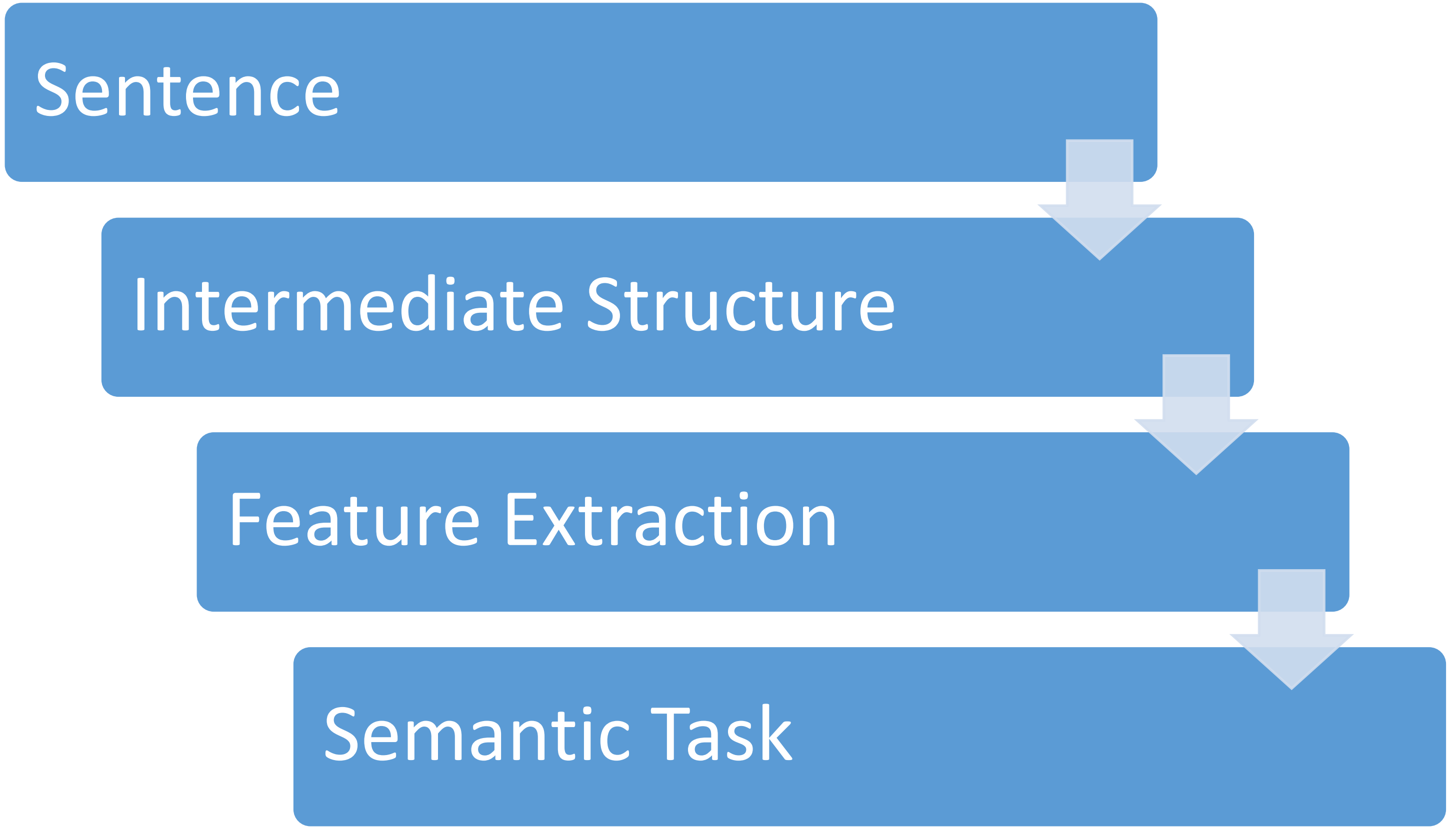
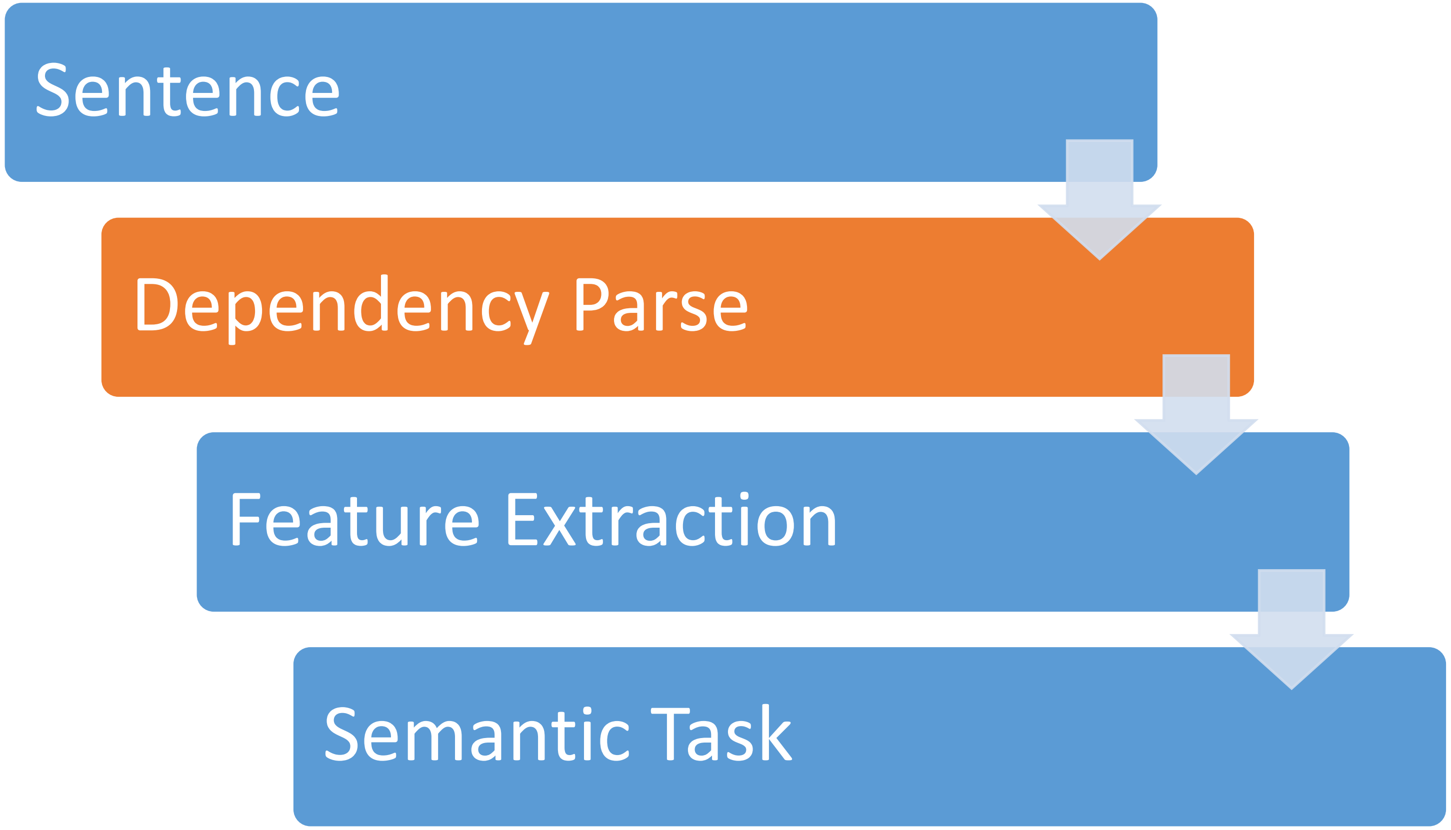
Open IE as an Intermediate Structure for Semantic Tasks

Gabriel Stanovsky, Ido Dagan and Mausam

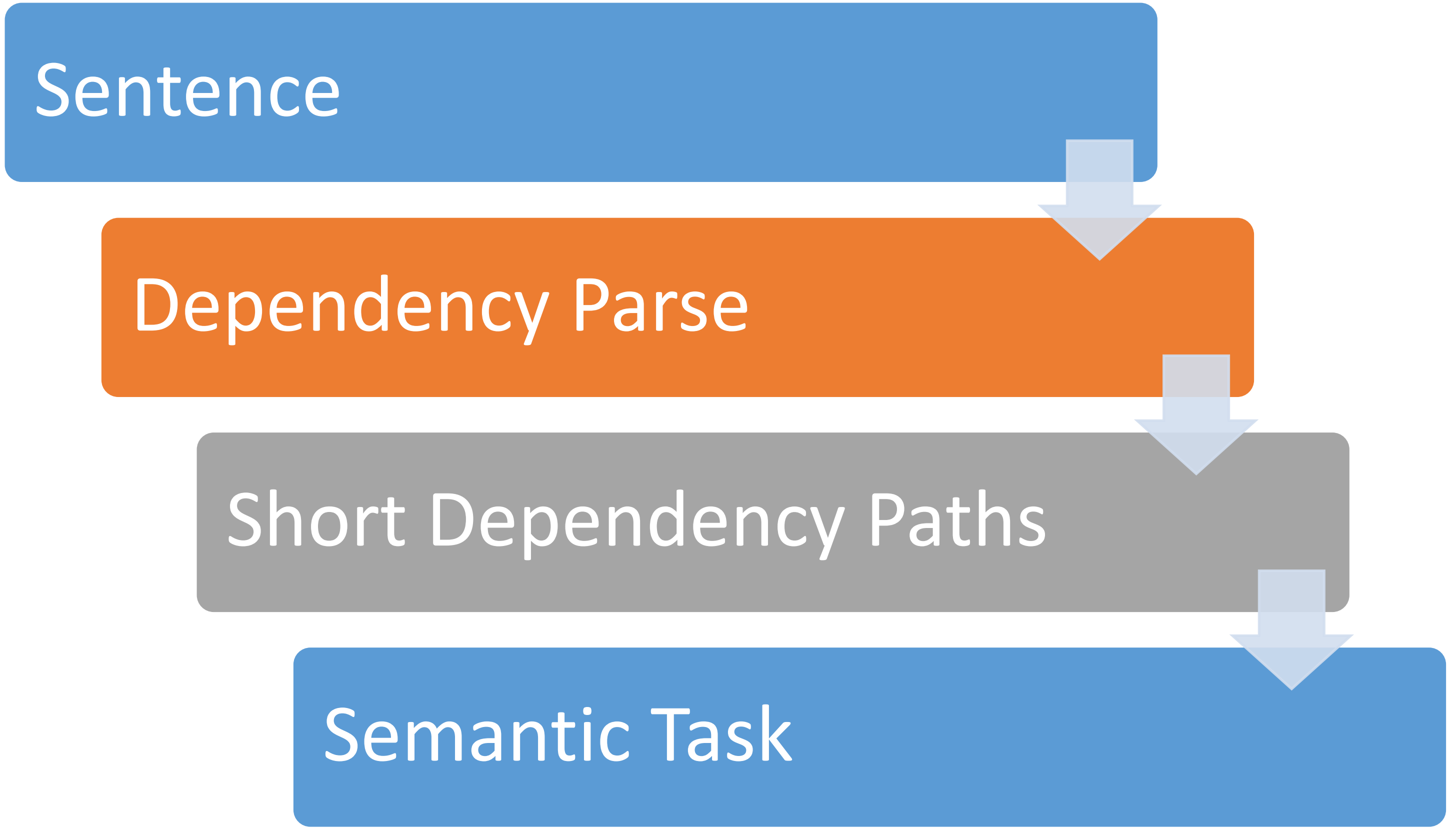
Sentence Level Semantic Application



Example: Sentence Compression



Example: Sentence Compression



Example: Sentence Compression

- Sentence
- Dependency Parse
- Short Dependency Paths
- Sentence Compression

Research Question

- **Open Information Extraction** was developed as an end-goal on itself
- ...Yet it makes **structural decisions**

*Can Open IE serve as a useful **intermediate representation**?*

Open Information Extraction

(John, **married**, Yoko)

(John, **wanted to leave**, the band)

(The Beatles, **broke up**)

Open Information Extraction

(John, **wanted to leave**, the band)

argument predicate argument

Open IE as Intermediate Representation

- Infinitives and multi word predicates

(John, **wanted to leave**, the band)

(The Beatles, **broke up**)

Open IE as Intermediate Representation

- Coordinative constructions

  "John decided to compose and perform solo albums"

(John, **decided to compose**, solo albums)

(John, **decided to perform**, solo albums)

Open IE as Intermediate Representation

- Appositions

  "Paul McCartney, founder of the Beatles, wasn't surprised"

(Paul McCartney, **wasn't surprised**)

(Paul McCartney, **[is] founder of**, the Beatles)

Open IE as Intermediate Representation

- Test Open IE versus:

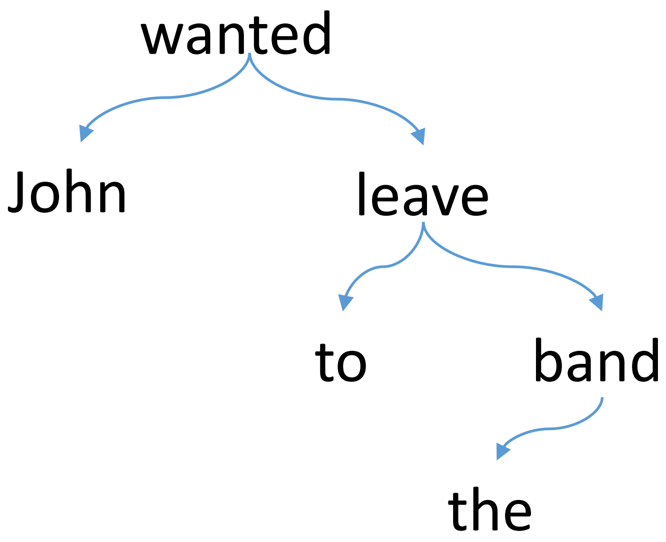
Open IE as Intermediate Representation

- Test Open IE versus:
  - Bag of words

John wanted to leave the band

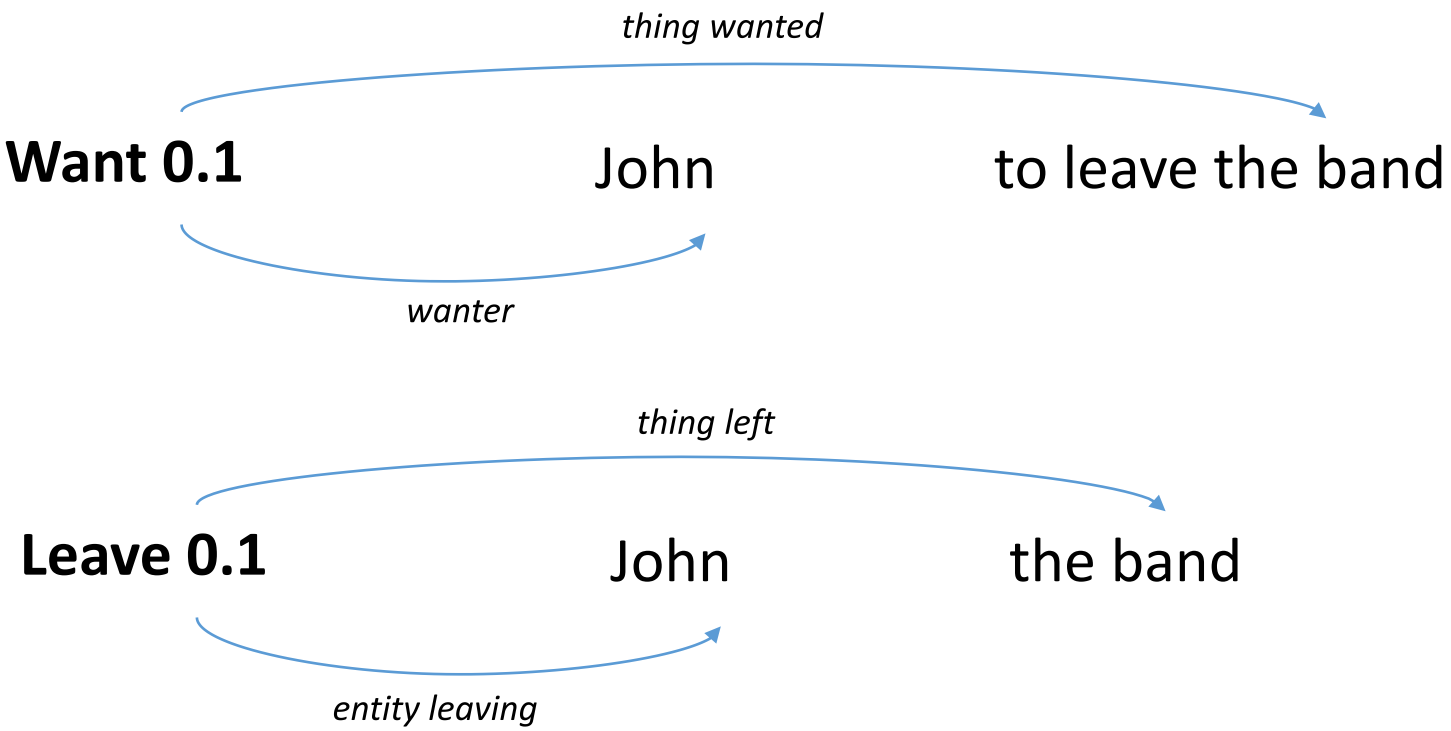
Open IE as Intermediate Representation

- Test Open IE versus:
  - Dependency parsing

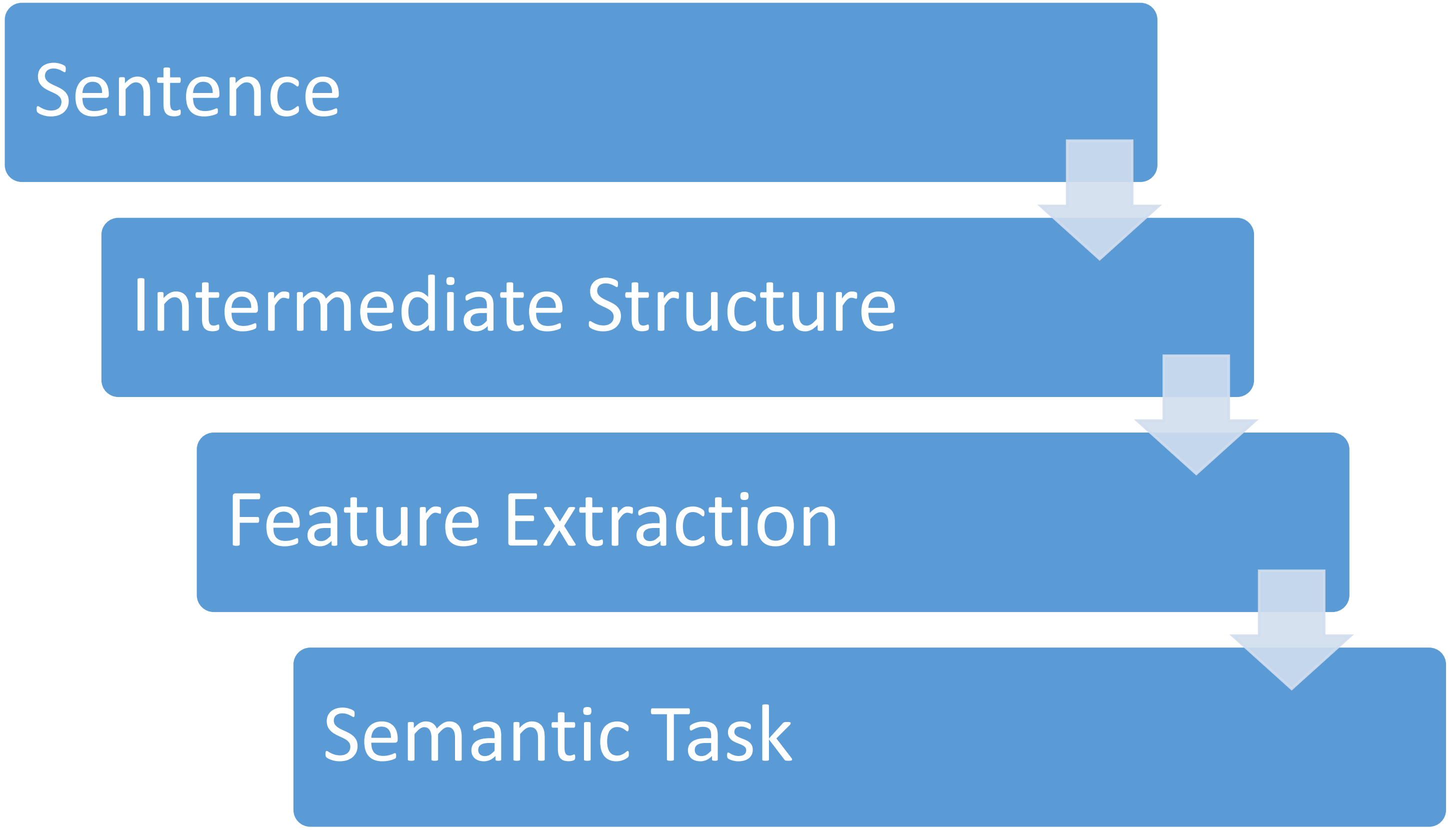


Open IE as Intermediate Representation

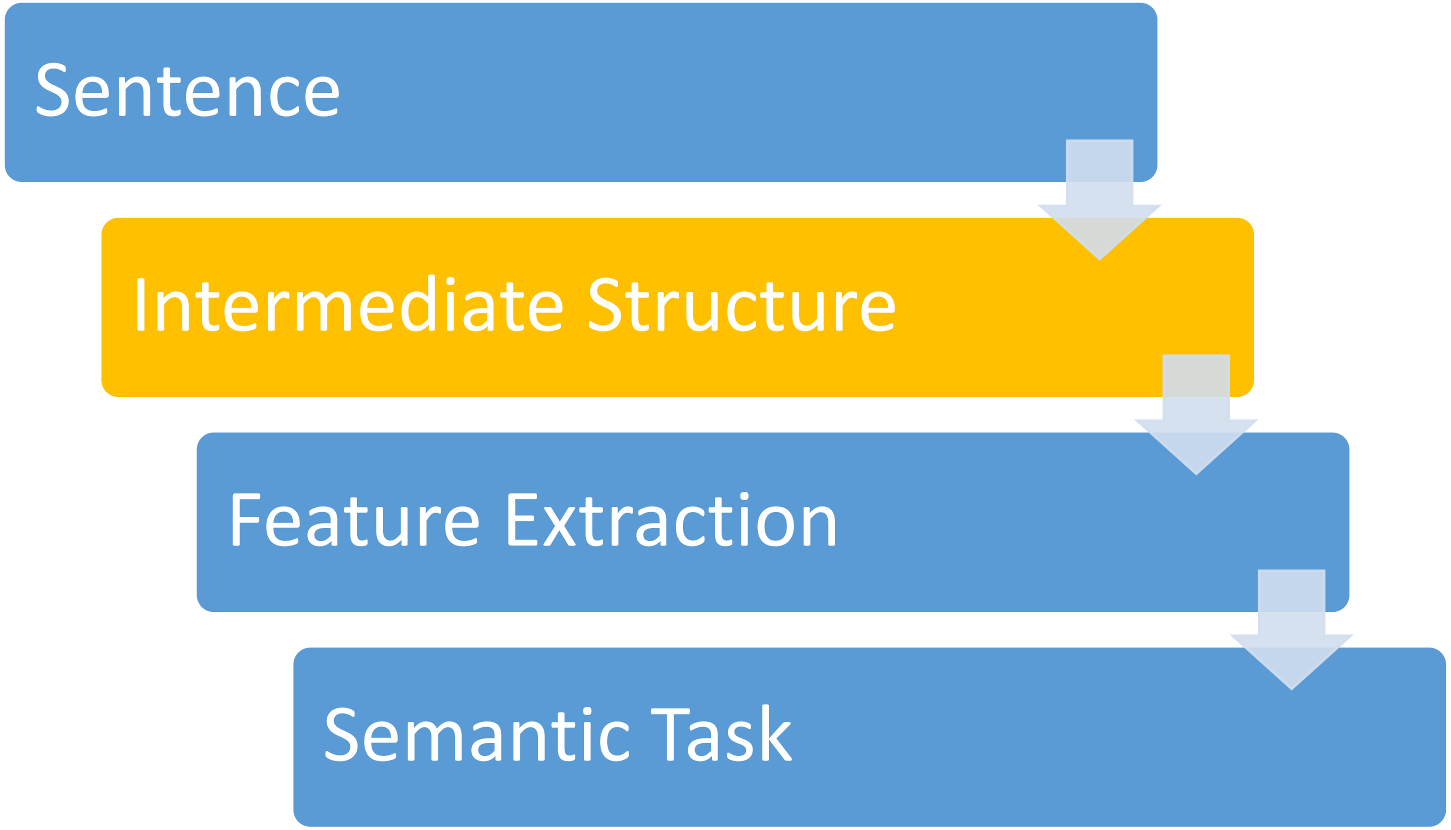
- Test Open IE versus:
  - Semantic Role Labeling



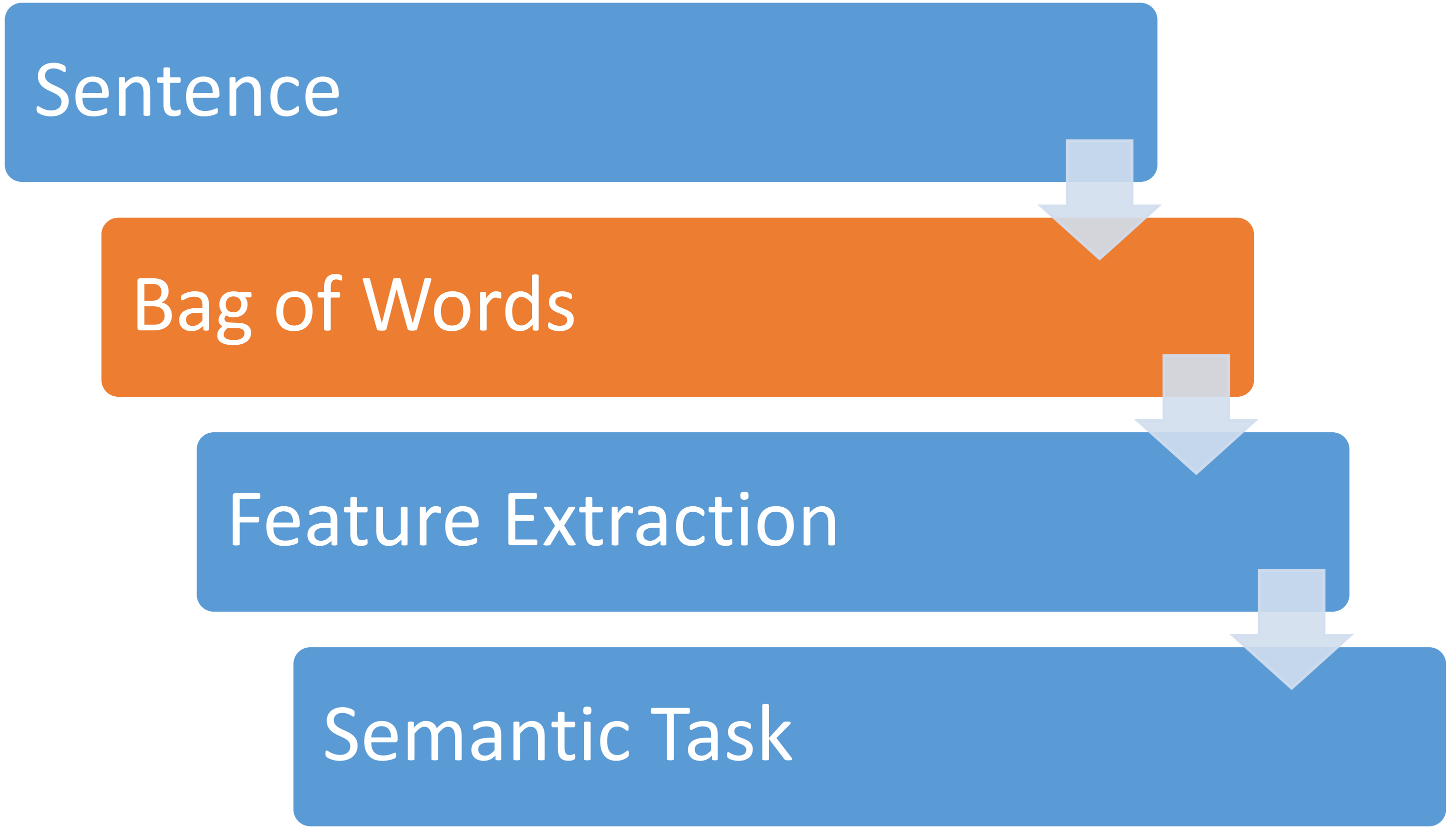
Quantitative Analysis



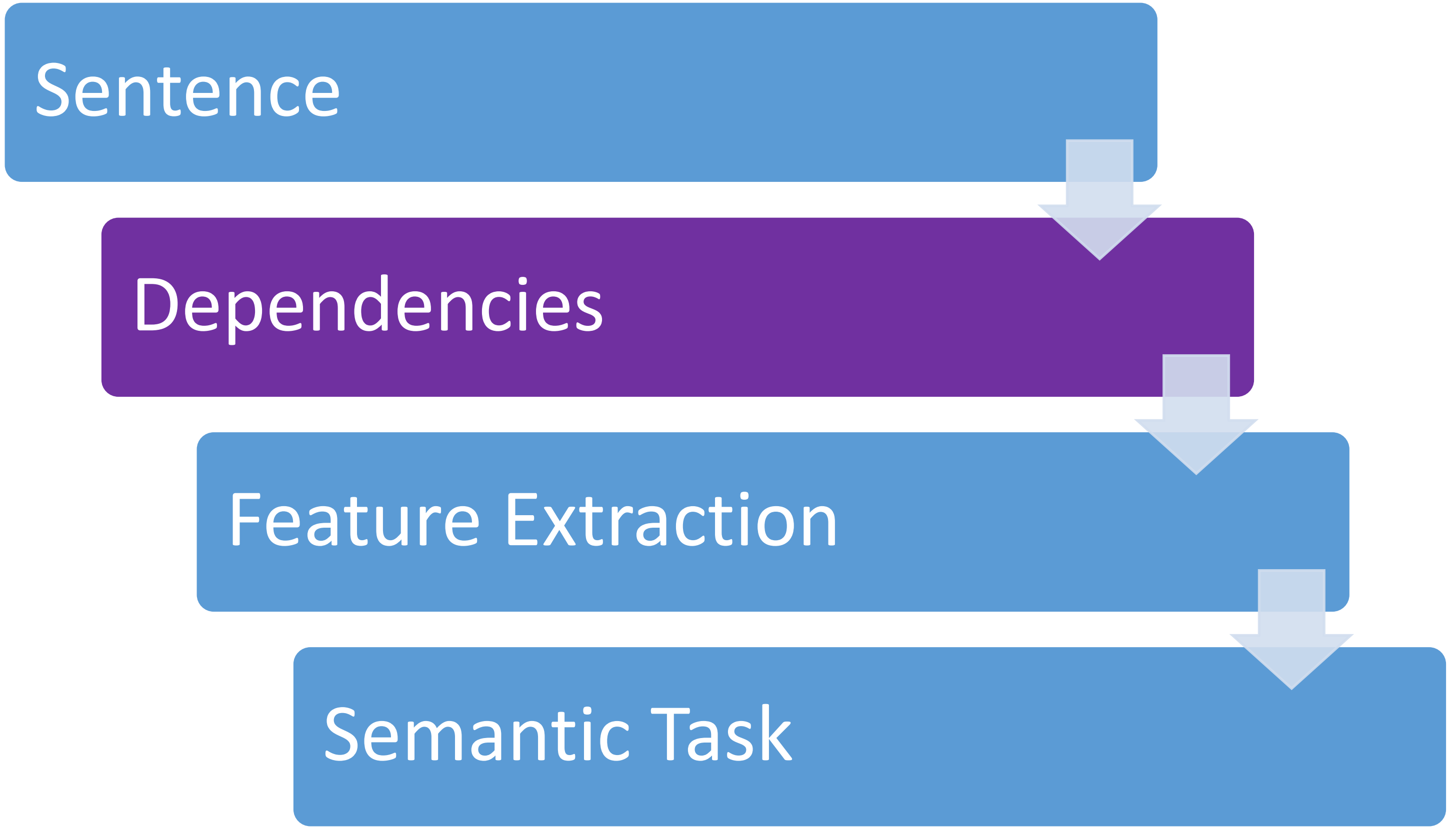
Quantitative Analysis



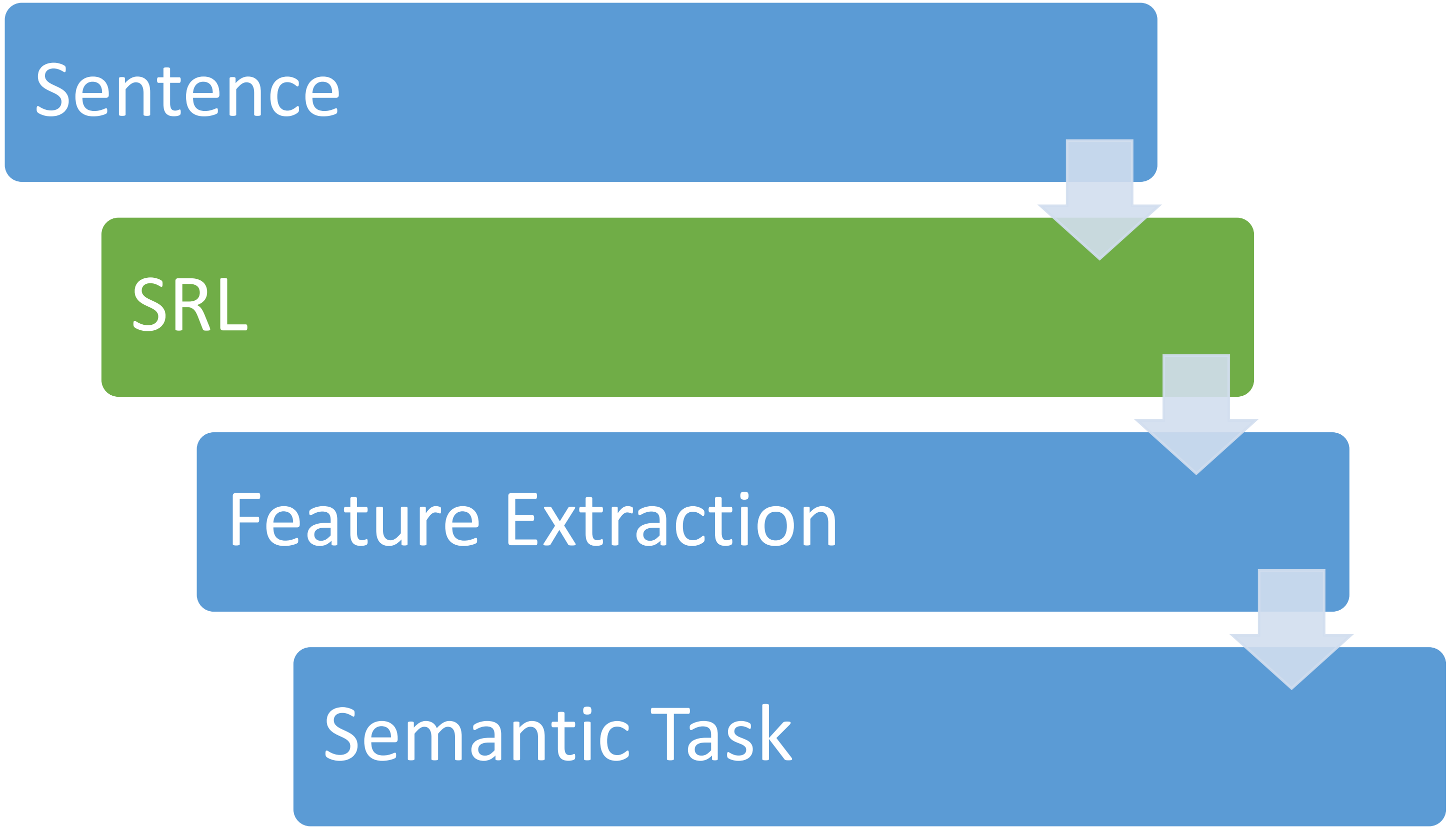
Quantitative Analysis



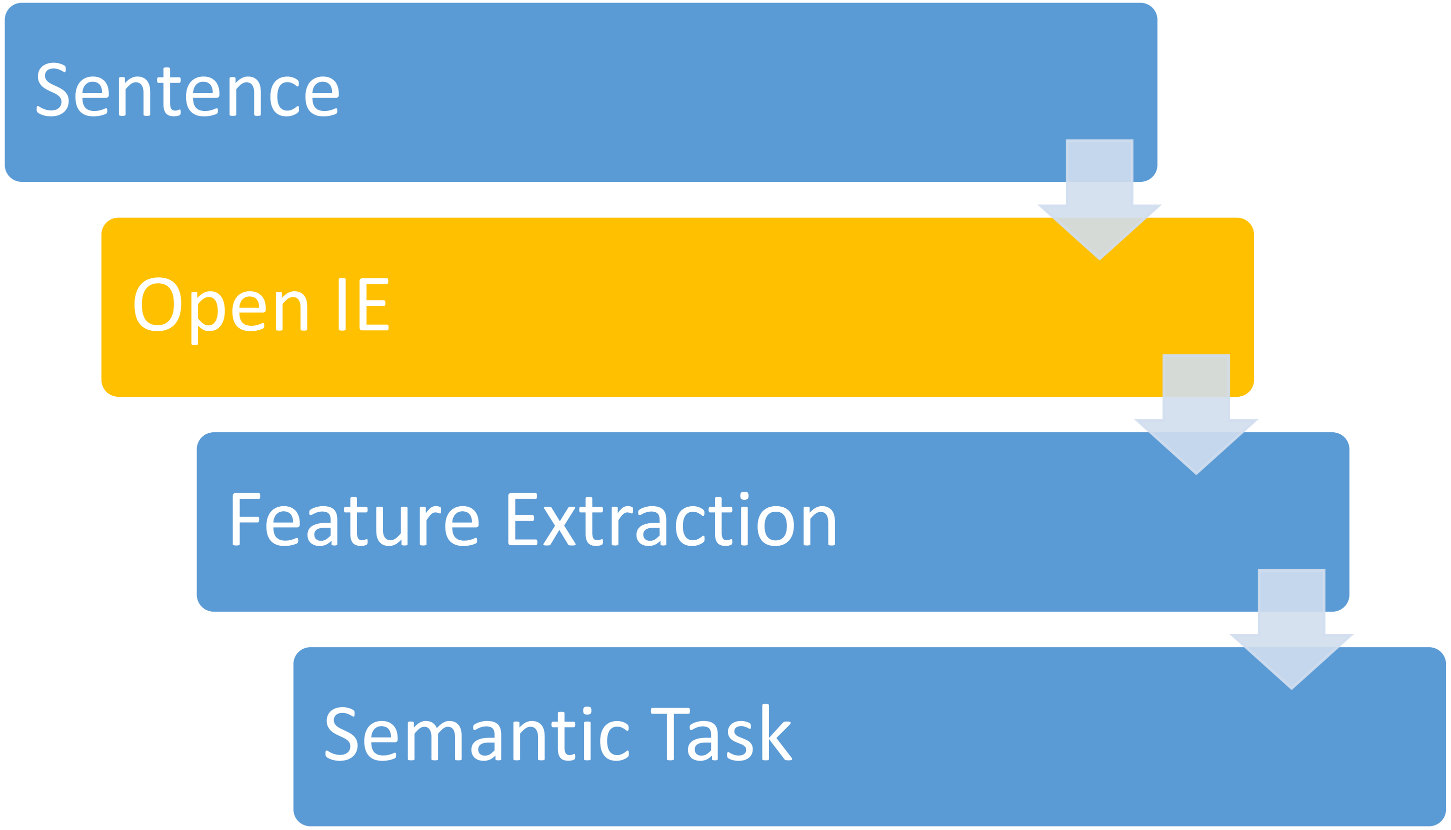
Quantitative Analysis



Quantitative Analysis



Quantitative Analysis



Textual Similarity

- Domain Similarity
  - *Carpenter ←→ hammer* **[Domain similarity]**
  - Various test sets:
    - Bruni (2012), Luong (2013), Radinsky (2011), and ws353 (Finkelstein et al., 2001)
    - ~5.5K instances

- Functional Simlarity
  - *Carpenter ←→ Shoemaker* **[Functional similarity]**
  - Dedicated test set:
    - Simlex999 (Hill et al, 2014)
    - ~1K instances

Word Analogies

- (**man** : **king**), (**woman** : **?**)

Word Analogies

- (**man** : **king**), (**woman** : **queen**)

Word Analogies

- (man : king), (woman : queen)
- (Athens : Greece), (Cairo : ?)

Word Analogies

- (**man** : **king**), (**woman** : **queen**)
- (**Athens** : **Greece**), (**Cairo** : **Egypt**)

Word Analogies

- (**man** : **king**), (**woman** : **queen**)
- (**Athens** : **Greece**), (**Cairo** : **Egypt**)
- Test sets:
  - Google (~195K instances)
  - MSR (~8K instances)

Reading Comprehension

- MCTest, (Richardson et. al., 2013)

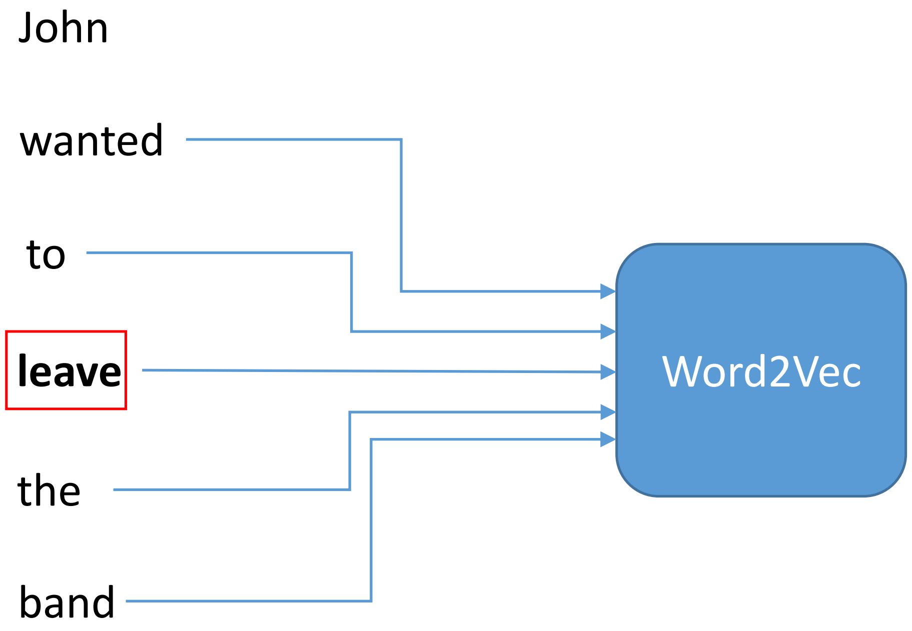
- Details in the paper!

Textual Similarity and Analogies

- Previous approaches used distance metrics over word embedding:
  - (Mikolov et al, 2013) - **lexical contexts**
  - (Levy and Goldberg, 2014) - **syntactic contexts**
- We compute embeddings for **Open IE** and **SRL** contexts
- Using the same training data for all embeddings (1.5B tokens Wikipedia dump)

Computing Embeddings

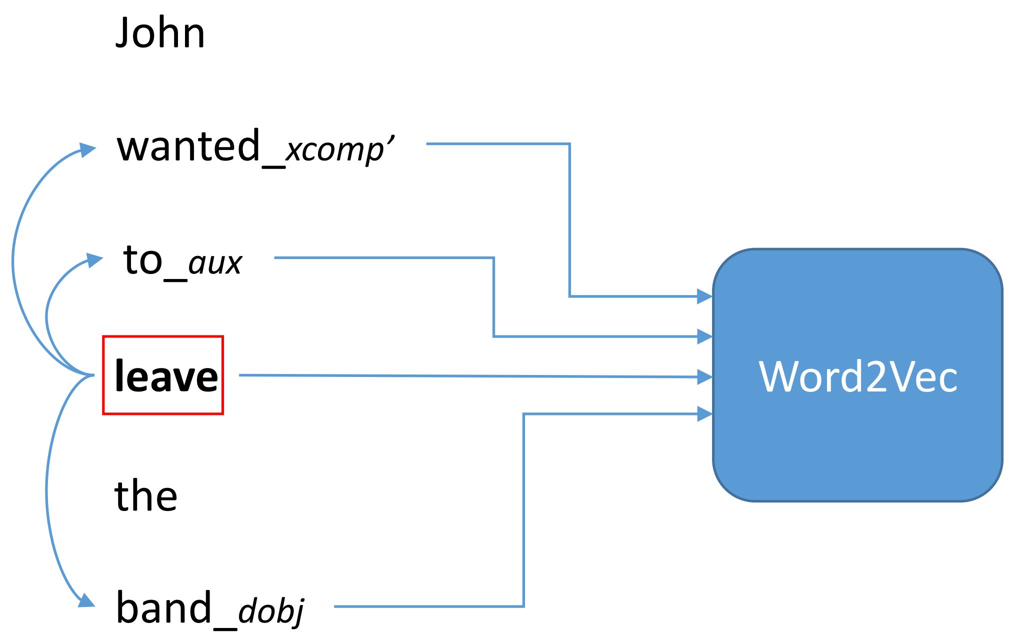
- **Lexical contexts**
  (for word **leave**)



(Mikolov et al., 2013)

Computing Embeddings

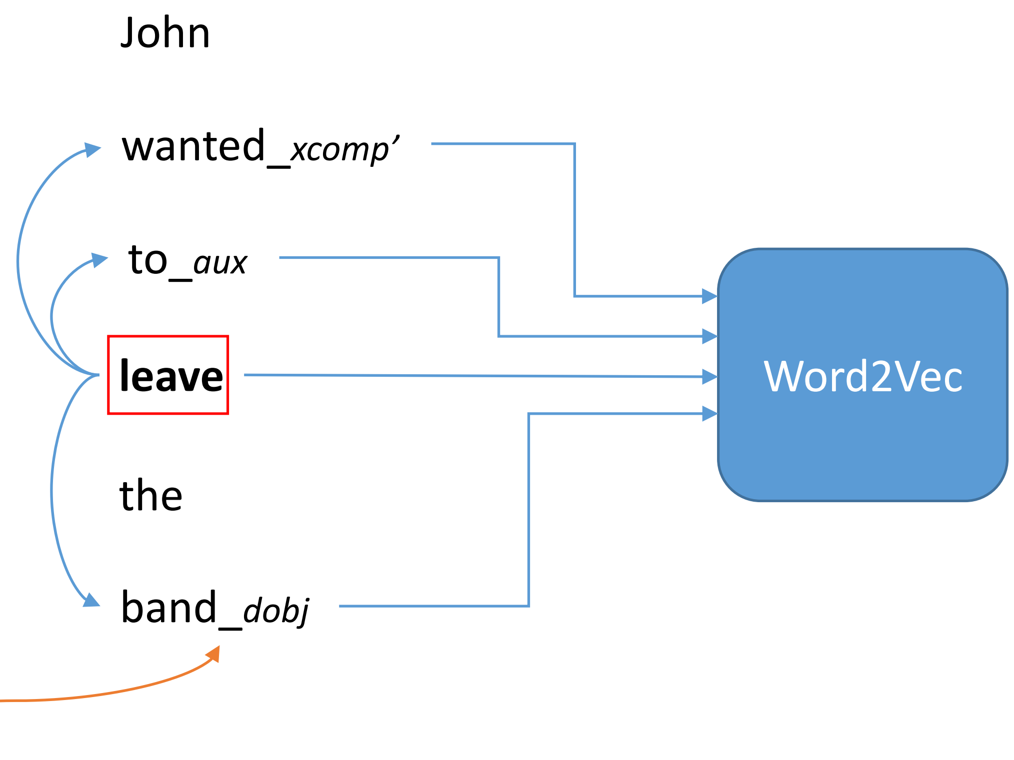
- **Syntactic contexts**
  (for word **leave**)



(Levy and Goldberg, 2014)

Computing Embeddings

- **Syntactic contexts**
  (for word **leave**)

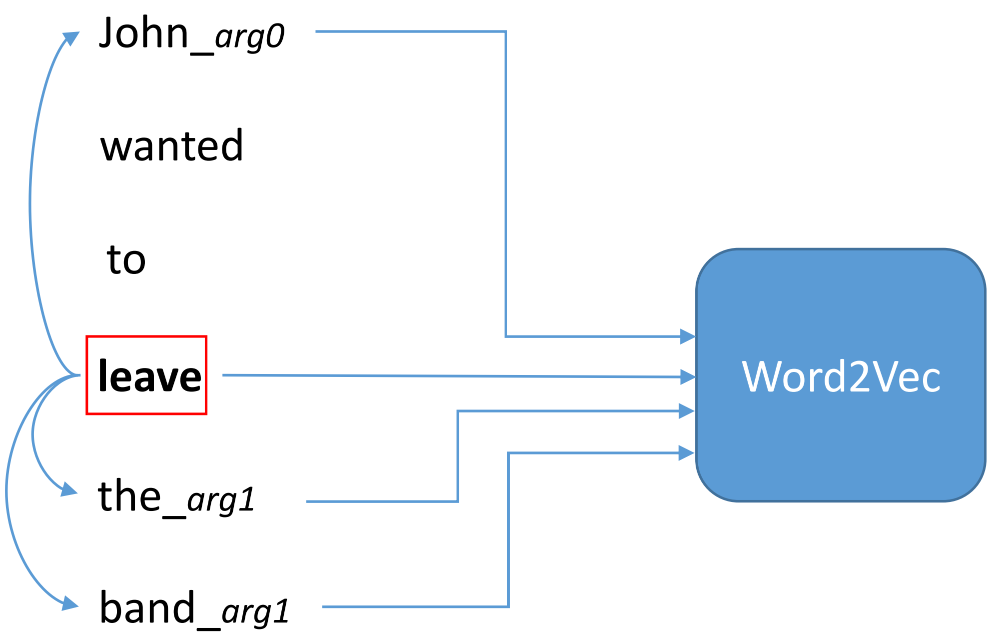


A context is formed of word + syntactic relation

(Levy and Goldberg, 2014)

Computing Embeddings

- **SRL contexts**
  (for word **leave**)

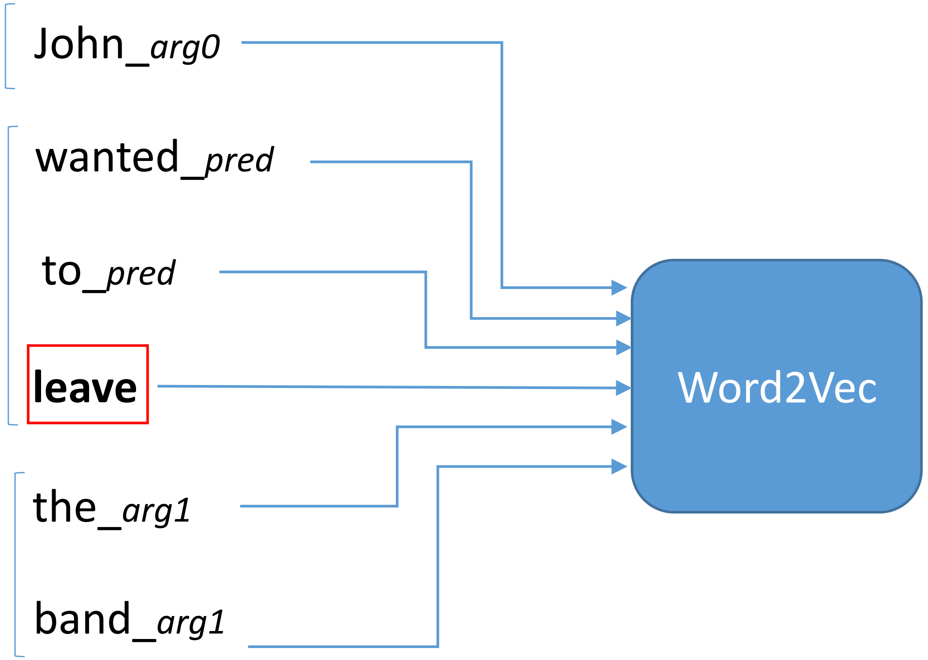


Available at author's website

Computing Embeddings

- **Open IE contexts**
  (for word **leave**)

(John, **wanted to leave**, the band)



Available at author's website

Results on Textual Similarity

	Open IE	Lexical	Deps	SRL
bruni	**.757**	.735	.618	.491
luong	**.288**	.229	.197	.171
radinsky	**.681**	.674	.592	.433
simlex	.39	.365	**.447**	.306
ws353-rel	**.647**	.64	.492	.551
ws353-sym	**.77**	.763	.759	.439
ws353-full	**.711**	.703	.629	.693

Results on Textual Similarity

	Open IE	Lexical	Deps	SRL
bruni	**.757**	.735	.618	.491
luong	**.288**	.229	.197	.171
radinsky	**.681**	.674	.592	.433
simlex	.39	.365	**.447**	.306
ws353-rel	**.647**	.64	.492	.551
ws353-sym	**.77**	.763	.759	.439
ws353-full	**.711**	.703	.629	.693

Syntactic does better on functional similarity

Results on Analogies

	Google		MSR	
	Add	Mul	Add	Mul
Open IE	**.714**	**.719**	**.529**	**.55**
Lexical	.651	.656	.438	.455
Deps	.34	.367	.4	.434
SRL	.352	.362	.389	.406

Additive

$$\arg\max_{b^* \in V} \left(\cos\left(b^*, b\right) - \cos\left(b^*, a\right) + \cos\left(b^*, a^*\right)\right)$$

Multiplicative

$$\arg\max_{b^* \in V} \frac{\cos\left(b^*, b\right)\cos\left(b^*, a^*\right)}{\cos\left(b^*, a\right) + \varepsilon}$$

Results on Analogies

	Google		MSR	
	Add	Mul	Add	Mul
Open IE	**.714**	**.719**	**.529**	**.55**
Lexical	.651	.656	.438	.455
Deps	.34	.367	.4	.434
SRL	.352	.362	.389	.406

State of the art with this amount of data

Additive

$$\arg\max_{b^* \in V} \left(\cos(b^*, b) - \cos(b^*, a) + \cos(b^*, a^*)\right)$$

Multiplicative

$$\arg\max_{b^* \in V} \frac{\cos(b^*, b)\cos(b^*, a^*)}{\cos(b^*, a) + \varepsilon}$$

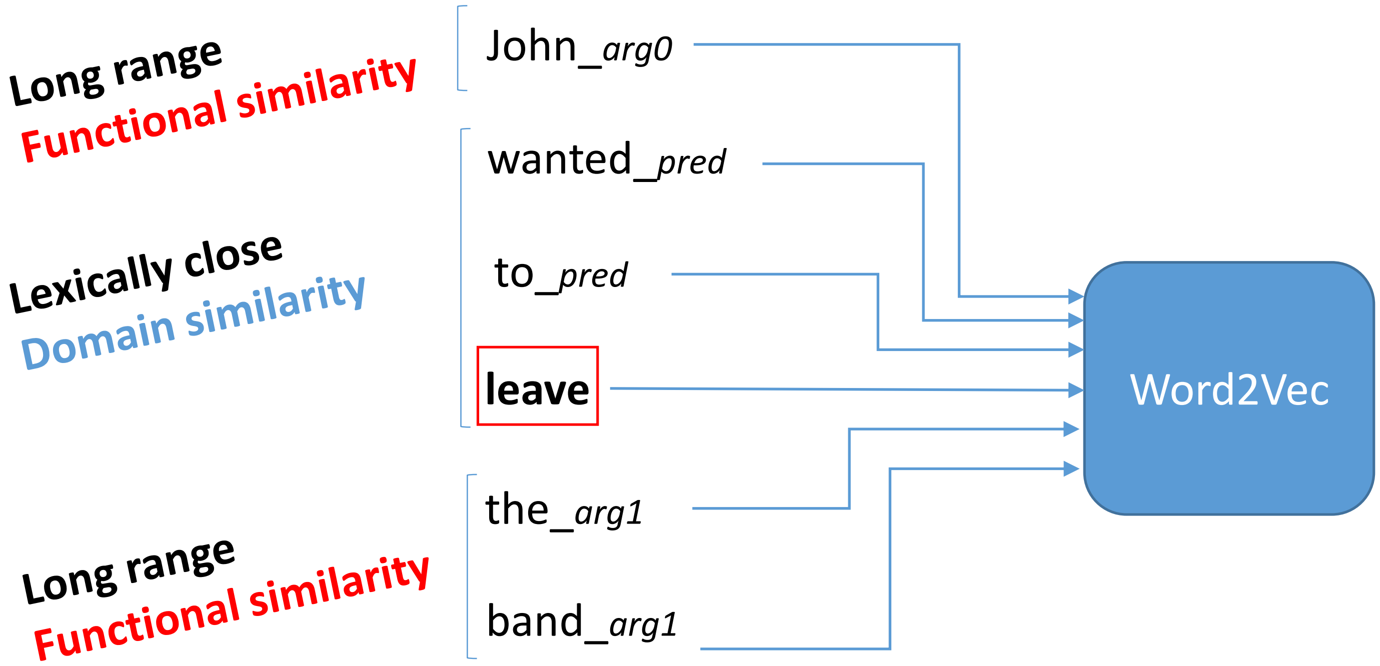
Domain vs. Functional Similarity

- Previous work has identified that:
  - Lexical contexts induce domain similarity
  - Syntactic contexts induce functional similarity

- ***What kind of similarity does Open IE induce?***

Computing Embeddings

- **Open IE contexts**
  (for word **leave**)



Open IE combines domain and functional similarity in a single framework!

Concluding Example

- (***gentlest***: gentler), (***loudest***:?)
  - Lexical: **higher-pitched** **X** **[Domain Similar]**
  - Syntactic: **thinnest** **X** **[Functionally Similar]**
  - SRL: **unbelievable** **X** **[Functionally Similar?]**
  - Open-IE: louder **V**

Conclusions

- Open IE makes different structural decisions
  - These can prove beneficial in certain tasks
- A key strength is Open IE's ability to balance lexical proximity with long range dependencies in a single representation
- Embeddings made available: www.cs.bgu.ac.il/~gabriels

Thank you!

Questions?