Semantic parsing with graph transformations

Ádám Kovács
Kinga Gémes
Gábor Recski

adaam.ko@gmail.com, kinga.andrea.gemes@gmail.com, recski@apollo.ai

AACS19
21.06.2019.

Our contribution

Our contribution

- Defining graph based semantic models

Our contribution

- Defining graph based semantic models
- Evaluation of our models on a machine comprehension task

Our contribution

- Defining graph based semantic models
- Evaluation of our models on a machine comprehension task
- Integrating our model to a state-of-the-art system

Our contribution

- Defining graph based semantic models
- Evaluation of our models on a machine comprehension task
- Integrating our model to a state-of-the-art system
- Results show some improvement on the task

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

Text: "Today we decided to paint the extra room in our house. Were going to have visitor coming next month so hopefully the painting ain't that smelly anymore. I made sure that the wall is clean and clear of all the nuisance. We already bought the pain and we decided the new wall pain is sky blue. My husband is putting newspaper on the floor to avoid any spill on our floor/carpet....."
Question: "Did anyone help him?"
1st answer: "It was the narrator and her husband"
2nd answer: "No, he worked alone."

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

- **2018-as Semeval Task**
  - *Machine comprehension using commonsense knowledge*

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

- **2018-as Semeval Task**
  - *Machine comprehension using commonsense knowledge*
- Simple multi choice questions based on short stories

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

- **2018-as Semeval Task**
  - *Machine comprehension using commonsense knowledge*
- Simple multi choice questions based on short stories

- The best two systems
  - `HFL-RC`
  - `Yuanfudao`

The machine comprehension task (Chen et al., 2018; Wang et al., 2018)

- **2018-as Semeval Task**
  - *Machine comprehension using commonsense knowledge*
- Simple multi choice questions based on short stories

- The best two systems
  - `HFL-RC`
  - `Yuanfudao`
- **MCScript**
  - training and test data

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

- type 0 edge

`4lang`: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

- type 0 edge
  - property: `dog` $\xrightarrow{0}$ `large`

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

- type 0 edge
  - property: dog $\xrightarrow{0}$ large
  - IS_A relation (hypernym): dog $\xrightarrow{0}$ mammal

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

- type 0 edge
  - property: $\texttt{dog} \xrightarrow{0} \texttt{large}$
  - IS_A relation (hypernym): $\texttt{dog} \xrightarrow{0} \texttt{mammal}$
  - predicate: $\texttt{dog} \xrightarrow{0} \texttt{bark}$

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

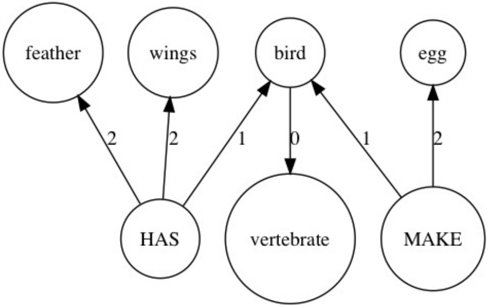
Directed meaning graphs with 3 edge type

- type 0 edge
    - property: $\texttt{dog} \xrightarrow{0} \texttt{large}$
    - IS_A relation (hypernym): $\texttt{dog} \xrightarrow{0} \texttt{mammal}$
    - predicate: $\texttt{dog} \xrightarrow{0} \texttt{bark}$
- type 1 and type 2 edges connect a binary predicate with the arguments.

4lang: the formalism (Kornai, 2010; Kornai et al., 2015)

Directed meaning graphs with 3 edge type

- type 0 edge
  - property: dog $\xrightarrow{0}$ large
  - IS_A relation (hypernym): dog $\xrightarrow{0}$ mammal
  - predicate: dog $\xrightarrow{0}$ bark
- type 1 and type 2 edges connect a binary predicate with the arguments.



The `4lang` service

The 4lang service

- Live service built on the `text_to_4lang` module

The `4lang` service

- Live service built on the `text_to_4lang` module
- Highly automated graph generation

The 4lang service

- Live service built on the `text_to_4lang` module
- Highly automated graph generation
- Online demo
  - `http://4lang.hlt.bme.hu`

The 4lang service

- Live service built on the text_to_4lang module
- Highly automated graph generation
- Online demo
  - http://4lang.hlt.bme.hu
- Github
  - https://github.com/adaamko/4lang

The 4lang service

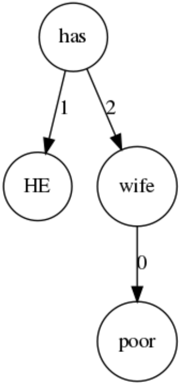
The `4lang` service

My poor wife!

The 4lang service

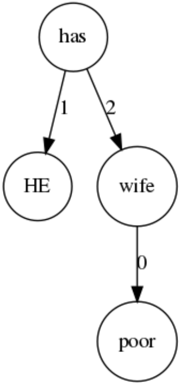
I feel bad for my wife!

My poor wife!

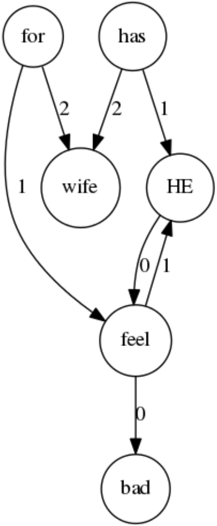


The 4lang service

My poor wife!



I feel bad for my wife!



Defined metrics

Defined metrics

- **Goal**
  - is the statement a consequence of the hypothesis?

Defined metrics

- **Goal**
  - is the statement a consequence of the hypothesis?
- **Implementation**
  - Graph similarity
  - What makes two graphs similar?

Defined metrics

- **Goal**
  - is the statement a consequence of the hypothesis?
- **Implementation**
  - Graph similarity
  - What makes two graphs similar?

- My poor wife! (G1)

Defined metrics

- **Goal**
  - is the statement a consequence of the hypothesis?
- **Implementation**
  - Graph similarity
  - What makes two graphs similar?

- My poor wife! (G1)
- I feel bad for my wife! (G2)

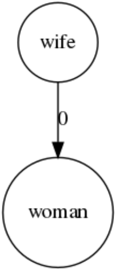
Defined metrics

- **Goal**
  - is the statement a consequence of the hypothesis?
- **Implementation**
  - Graph similarity
  - What makes two graphs similar?

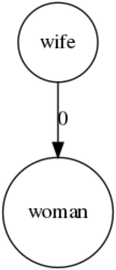
- My poor wife! (G1)
- I feel bad for my wife! (G2)
- $$\frac{|E(G_1) \cap E(G_2)|}{|E(G_2)|}$$

Expanded graphs, the "expand" function

Expanded graphs, the "expand" function

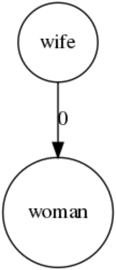


Expanded graphs, the "expand" function

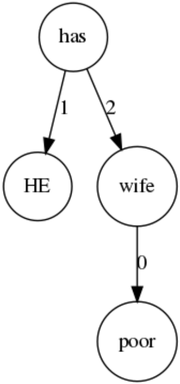


$\Rightarrow$

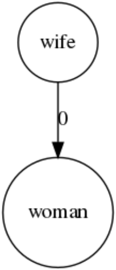
Expanded graphs, the "expand" function



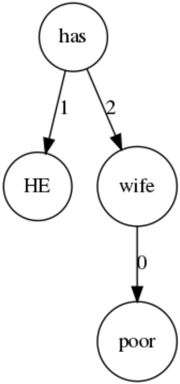
$\Rightarrow$



Expanded graphs, the "expand" function

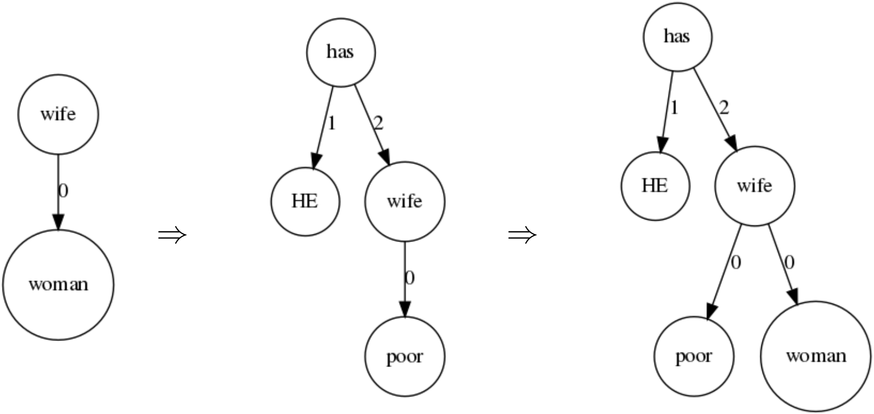


$\Rightarrow$



$\Rightarrow$

Expanded graphs, the "expand" function

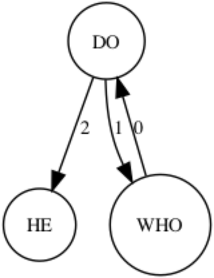


Baseline

- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score

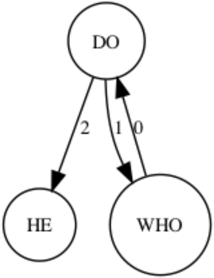
Baseline

- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score



Baseline

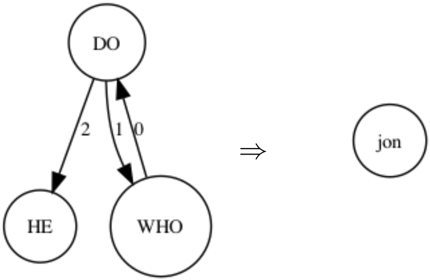
- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score



$\Rightarrow$

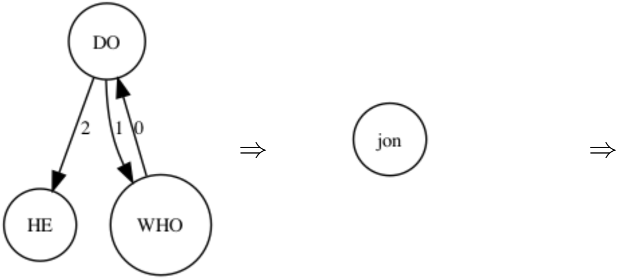
Baseline

- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score



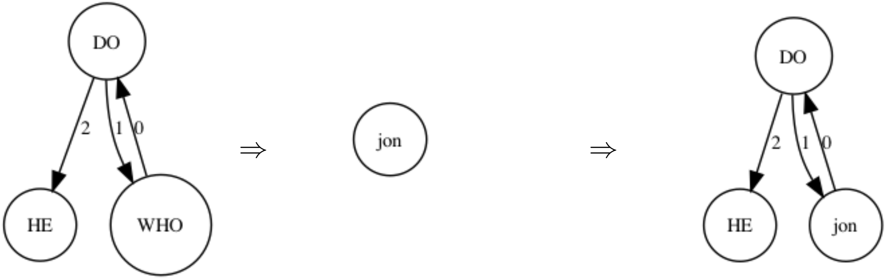
Baseline

- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score



Baseline

- Merged graph for all question-answer pair
- The "more" similar is the correct answer
- **68,3** accuracy score



Deep learning

Deep learning

- Widely used method

Deep learning

- Widely used method
- State-of-the-art in Natural Language Processing

Deep learning

- Widely used method
- State-of-the-art in Natural Language Processing
- RNN handles sequential data the best

Deep learning

- Widely used method
- State-of-the-art in Natural Language Processing
- RNN handles sequential data the best
    - LSTM

Deep learning

- Widely used method
- State-of-the-art in Natural Language Processing
- RNN handles sequential data the best
  - LSTM
  - Attention

Yuanfudao system

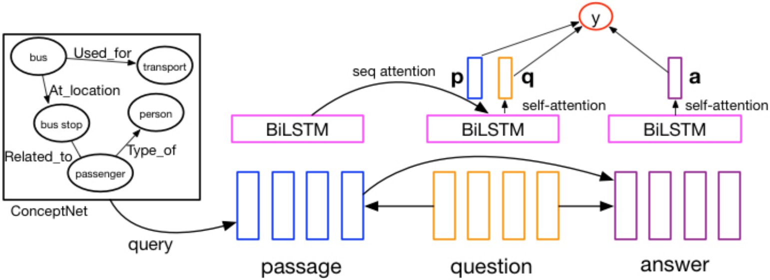


Figure: The structure of the system

Modifications

Modifications

- `4lang` similarity score calculated between words during the pre-processing

Modifications

- `4lang` similarity score calculated between words during the pre-processing
- New embedding layer for the `4lang` similarities

Modifications

- `4lang` similarity score calculated between words during the pre-processing
- New embedding layer for the `4lang` similarities
- Expansion of the RNN layer, that the corresponding RNN gets the `4lang` embedding's output

Results

model	dev	test
TriAN, no ConceptNet	82.8%	80.2%
TriAN, with ConceptNet	82.7%	80.5%
TriAN, with 4lang	**83.2%**	**80.9%**
TriAN, with both	83.1%	80.8%

Table: Effect of `4lang` and `ConceptNet` on results

Results

model	dev	test
TriAN, no ConceptNet	83.7%	81.9%
TriAN, with ConceptNet	82.5%	80.3%
TriAN, with 4lang	84.2%	81.5%
TriAN, with both	**83.4%**	**82.9%**

Table: Effect of `4lang` and `ConceptNet` on the pretrained models

Results

pretrained, ensembled model	test
TriAN, no ConceptNet	82.95%
TriAN, with ConceptNet	83.697%
TriAN, with 4lang	82.8%
TriAN, with both	**83.73%**

Table: The effect of `4lang` and `ConceptNet` on the pretrained and ensembled models

Summary

Summary

- Automated method and metric for measuring similarity.

Summary

- Automated method and metric for measuring similarity.
- Strong baseline on the machine comprehension task

Summary

- Automated method and metric for measuring similarity.
- Strong baseline on the machine comprehension task
- Using out baseline method in the state-of-the-art system, seeing some improvement

Summary

- Automated method and metric for measuring similarity.
- Strong baseline on the machine comprehension task
- Using out baseline method in the state-of-the-art system, seeing some improvement
- A code is available at
  `https://github.com/adaamko/4lang`
  `https://github.com/GKingA/commonsense-rc`

Summary

- Automated method and metric for measuring similarity.
- Strong baseline on the machine comprehension task
- Using out baseline method in the state-of-the-art system, seeing some improvement
- A code is available at
  `https://github.com/adaamko/4lang`
  `https://github.com/GKingA/commonsense-rc`
- Live Rest API and online demo
  `http://4lang.hlt.bme.hu`

Thank you for your attention!

Chen, Zhipeng et al. (2018). "HFL-RC System at SemEval-2018 Task 11: Hybrid Multi-Aspects Model for Commonsense Reading Comprehension". In: *arXiv preprint arXiv:1803.05655*.

Kornai, András (2010). "The algebra of lexical semantics". In: *Proceedings of the 11th Mathematics of Language Workshop*. Ed. by Christian Ebert, Gerhard Jäger, and Jens Michaelis. LNAI 6149. Springer, pp. 174–199.

Kornai, András et al. (2015). "Competence in lexical semantics". In: *Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics (*SEM 2015)*. Denver, Colorado: Association for Computational Linguistics, pp. 165–175.

Wang, Liang et al. (2018). "Yuanfudao at SemEval-2018 Task 11: Three-way Attention and Relational Knowledge for Commonsense Machine Comprehension". In: *arXiv preprint arXiv:1803.00191*.