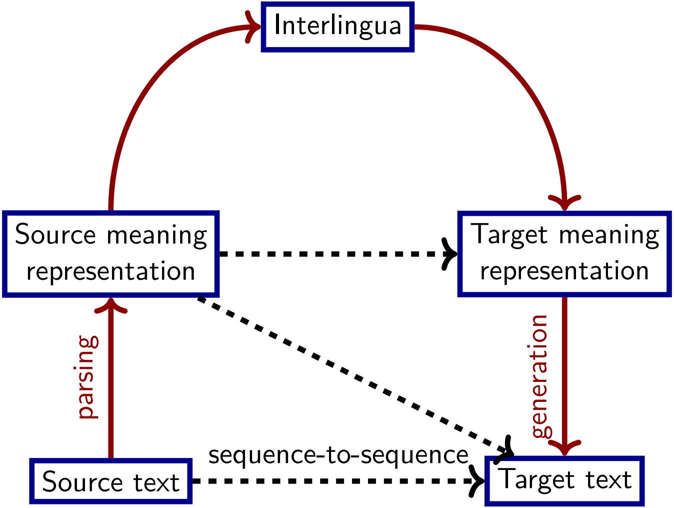


Better machine translation

(Simplified) *Vauquois triangle*:



(Simplified) *Vauquois triangle*:



Related courses

DIKU:

- Natural Language Processing
- Advanced Topics in Natural Language Processing
- Elements of Machine Learning
- Machine Learning
- Data Science

Linguistics:

- Semantics and pragmatics
- Language Processing 2
- Language 3 – Semantics, Interaction Analysis, and Linguistic Theory

(I am not teaching yet.)

Contact me...

... if you are interested in a project on

- Multilingual Enhanced Universal Dependency Parsing
- Meaning Representation Encoding for Machine Translation
- Semantic Dependency Probing of Pretrained Encoders
- Linguistic Analysis of Pretraining Methods
- Recursive Composition in Stack Pointer Parsers
- Phase Transitions in Word Representation
- Training Parsers with Translation Signals

`dh@di.ku.dk`