NL2SPARQL

by Nico Lutz

A journey

X

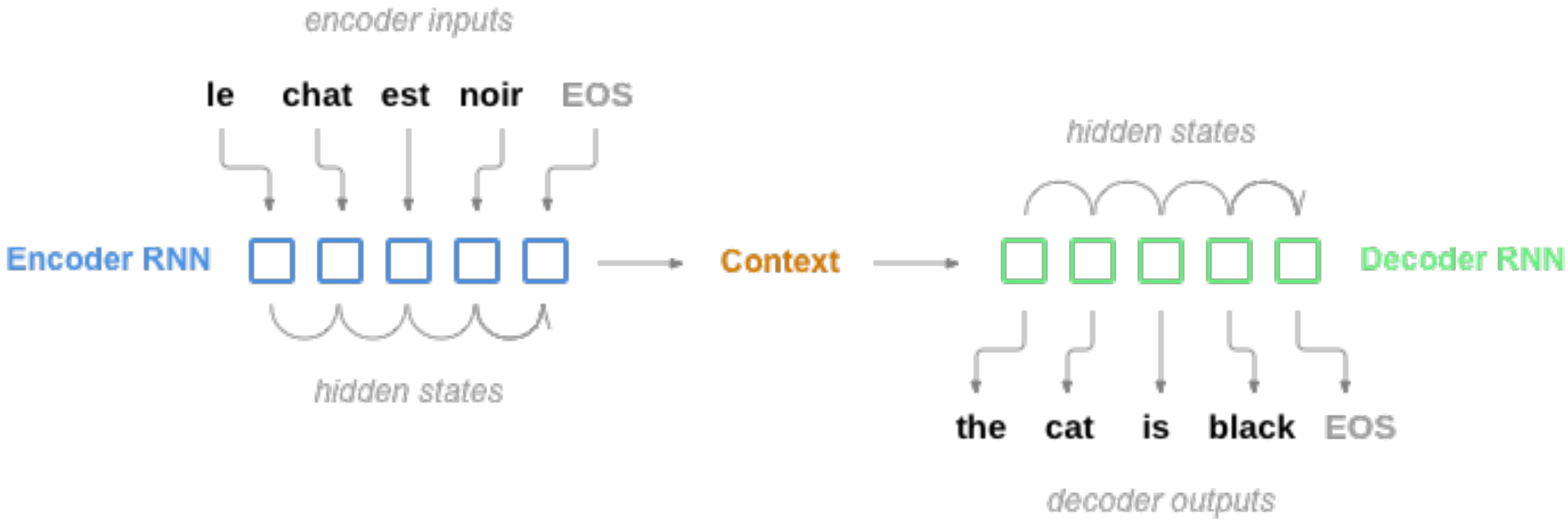
Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base

Natural Language :
Which comic characters are painted by Bill Finger ?

```
SELECT DISTINCT ?uri
WHERE {
?uri <http://dbpedia.org/ontology/creator> <http://
dbpedia.org/resource/Bill_Finger>  .
?uri <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://dbpedia.org/ontology/ComicsCharacter>
}
```

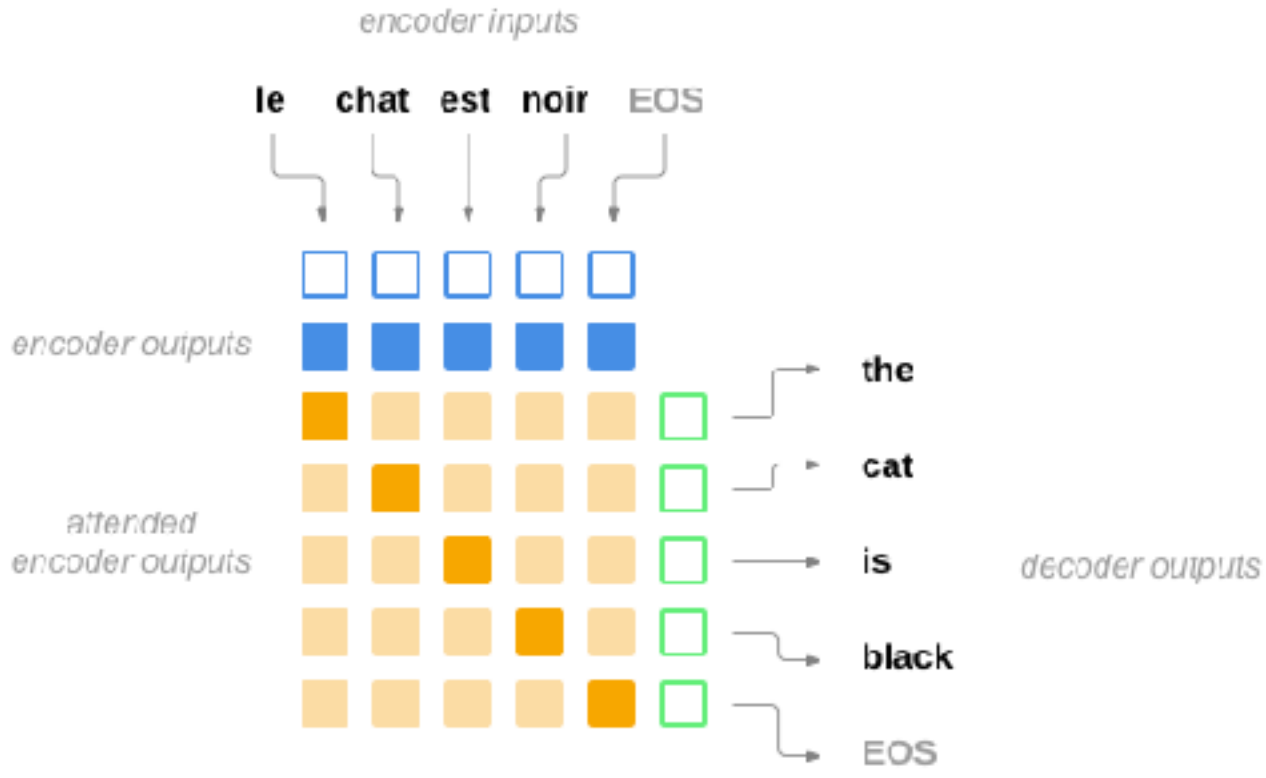
Approach: NMT

Encoder Decoder


encoder inputs
le
chat
est
noir
EOS
Encoder RNN
hidden states
Context
hidden states
Decoder RNN
the
cat
is
black
EOS
decoder outputs

+

Attention



encoder inputs

le chat est noir EOS

encoder outputs

attended encoder outputs

the

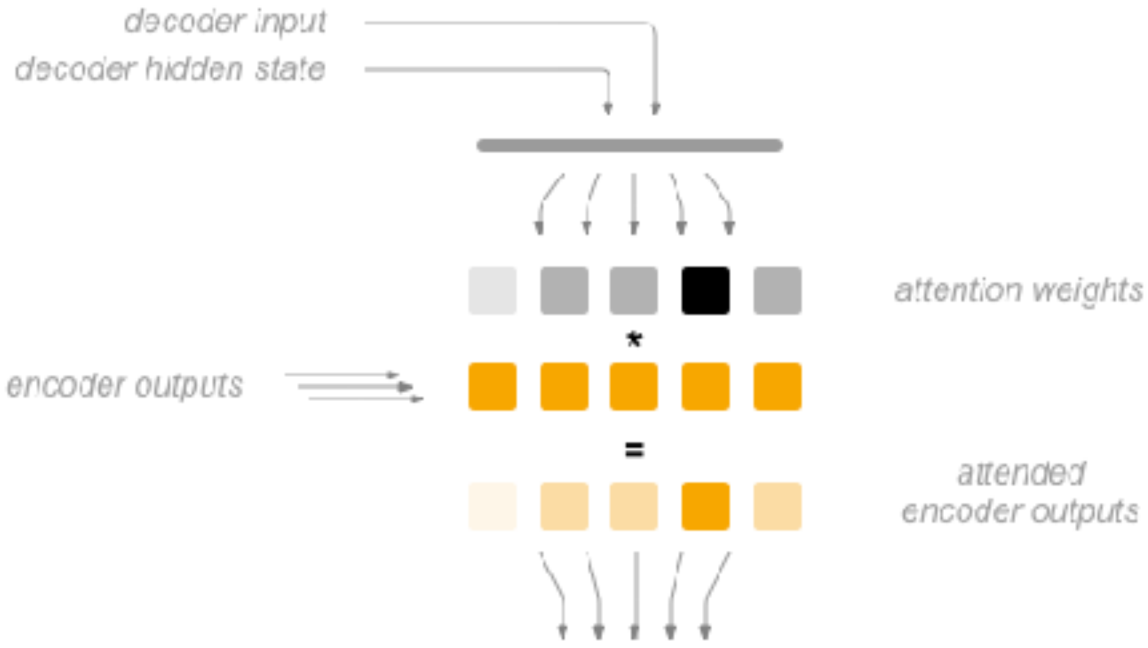
cat

is

black

EOS

decoder outputs



decoder input

decoder hidden state

attention weights

encoder outputs

*

=

attended encoder outputs

Preprocessing

- Lower case
- Adding special tokens:
    - '<curly_open> ': '{'
    - ' <curly_close>': '}'
    - '<croc_open> ':'<'
    - ' <croc_close>’:'>'
    - ' <punc>': '.'
    - ' <ques>':' ?’
    - EOS
    - SOS
    - PAD
- Padding
- Eval / Train split

which comic characters are painted by bill finger <ques>

```
SELECT DISTINCT  <ques>uri
WHERE <curly_open>
<ques>uri <croc_open> http://dbpedia.org/ontology/creator <croc_close>
<croc_open> http://dbpedia.org/resource/bill_finger <croc_close>  <punc>

<ques>uri <croc_open> http://www.w3.org/1999/02/22-rdf-syntax-ns#type
<croc_close> <croc_open> http://dbpedia.org/ontology/comicscharacter <croc_close>

<curly_close>
```

English - German

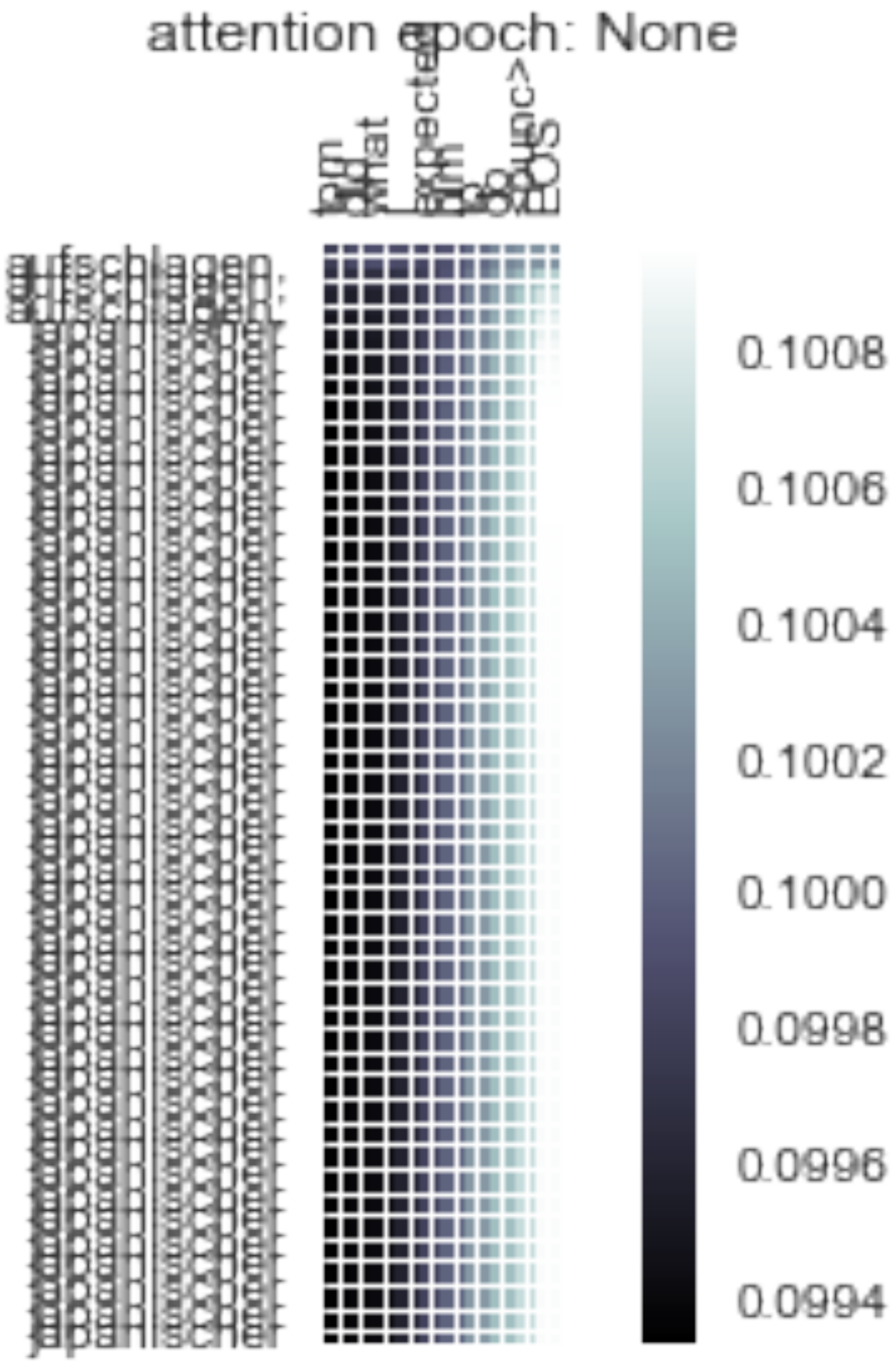
- Why? -> Dummy testing the implemented networks. See if everything works as it should

Examples

```
tom did what I expected him to do <punc> EOS PAD PAD PAD PAD PAD PAD PAD
```

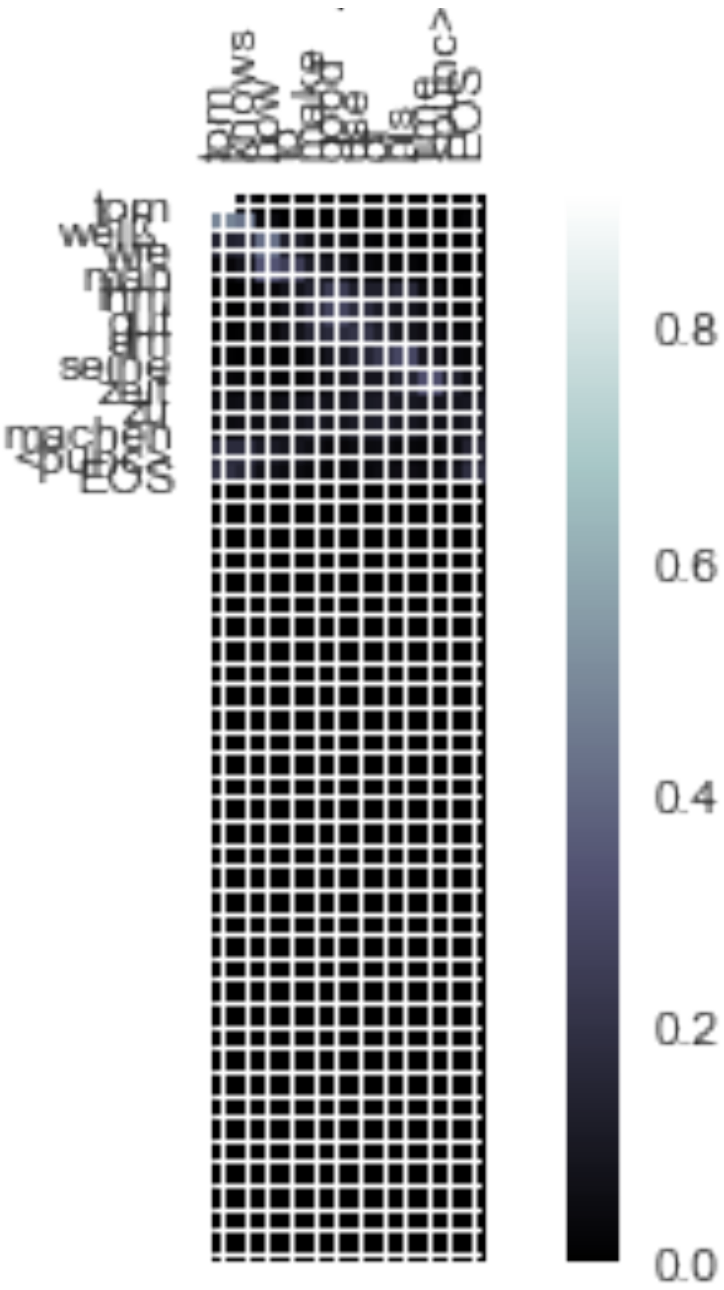
```
=target-sentence= tom tat, was ich von ihm erwartete <punc> EOS PAD PAD PAD PAD PAD
```

```
< output-sentence < aufschlagen, aufschlagen, aufschlagen, japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
japanischer japanischer japanischer japanischer japanischer japanischer japanischer
```

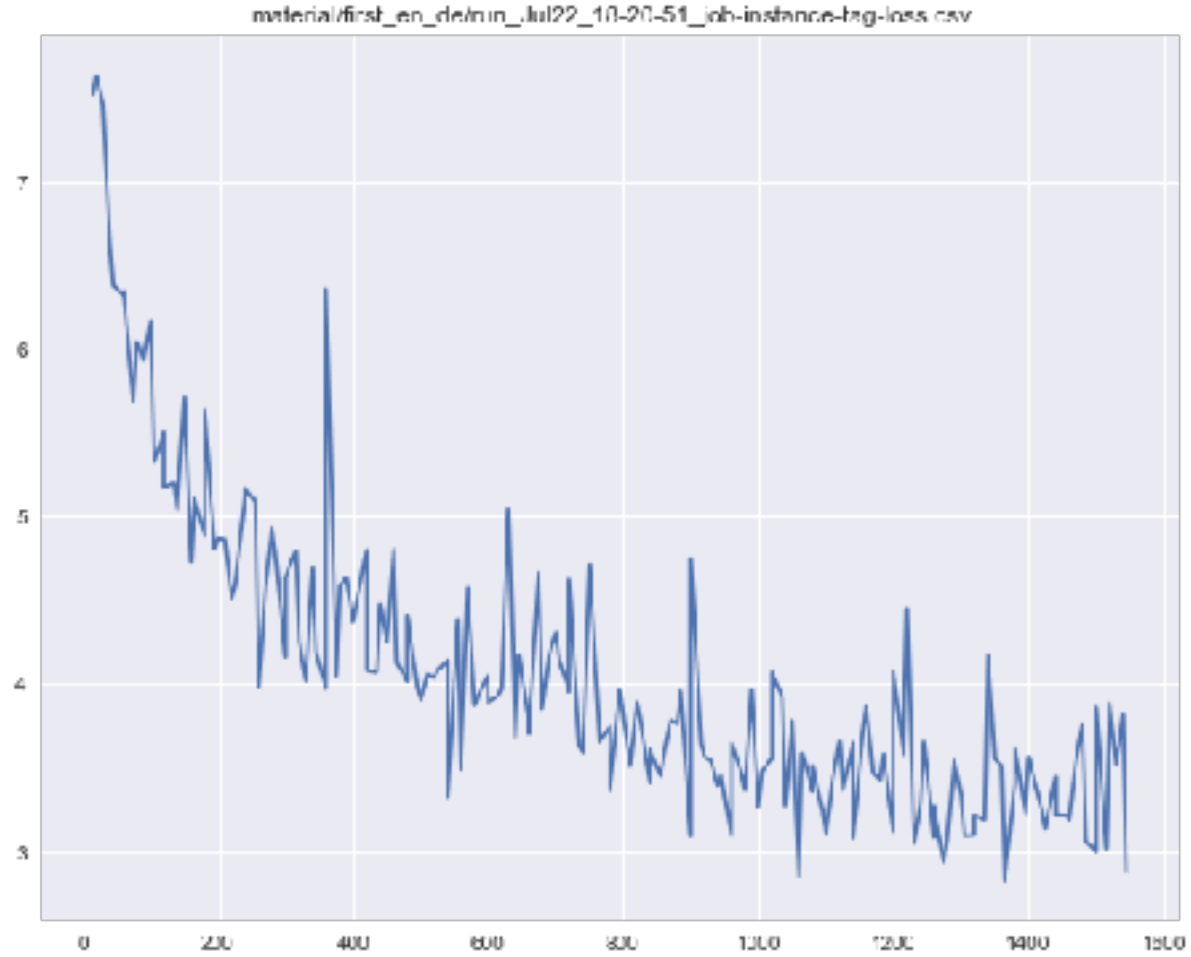


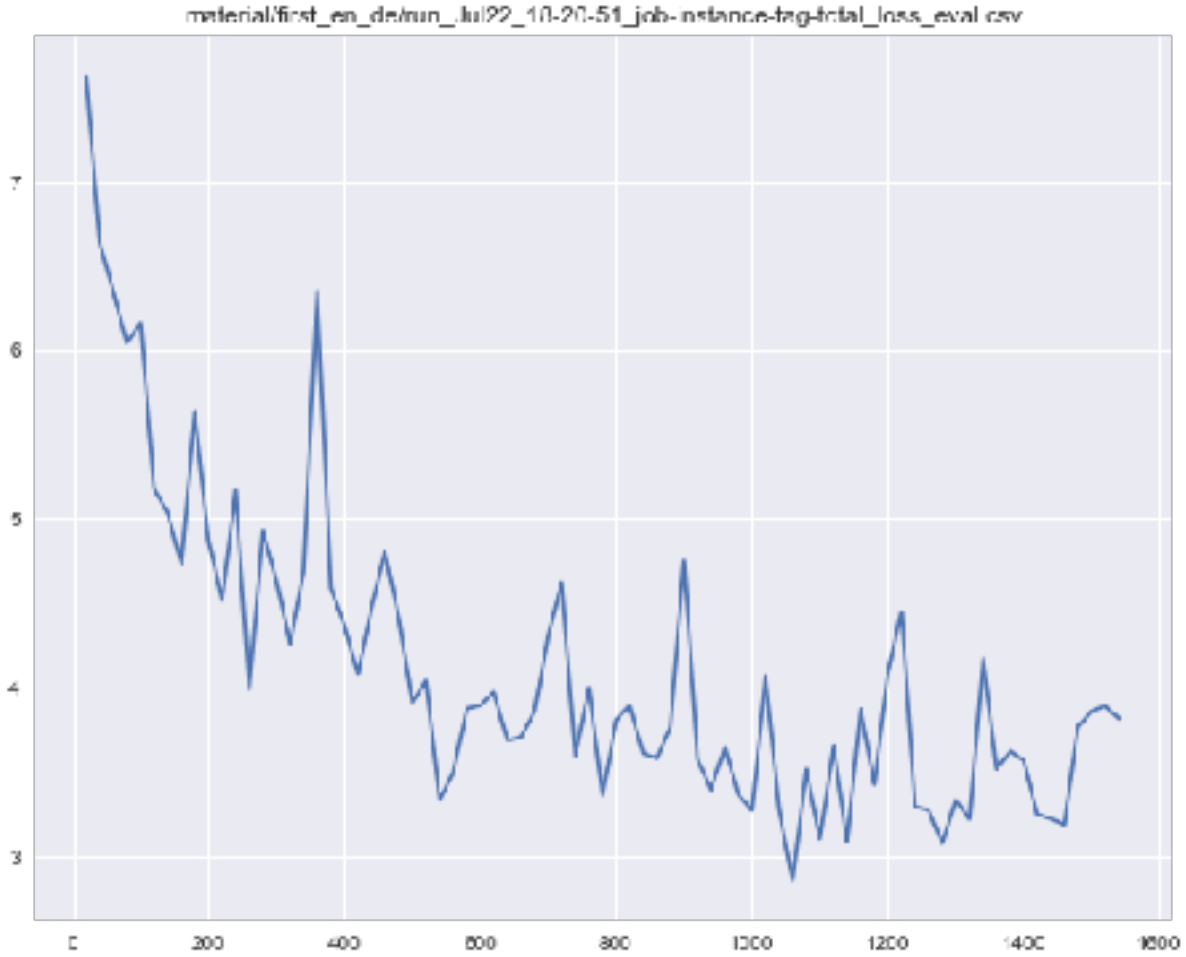
Examples

```
> input-sentence: > tom knows how to make good use of his time <punc>
EOS PAD PAD PAD PAD PAD PAD PAD


=target-sentence= tom versteht es, seine zeit gut einzuteilen <punc> EOS
PAD PAD PAD PAD PAD PAD PAD PAD PAD PAD


< output-sentence < tom weiß, wie man ihm gut um seine zeit zu machen
<punc> EOS
```

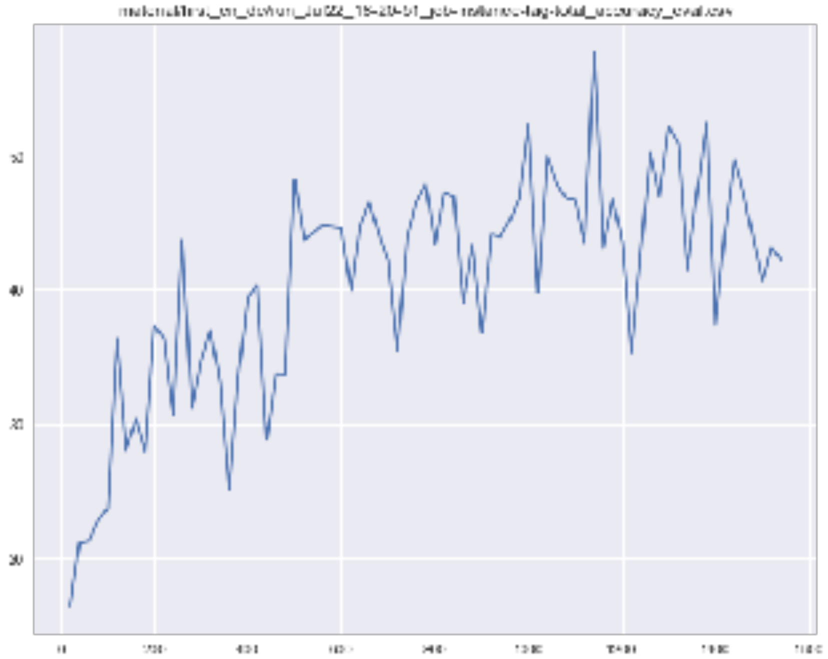


Loss


material/first_en_de/run_Jul22_10-20-51_job-instance-tag-loss.csv


material/first_en_de/run_Jul22_10-20-51_job-instance-tag-total_loss_eval.csv

Accuracy


material/first_en_de/run_Jul22_10-20-51_job-instance-tag-total_accuracy_eval.csv

English - Sparql

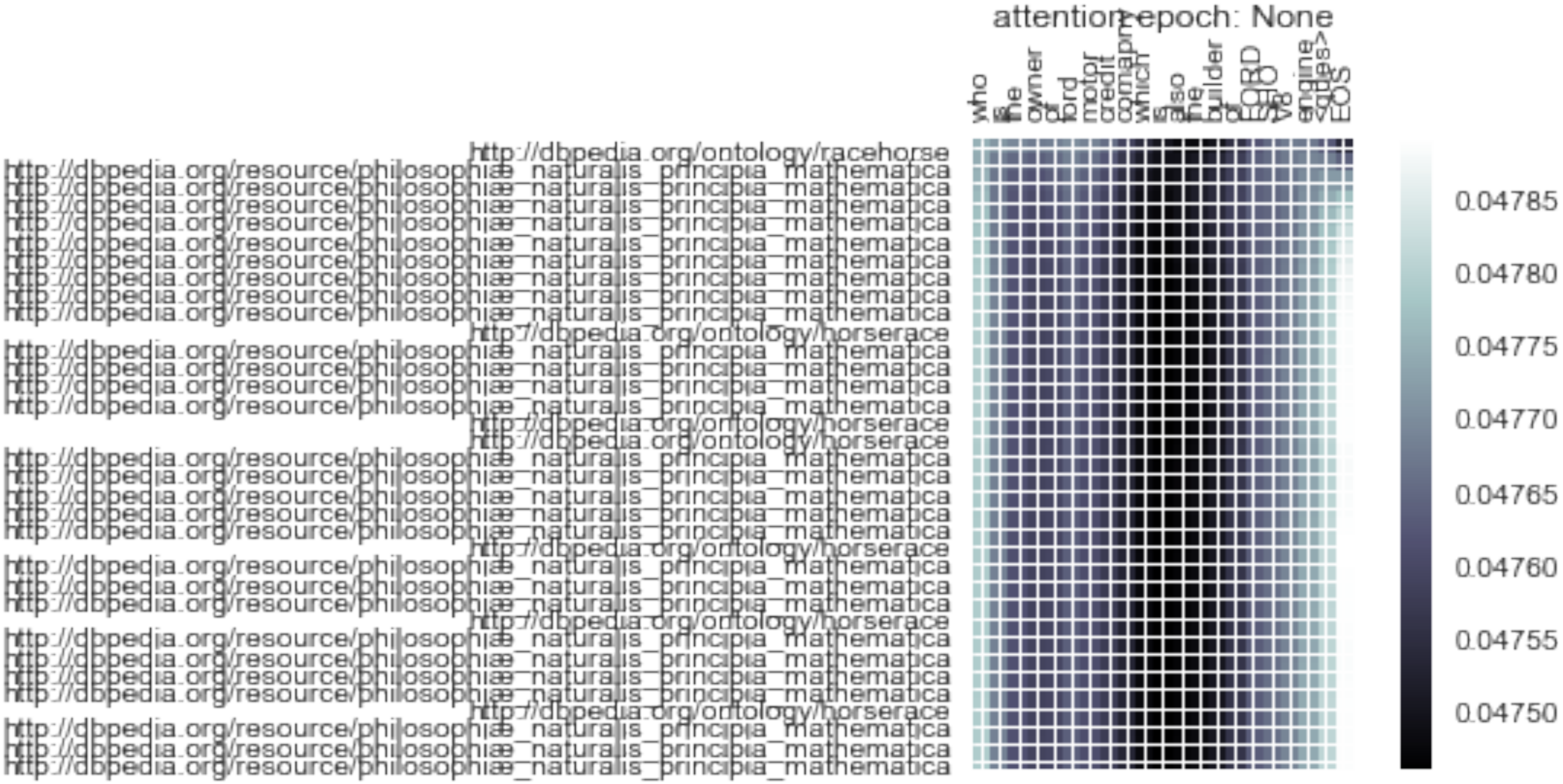
Examples

```
> input-sentence: > who is the owner of ford motor credit comapny which is also the builder of FORD SHO V8 engine <ques>
EOS
```

```
=target-sentence= SELECT DISTINCT <ques>uri WHERE <curly_open> <croc_open> http://dbpedia.org/resource/
ford_motor_credit_company <croc_close> <croc_open> http://dbpedia.org/property/owner <croc_close> <ques>uri. <croc_open>
http://dbpedia.org/resource/ford_SHO_V8_engine <croc_close> <croc_open> http://dbpedia.org/ontology/manufacturer
<croc_close> <ques>uri <curly_close> EOS PAD PAD PAD PAD PAD PAD PAD PAD PAD
```

```
< output-sentence < http://dbpedia.org/ontology/racehorse http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/ontology/horserace http://dbpedia.org/
resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/ontology/horserace http://dbpedia.org/
ontology/horserace http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/ontology/horserace http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/ontology/horserace http://dbpedia.org/
resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://
dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/ontology/horserace http://dbpedia.org/
resource/philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/
philosophiæ_naturalis_principia_mathematica http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
```

Examples


attention-epoch: None
who
is
the
owner
of
ford
motor
credit
company
which
is
also
the
builder
of
FORD
SOHC
V8
engine
<ales>
EOS
http://dbpedia.org/ontology/racehorse
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/ontology/horserace
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/ontology/horserace
http://dbpedia.org/ontology/horserace
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/ontology/horserace
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/ontology/horserace
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/ontology/horserace
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
http://dbpedia.org/resource/philosophiæ_naturalis_principia_mathematica
0.04785
0.04780
0.04775
0.04770
0.04765
0.04760
0.04755
0.04750

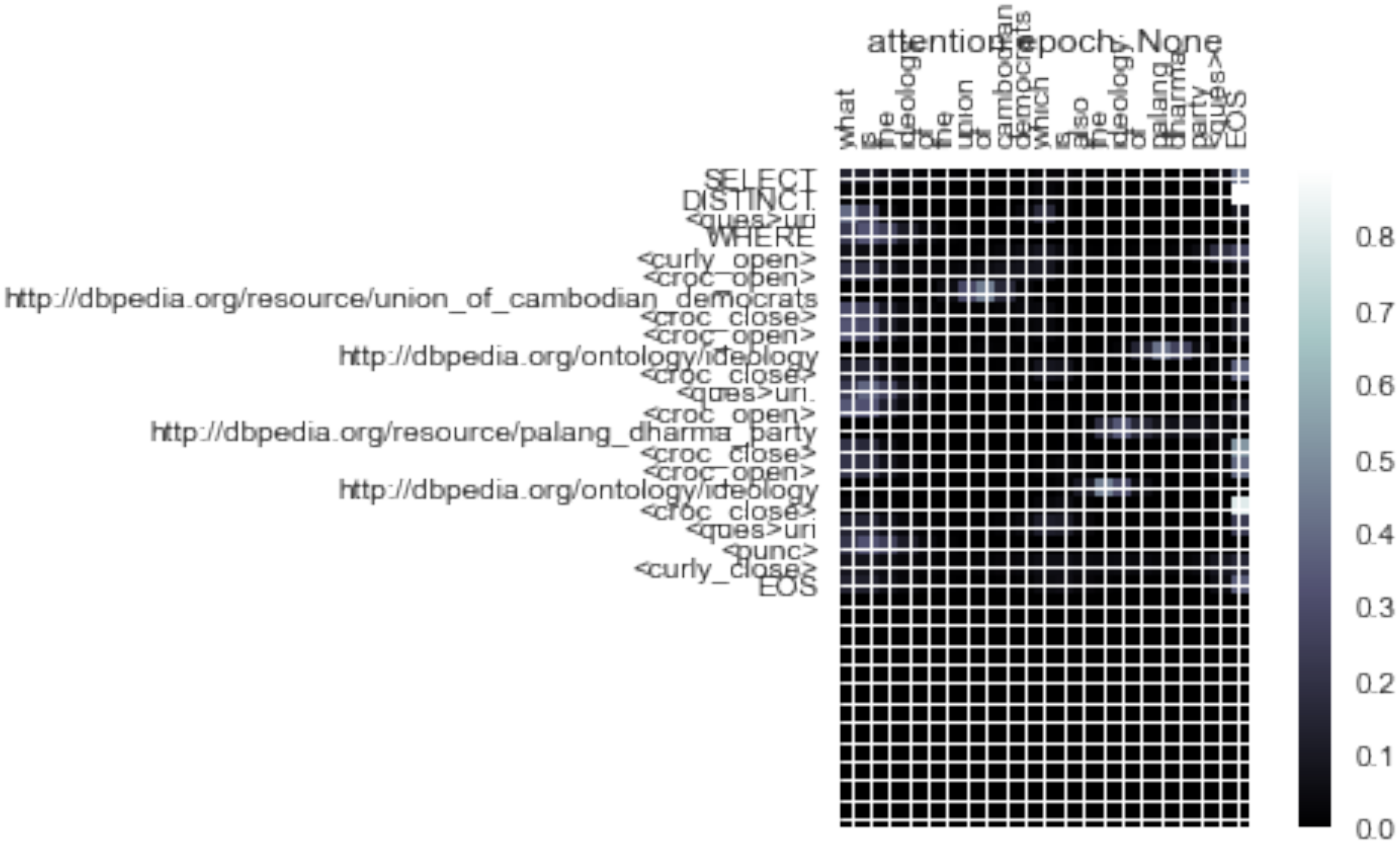
Examples (on train set after 2100 epochs)

```
what is the ideology of the union of cambodian democrats which is also the
ideology of palang dharma party <ques> EOS PAD
```

```
=target-sentence= SELECT DISTINCT <ques>uri WHERE <curly_open> <croc_open>
http://dbpedia.org/resource/union_of_cambodian_democrats <croc_close>
<croc_open> http://dbpedia.org/ontology/ideology <croc_close> <ques>uri.
<croc_open> http://dbpedia.org/resource/palang_dharma_party <croc_close>
<croc_open> http://dbpedia.org/ontology/ideology <croc_close> <ques>uri <punc>
<curly_close> EOS PAD PAD PAD PAD PAD PAD PAD PAD
```

```
< output-sentence < SELECT DISTINCT <ques>uri WHERE <curly_open> <croc_open>
http://dbpedia.org/resource/union_of_cambodian_democrats <croc_close>
<croc_open> http://dbpedia.org/ontology/ideology <croc_close> <ques>uri.
<croc_open> http://dbpedia.org/resource/palang_dharma_party <croc_close>
<croc_open> http://dbpedia.org/ontology/ideology <croc_close> <ques>uri <punc>
<curly_close> EOS
```

Examples (on train set after 2100 epochs)



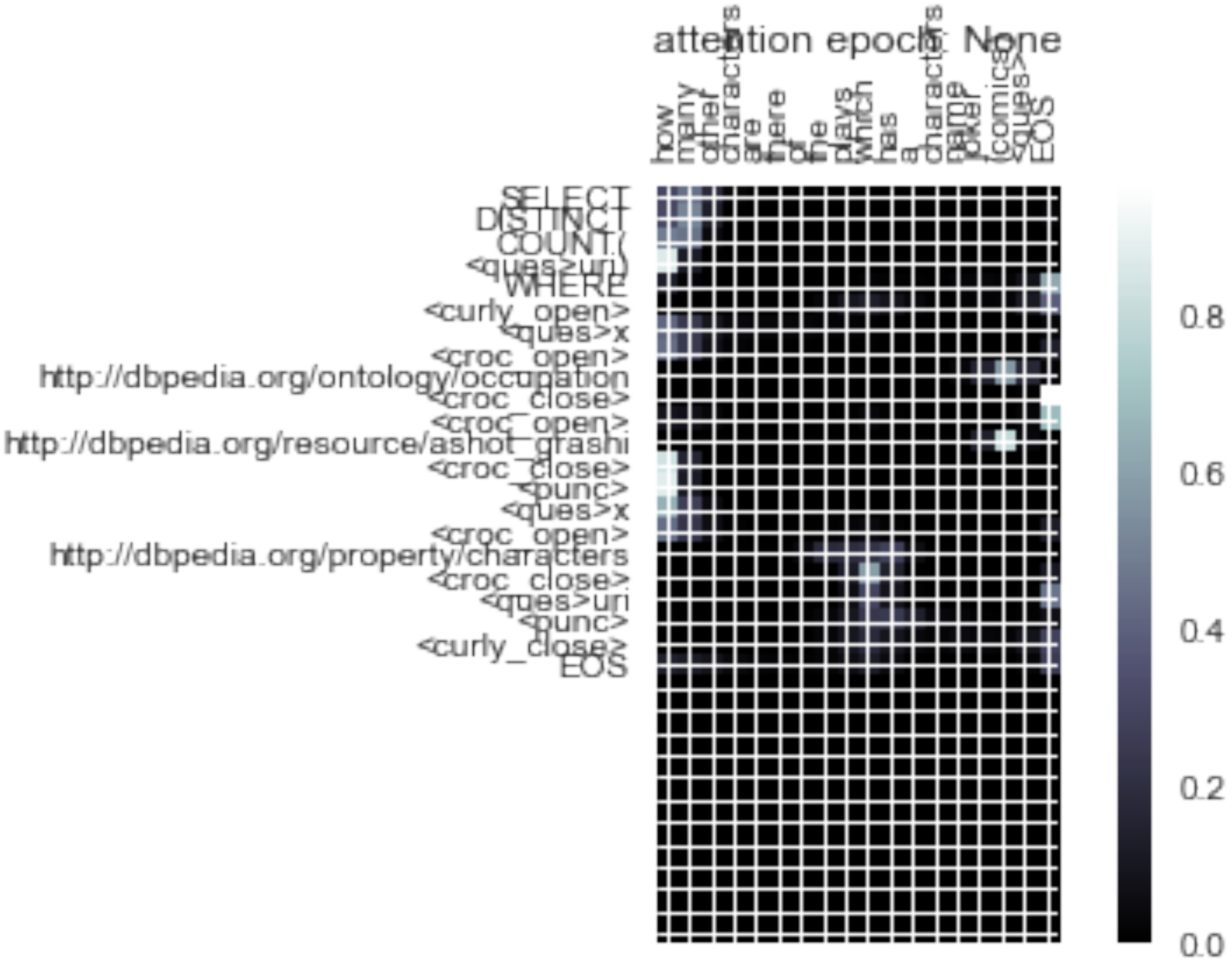
Examples (on validation set after 2100 epochs)

```
> input-sentence: > how many other characters are there of the plays which has
a characters name joker (comics) <ques> EOS PAD PAD PAD
```

```
=target-sentence= SELECT DISTINCT COUNT( <ques>uri) WHERE <curly_open> <ques>x
<croc_open> http://dbpedia.org/property/characters <croc_close> <croc_open>
http://dbpedia.org/resource/joker_(comics) <croc_close> <punc> <ques>x
<croc_open> http://dbpedia.org/property/characters <croc_close> <ques>uri
<curly_close> EOS PAD PAD PAD PAD PAD PAD PAD PAD PAD
```

```
< output-sentence < SELECT DISTINCT COUNT( <ques>uri) WHERE <curly_open>
<ques>x <croc_open> http://dbpedia.org/ontology/occupation <croc_close>
<croc_open> http://dbpedia.org/resource/ashot_grashi <croc_close> <punc>
<ques>x <croc_open> http://dbpedia.org/property/characters <croc_close>
<ques>uri <punc> <curly_close> EOS
```

Examples (on validation set after 2100 epochs)



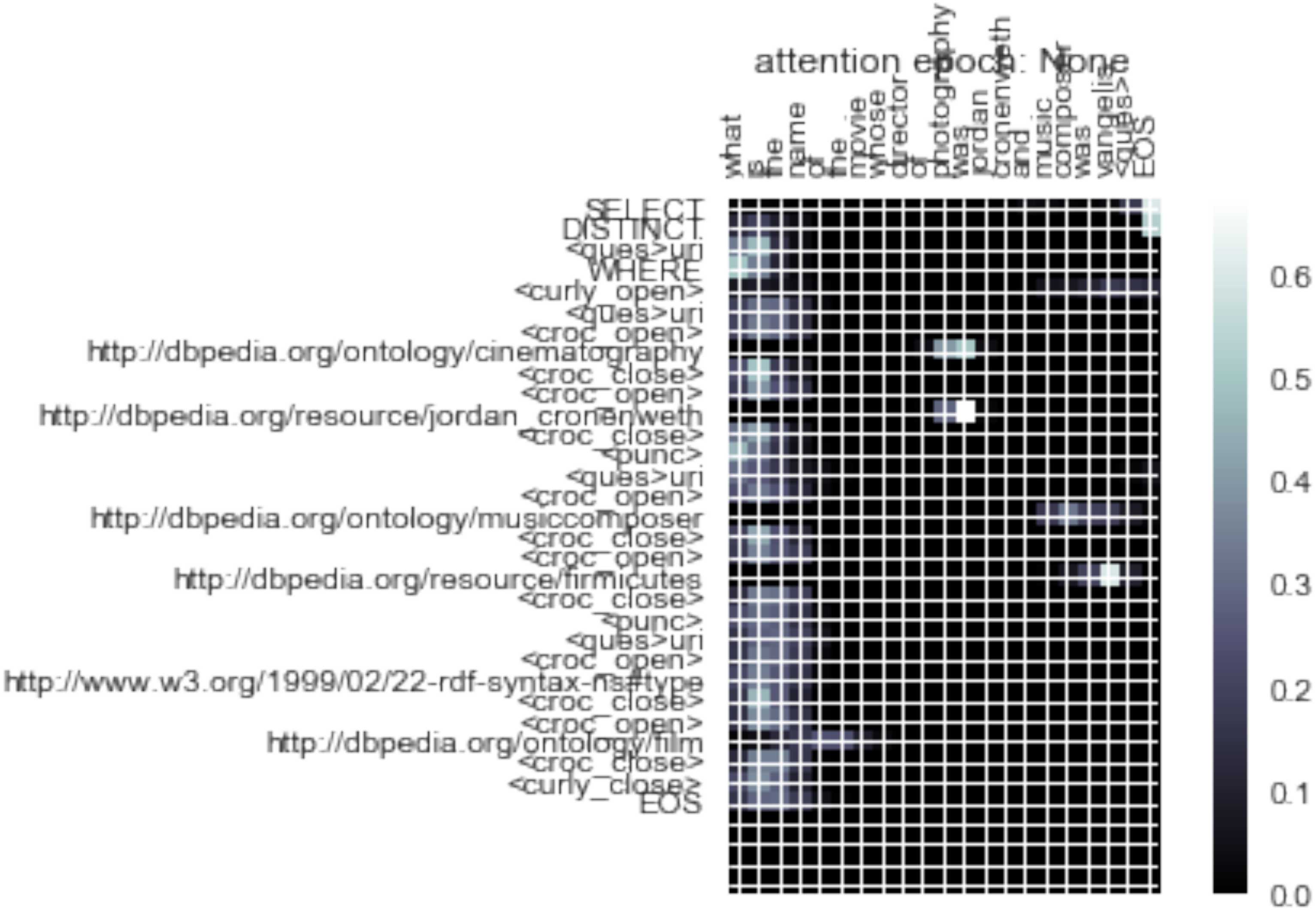
Examples (on validation set after 2100 epochs)

```
> input-sentence: > what is the name of the movie whose director of photography
was jordan cronenweth and music composer was vangelis <ques> EOS
```

```
=target-sentence= SELECT DISTINCT <ques>uri WHERE <curly_open> <ques>uri
<croc_open> http://dbpedia.org/property/cinematography <croc_close> <croc_open>
http://dbpedia.org/resource/jordan_cronenweth <croc_close> <punc> <ques>uri
<croc_open> http://dbpedia.org/ontology/musiccomposer <croc_close> <croc_open>
http://dbpedia.org/resource/vangelis <croc_close> <punc> <ques>uri <croc_open>
http://www.w3.org/1999/02/22-rdf-syntax-ns#type <croc_close> <croc_open>
http://dbpedia.org/ontology/film <croc_close> <curly_close> EOS
```

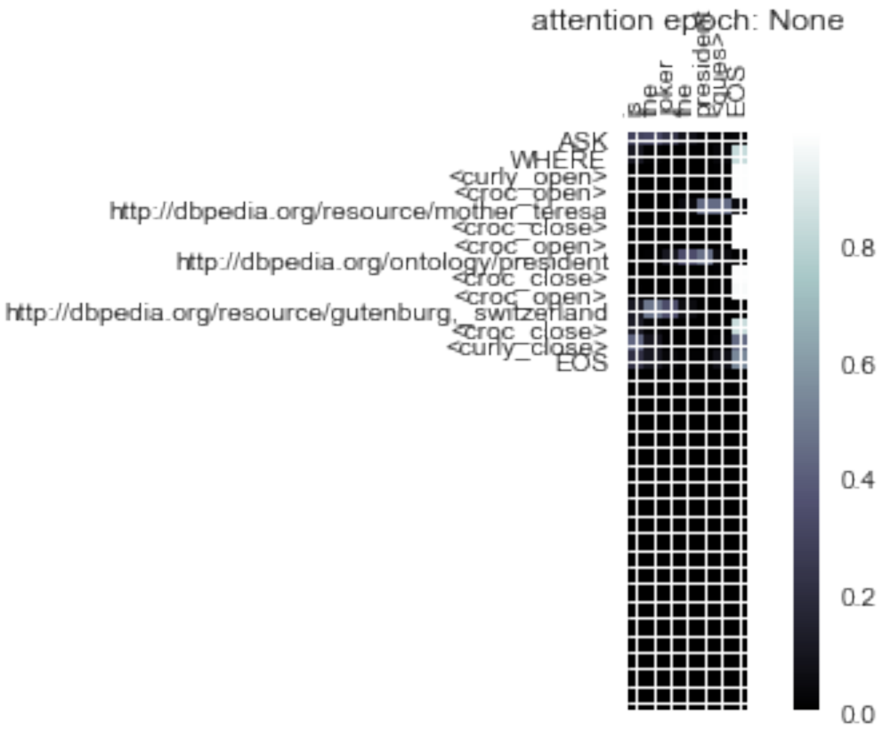
```
< output-sentence < SELECT DISTINCT <ques>uri WHERE <curly_open> <ques>uri
<croc_open> http://dbpedia.org/ontology/cinematography <croc_close> <croc_open>
http://dbpedia.org/resource/jordan_cronenweth <croc_close> <punc> <ques>uri
<croc_open> http://dbpedia.org/ontology/musiccomposer <croc_close> <croc_open>
http://dbpedia.org/resource/firmicutes <croc_close> <punc> <ques>uri
<croc_open> http://www.w3.org/1999/02/22-rdf-syntax-ns#type <croc_close>
<croc_open> http://dbpedia.org/ontology/film <croc_close> <curly_close> EOS
```

Examples (on validation set after 2100 epochs)

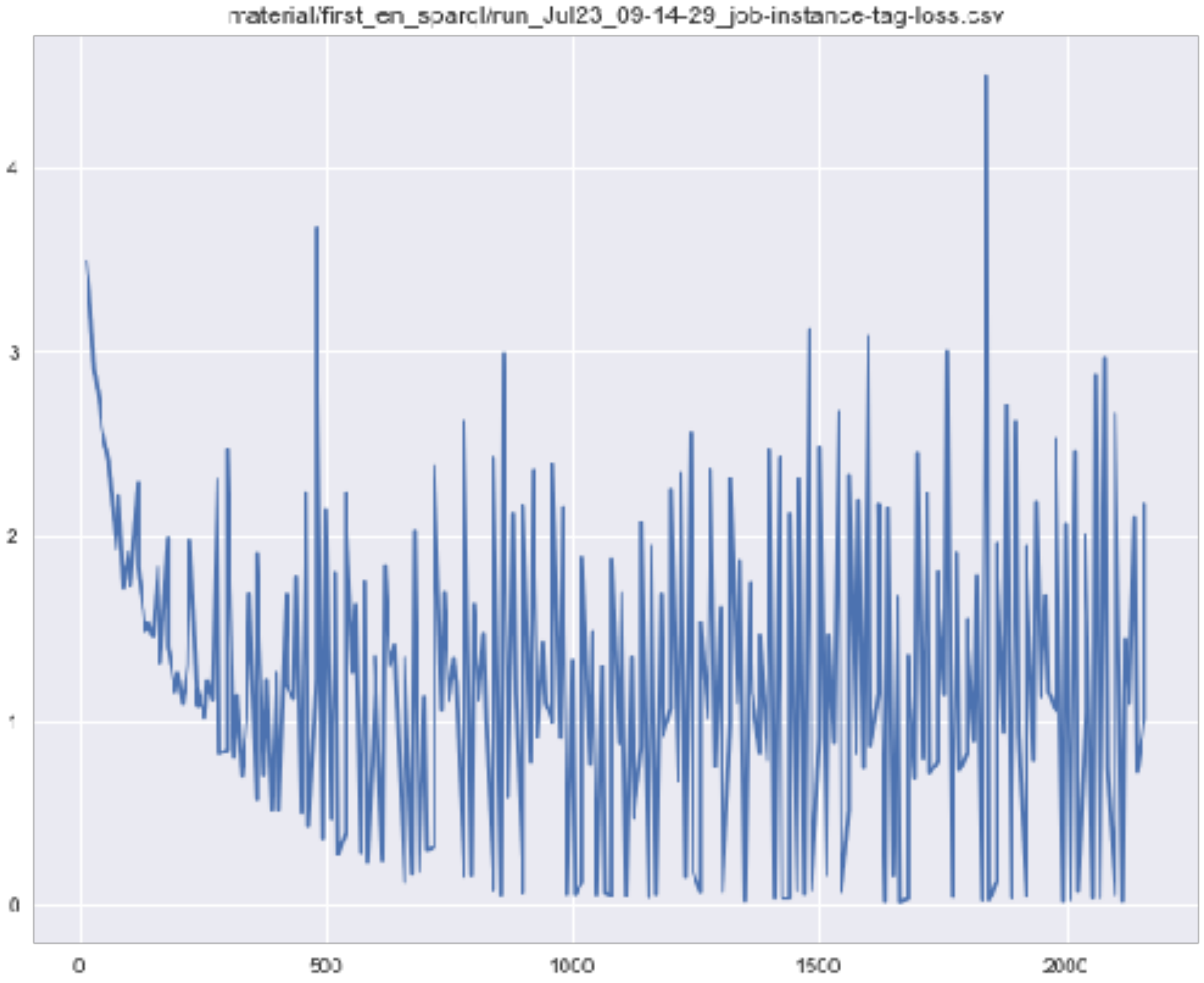


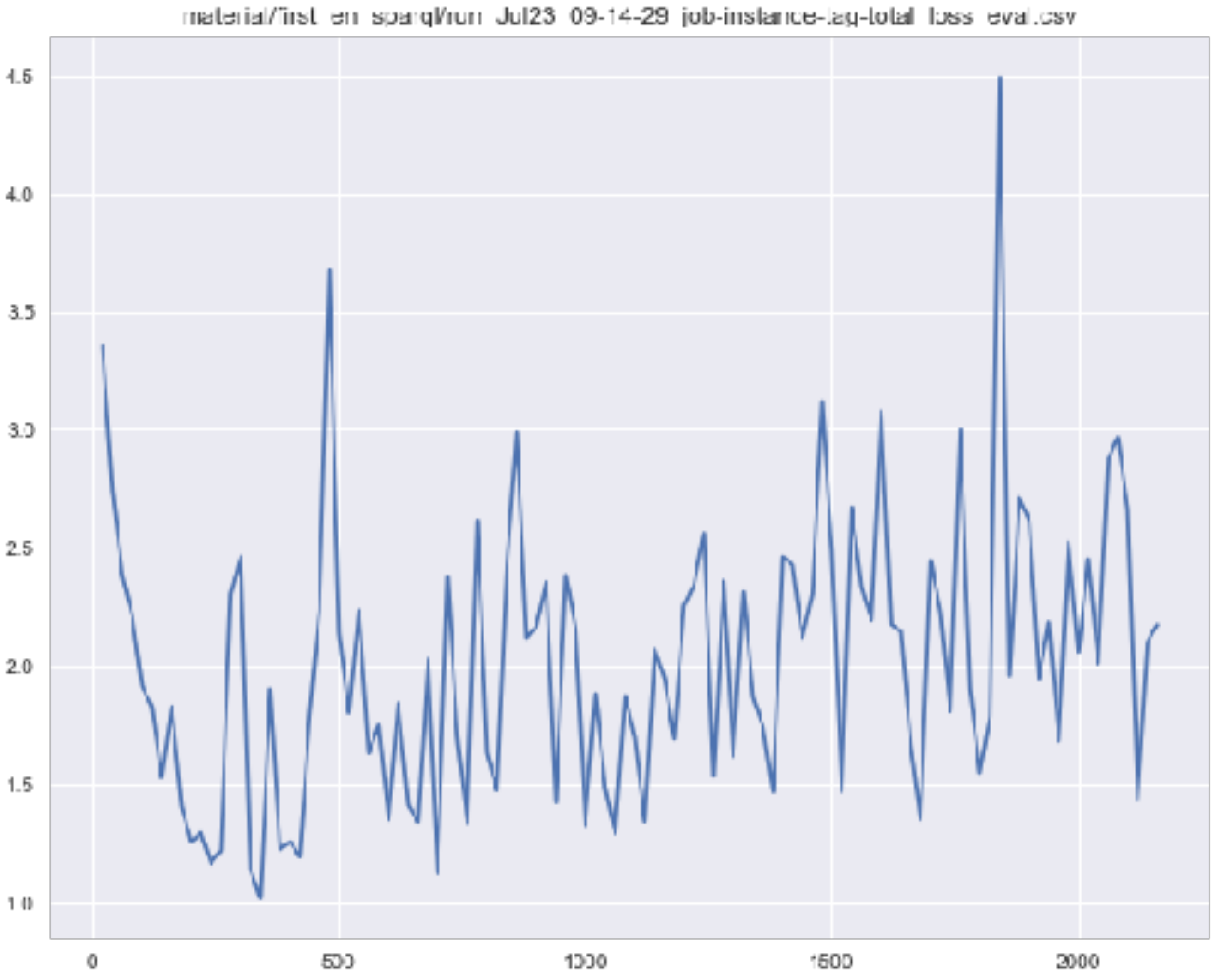
Examples (random question)

```
> input-sentence: > is the joker the president <ques> EOS



< output-sentence < ASK WHERE <curly_open> <croc_open> http://dbpedia.org/
resource/mother_teresa <croc_close> <croc_open> http://dbpedia.org/ontology/
president <croc_close> <croc_open> http://dbpedia.org/resource/
gutenburg,_switzerland <croc_close> <curly_close> EOS



FORMATED:

ASK WHERE {<http://dbpedia.org/resource/mother_teresa> <http://dbpedia.org/
ontology/president> <http://dbpedia.org/resource/gutenburg,_switzerland>} EOS
```

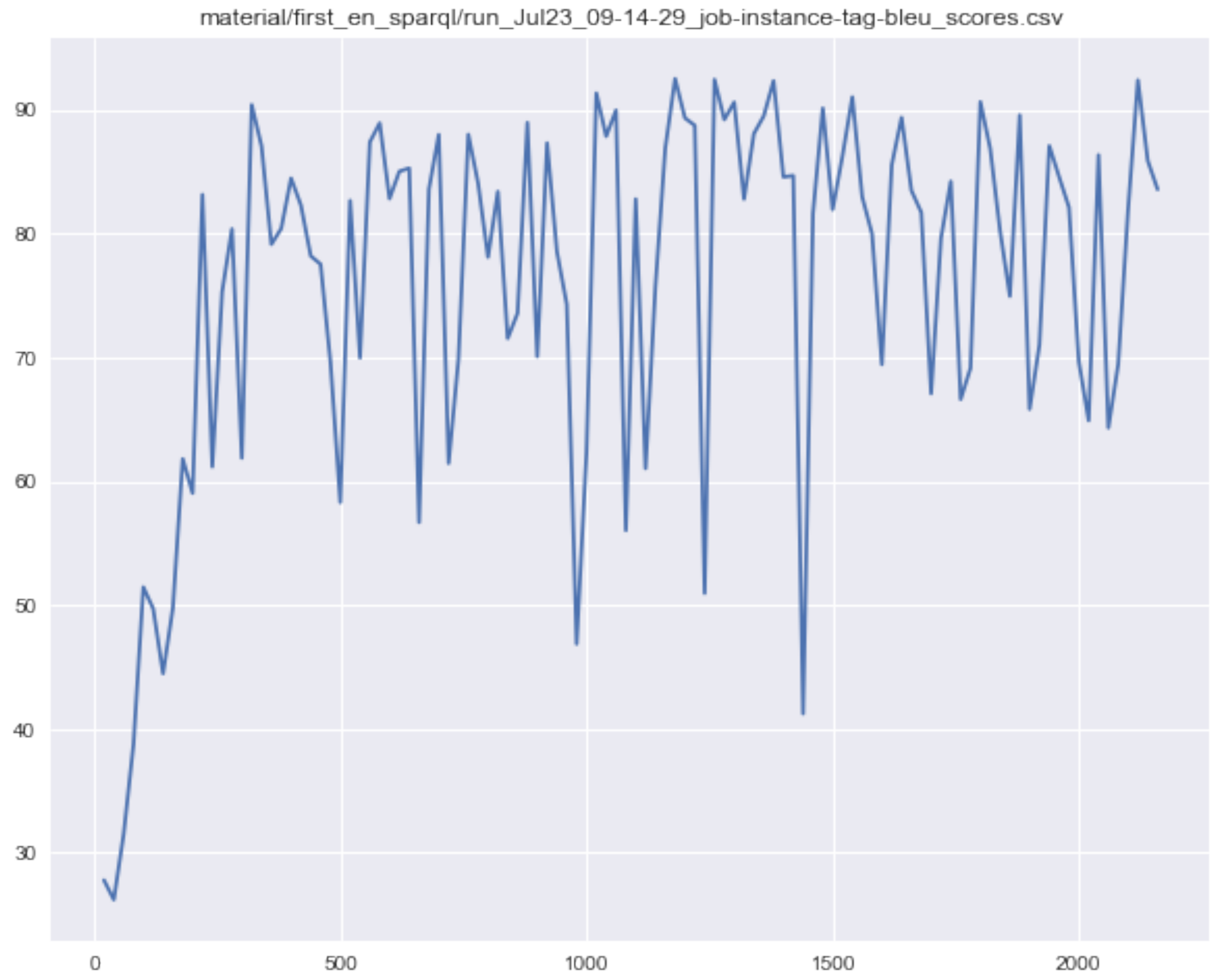


Loss


material/first_en_sparql/run_Jul23_09-14-29_job-instance-tag-loss.csv
4
3
2
1
0
0
500
1000
1500
2000


material/first_en_sparql/run_Jul23_09-14-29_job-instance-tag-total_loss_eval.csv
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
500
1000
1500
2000

Accuracy


material/first_en_sparql/run_Jul23_09-14-29_job-instance-tag-bleu_scores.csv
90
80
70
60
50
40
30
0
500
1000
1500
2000

Lessons learned

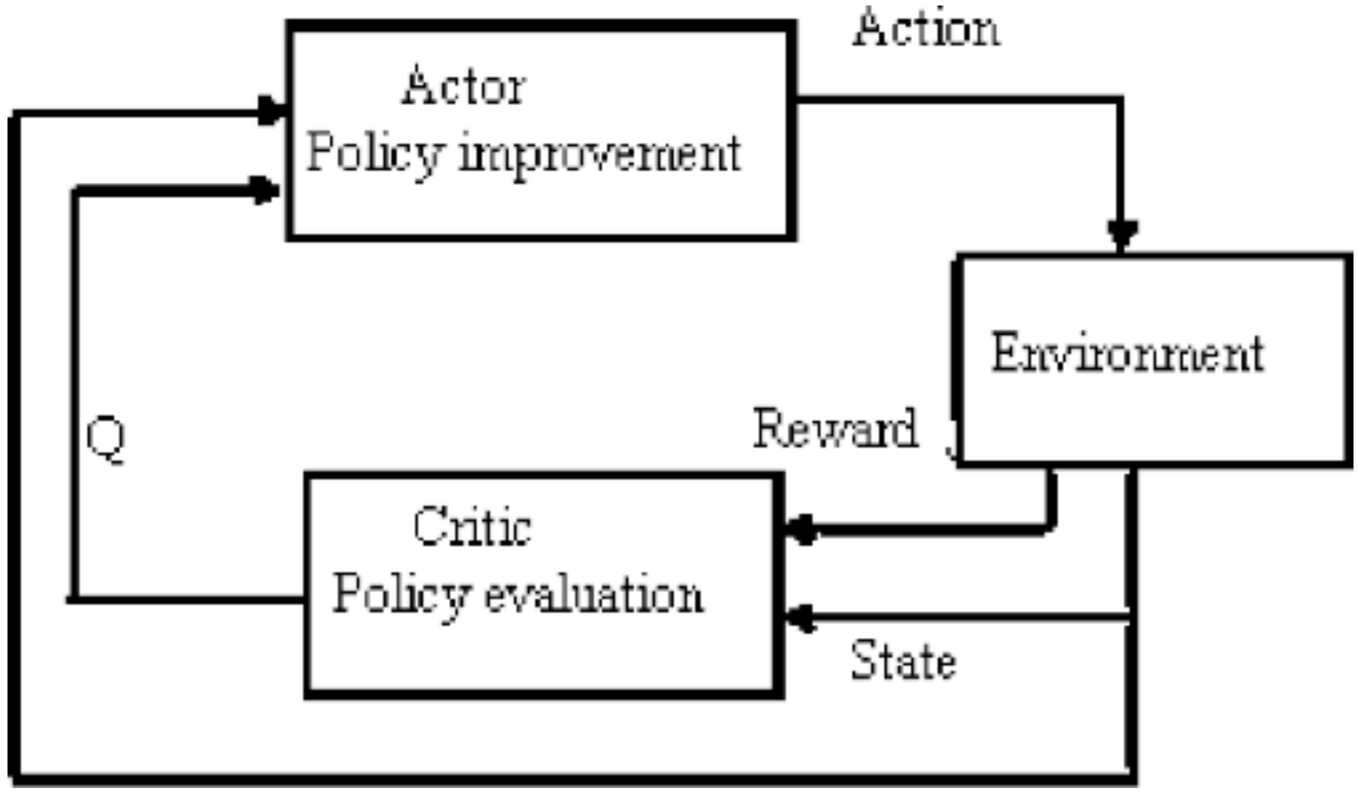
- Lower case seems not to be an improvement
- From very early on the network captures the “grammar” (SELECT COUNT WHERE ASK <> {})
- The mapping from word to its entity … is the hardest part for the network
- Take time for hyper parameter search

Flaws & Improvements

- Works only for words in the embeddings: use a shared embedding or a pretained (fasttext, glove,…)
- Mapping still hard -> maybe other embeddings help
- Training with larger embedding
- Without / degraded teacher forcing ?
- Beam search

Reinforced Training

- Next steps use the spraql query results to update (based on bandit-nmt)


Action
Actor
Policy improvement
Environment
Q
Reward
Critic
Policy evaluation
State

References

- https://github.com/spro/practical-pytorch/blob/master/seq2seq-translation/seq2seq-translation.ipynb
- https://github.com/spro/practical-pytorch/blob/master/seq2seq-translation/seq2seq-translation-batched.ipynb
- https://arxiv.org/pdf/1707.07402.pdf
- https://github.com/khanhptnk/bandit-nmt
- http://opennmt.net/