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
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.
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.
I\textsubscript{1} had seen little of Holmes\textsubscript{2} lately. My\textsubscript{1} marriage had drifted us\textsubscript{3} away from each other. He\textsubscript{2} was still, as ever, deeply attracted by the study of crime.

1 \rightarrow John Watson  
2 \rightarrow Sherlock Holmes  
3 \rightarrow Watson + Holmes
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

John Watson

Sherlock Holmes

Webster et al 2014, Xia et al 2020

Entities
Entity-Mention Models

John Watson

Sherlock Holmes

He

Webster et al 2014, Xia et al 2020
Entity-Mention Models

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

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

![Bar chart showing CoNLL F-score for different models: UB-MEM 78.1, RB-MEM 78.0, LB-MEM 77.9, Xia (2020) 79.4, Joshi (2019) 79.6, Wu (2020) 83.1]
Memory and Inference Time Comparison

- UB-MEM
- RB-MEM
- LB-MEM

Peak Training Mem. (in GB)

Inference Time (in sec)
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.


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.