Learning to Ignore: Long Document Coreference with Bounded Memory Neural Networks

Shubham Toshniwal Sam Wiseman Allyson Ettinger
Karen Livescu Kevin Gimpel

github.com/shtoshni92/long-doc-coref

Coreference Resolution

I had seen little of Holmes lately. My marriage had drifted us away from each other. He was still, as ever, deeply attracted by the study of crime.

Coreference Resolution

I$_1$ had seen little of Holmes$_2$ lately. My$_1$ marriage had drifted us$_3$ away from each other. He$_2$ was still, as ever, deeply attracted by the study of crime.

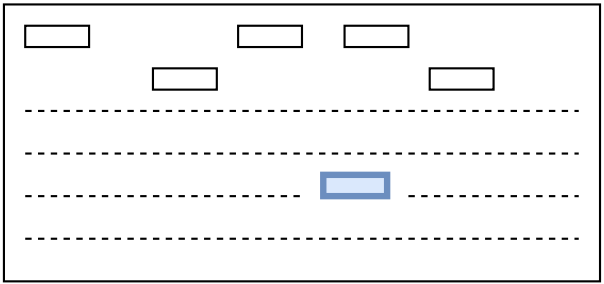
Coreference Resolution

I_1 had seen little of $Holmes_2$ lately. My_1 marriage had drifted us_3 away from each other. He_2 was still, as ever, deeply attracted by the study of crime.

$$
\begin{aligned}
1 &\longrightarrow \text{John Watson} \\
2 &\longrightarrow \text{Sherlock Holmes} \\
3 &\longrightarrow \text{Watson} + \text{Holmes}
\end{aligned}
$$

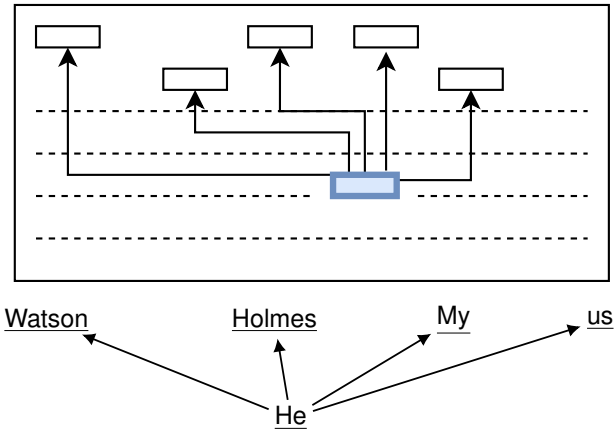
Coreference Resolution Models

Mention Ranking Models



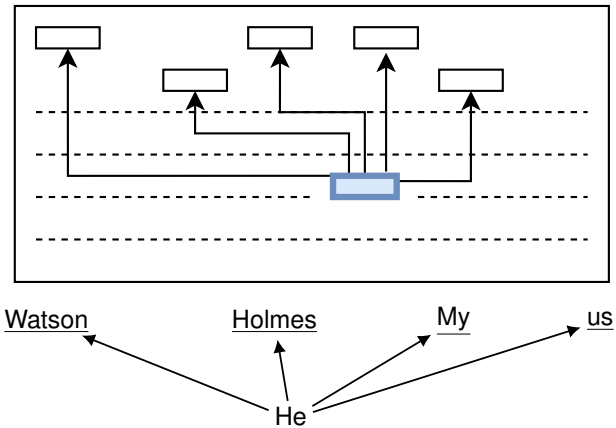
Lee et al 2018, Joshi et al 2019,
Wu et al 2020

Mention Ranking Models



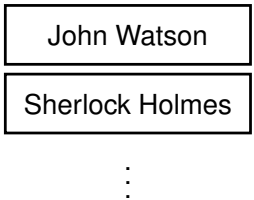
Lee et al 2018, Joshi et al 2019,
Wu et al 2020

Mention Ranking Models



Impractical for long documents!
Quadratic runtime!

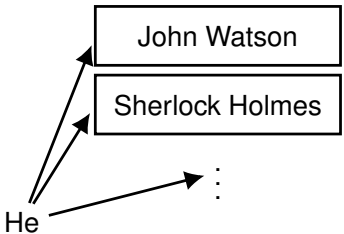
Entity-Mention Models



Entities

Webster et al 2014, Xia et al 2020

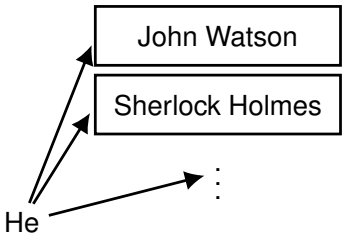
Entity-Mention Models



Entities

Webster et al 2014, Xia et al 2020

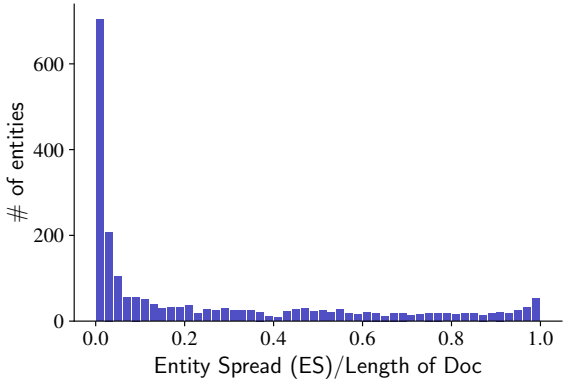
Entity-Mention Models



Entities

Number of entities can be quite large!

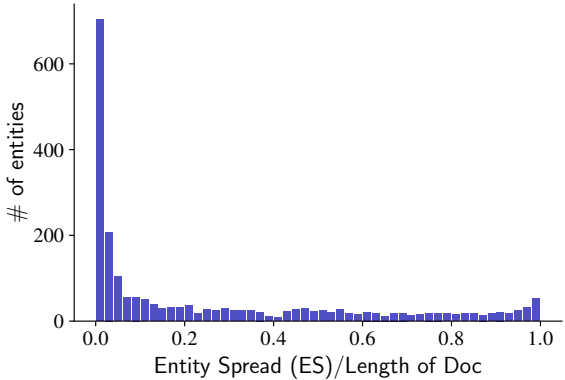
Most Entities Are Transient



LitBank Entity Spread Histogram

Most entities have a small “spread”: # of tokens between first and last mention.

Most Entities Are Transient



LitBank Entity Spread Histogram

Most entities have a small “spread”: # of tokens between first and last mention.

Not necessary to keep all entities in memory all the time!

Bounded Memory Model: Ignore and Evict

- Track a small, bounded number of entities.
- When the model's memory is full, and a mention corresponding to a new (currently untracked) entity comes next, then :
  - *Evict*: Remove an entity already being tracked, and start tracking this new entity.
  - *Ignore*: Ignore the mention.
- Learns to ignore and evict by mimicking oracle actions.

Results

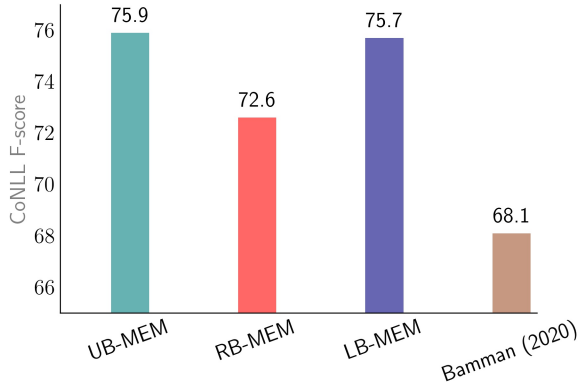
Model Variants

We compare the following model variants:

- Unbounded Memory (UB-MEM)
- Rule-based Bounded Memory (RB-MEM): Uses LRU heuristic to select eviction candidate.
- Learned Bounded Memory (LB-MEM): Proposed model which learns to ignore/evict entities.

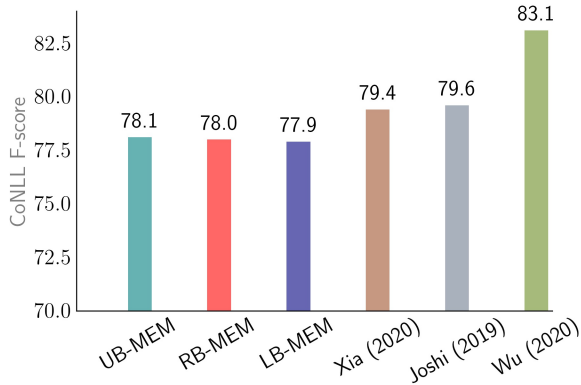
For bounded memory models, we only show results with memory capacity of 20 entities.

LitBank Results

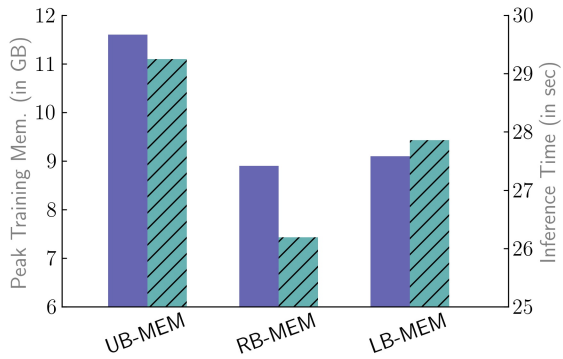


State of the art results for LitBank!

OntoNotes Results



Memory and Inference Time Comparison



Conclusion

- Proposed a memory model that tracks a small, bounded number of entities.
- The model is competitive with prior work on OntoNotes and LitBank.
- The model guarantees linear runtime in length of document, and reduces memory usage during training.

Bibliography

David Bamman, Olivia Lewke, and Anya Mansoor. 2020.
An Annotated Dataset of Coreference in English Literature.
In *LREC*.

Mandar Joshi, Omer Levy, Luke Zettlemoyer, and Daniel Weld. 2019.
BERT for Coreference Resolution: Baselines and Analysis.
In *EMNLP*.

Kellie Webster and James R. Curran. 2014.
Limited memory incremental coreference resolution.
In *COLING*.

Wei Wu, Fei Wang, Arianna Yuan, Fei Wu, and Jiwei Li. 2020.
Coreference Resolution as Query-based Span Prediction.
In *ACL*.

Patrick Xia, João Sedoc, and Benjamin Van Durme. 2020.
Revisiting Memory-Efficient Incremental Coreference Resolution.
In *EMNLP*.