Statistical Script Learning with Recurrent Neural Networks

Karl Pichotta and Raymond J. Mooney
The University of Texas at Austin
{pichotta,mooney}@cs.utexas.edu

Simplifying Events to Make Event Co-occurrence Learnable

- Scripts [Schank & Abelson 1975] are an influential early encoding of situation-specific world event

```
script:  restaurant
roles:   customer, waitress, chef, cashier
reason:  to get food so as to go up in pleasure
            and down in hunger

scene 1: entering
         PTRANS self into restaurant
         ATTEND eyes to where empty tables are
         MBUILD where to sit
         PTRANS self to table
         MOVE sit down

scene 2: ordering
         ATRANS receive menu
         MTRANS read menu
         MBUILD decide what self wants
         MTRANS order to waitress

scene 3: eating
         ATRANS receive food
         INGEST food

scene 4: exiting
         MTRANS ask for check
         ATRANS receive check
         ATRANS tip to waitress
         PTRANS self to cashier
         ATRANS money to cashier
         PTRANS self out of restaurant
```

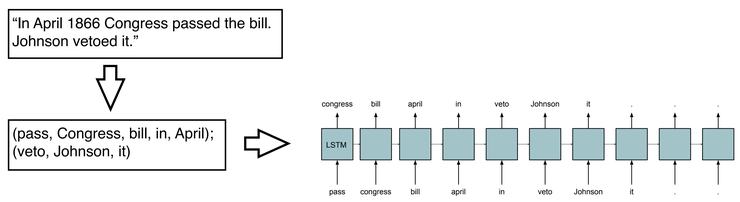
(From [Schank & Abelson 1975])

- In order to **learn** scripts automatically **from data**, one must simplify events:
  - [Chambers & Jurafsky 2008]: (verb, dependency) pairs (e.g. (*receive*, subject))
  - [Pichotta & Mooney 2014]: *n*-ary verb-with-arguments relations (e.g. (receive, *X*, *Y*))
- Build **statistical models** of event co-occurrence from large corpus of documents.

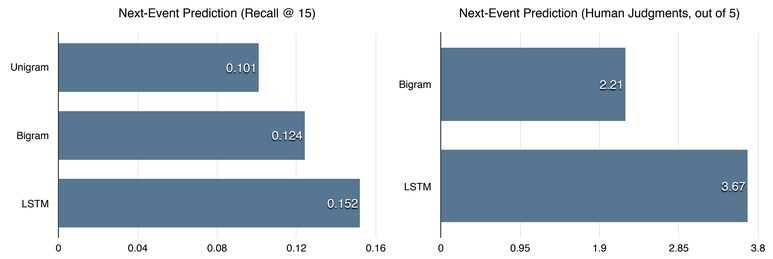
RNN Event Co-occurrence Models

[Pichotta & Mooney 2016a]

- Use an **LSTM sequence model** to model sequences of events in text.
- To **infer events**, input document events and use **beam search** over next event's components.
- Evaluate using a **cloze** test: hold an event out and try to infer it, given other events.



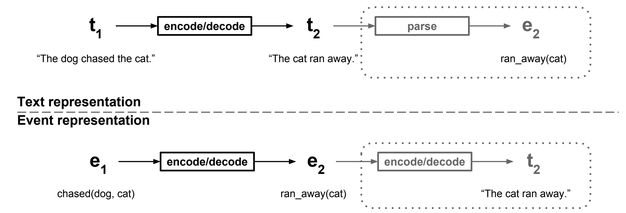
Systems:
- **Unigram**: Always guess most common events.
- **Bigram**: Smoothed event-level bigram Language Model
- **LSTM**: RNN event sequence model.

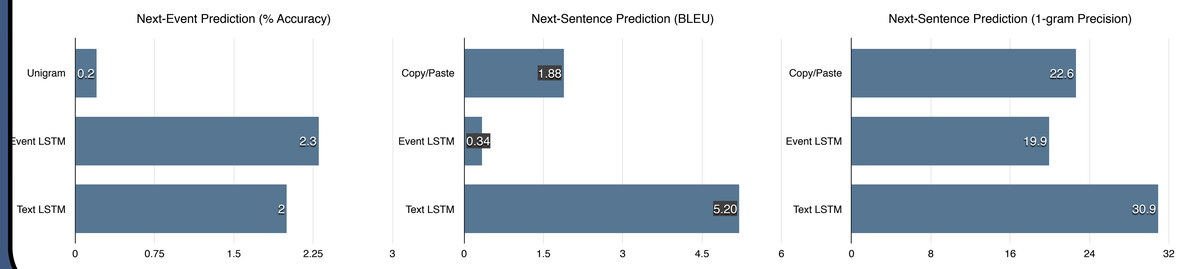


Modeling and Predicting Raw Text

[Pichotta & Mooney 2016b]

- What if we simply use **raw text** as our representation?
- Use the skip-thought approach [Kiros et al. 2015]: **sentence-level encoder/decoder** of raw text sequences.
  - Train LSTM sequence-to-sequence model to decode **whole sentences**.
  - Predict **text**, parse into **events**.
- No parser needed.
- Can evaluate on ability to predict either **events** (accuracy) or **raw text** (BLEU).





Future Work

- The competitiveness of the raw-text models indicates we should investigate other **extrinsic evaluations.**
- Modeling event sequences in **documents** and in the **world** are different tasks.
- How to **enrich event representation**?
- How to **enrich co-occurrence structure**?

References

[Chambers & Jurafsky 2008] Nathanael Chambers and Daniel Jurafsky. 2008. Unsupervised learning of narrative event chains. ACL 2008.

[Kiros et al. 15] Ryan Kiros, Yukun Zhu, Ruslan Salakhutdinov, Richard S. Zemel, Antonio Torralba, Raquel Urtasun, and Sanja Fidler. 2015. Skip-thought vectors. NIPS 2015.

[Pichotta & Mooney 2014] Karl Pichotta and Raymond J. Mooney. 2014. Statistical script learning with multi-argument events. EACL 2014.

[Pichotta & Mooney 2016a] Karl Pichotta and Raymond J. Mooney. 2016. Learning statistical scripts with LSTM recurrent neural networks. AAAI 2016.

[Pichotta & Mooney 2016b] Karl Pichotta and Raymond J. Mooney. 2016. Using sentence-level LSTM language models for script inference. ACL 2016.

[Schank & Abelson 1975] Roger C. Schank, and Robert P. Abelson. 1975. Scripts, plans, and knowledge. Yale University.