What Happens Next?
Event Prediction Using a Compositional Neural Network Model

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Mark Granroth-Wilding
mark.granroth-wilding@cl.cam.ac.uk

Stephen Clark
stephen.clark@cl.cam.ac.uk

Computer Laboratory, University of Cambridge, UK

whim
The What-if Machine

Introduction: event prediction

- Knowledge of what **events** are likely in particular contexts is fundamental to human reasoning
- Important for **narrative generation**, piecing together events in narrative sequence.

E.g., *X is arrested* → *X is put in jail*: highly plausible

- Chambers and Jurafsky (2008) extract event knowledge from text
- We build on this to learn a **robust model** of event associations, using **richer information from text**.

Our contributions

1. **Multiple choice narrative close** to evaluate event prediction for models using rich information from text
2. EVENT-COMP neural network model of event chain coherence

Unsupervised event chain learning (C&J08)

- **Coreference resolution** finds repeated mentions of an entity in text (red and blue below).
- Extract **event chains** – sequences of events concerning a particular entity: find verbs with the same entity as an argument.

Royal Welsh appoints new regimental goat Fusilier Llywelyn.

Fusilier Llywelyn was selected from the Royal herd on Llandudno's Great Orme and has been prepared for Army life. He will live at the regiment's base at Lucknow Barracks, Wiltshire, and will accompany them on ceremonial duties.

appoint X → select X → prepare X → X live → X accompany

C&J08 event prediction

Measure association between **context event** and **possible next event** as *pointwise mutual information* (PMI) between predicates.

PMI(X die, X modernize)

Limitations:

1. Only looks at verbs, not other information available in text — *X is put in ~~jail~~*
2. Needs lots of data: poor predictions for rare verbs

Multiple choice narrative cloze (MCNC)

- New evaluation of event prediction models
- Suitable for models that use richer information about events
- Given context of previous events
- Choice of 5 next events
  - Actual next event (**bold**)
  - 4 randomly chosen distractors

x_0 = Giardino x_1 = chairman, him

die(x_0), attend(x_0, reunion), specialize(x_0, as partner), describe(x_0, x_1, as product), hold(x_0, position), appoint(–, x_0, to the board), lead(x_0, effort), ____

c_1: receive(x_0, response)
c_2: drive(x_0, mile)
c_3: seem(x_0)
c_4: discover(x_0, truth)
c_5: **modernize(x_0, procedure)**

?

Models

C&J08:
Count cooccurrence of predicates in chains.
Coherence = PMI of predicates

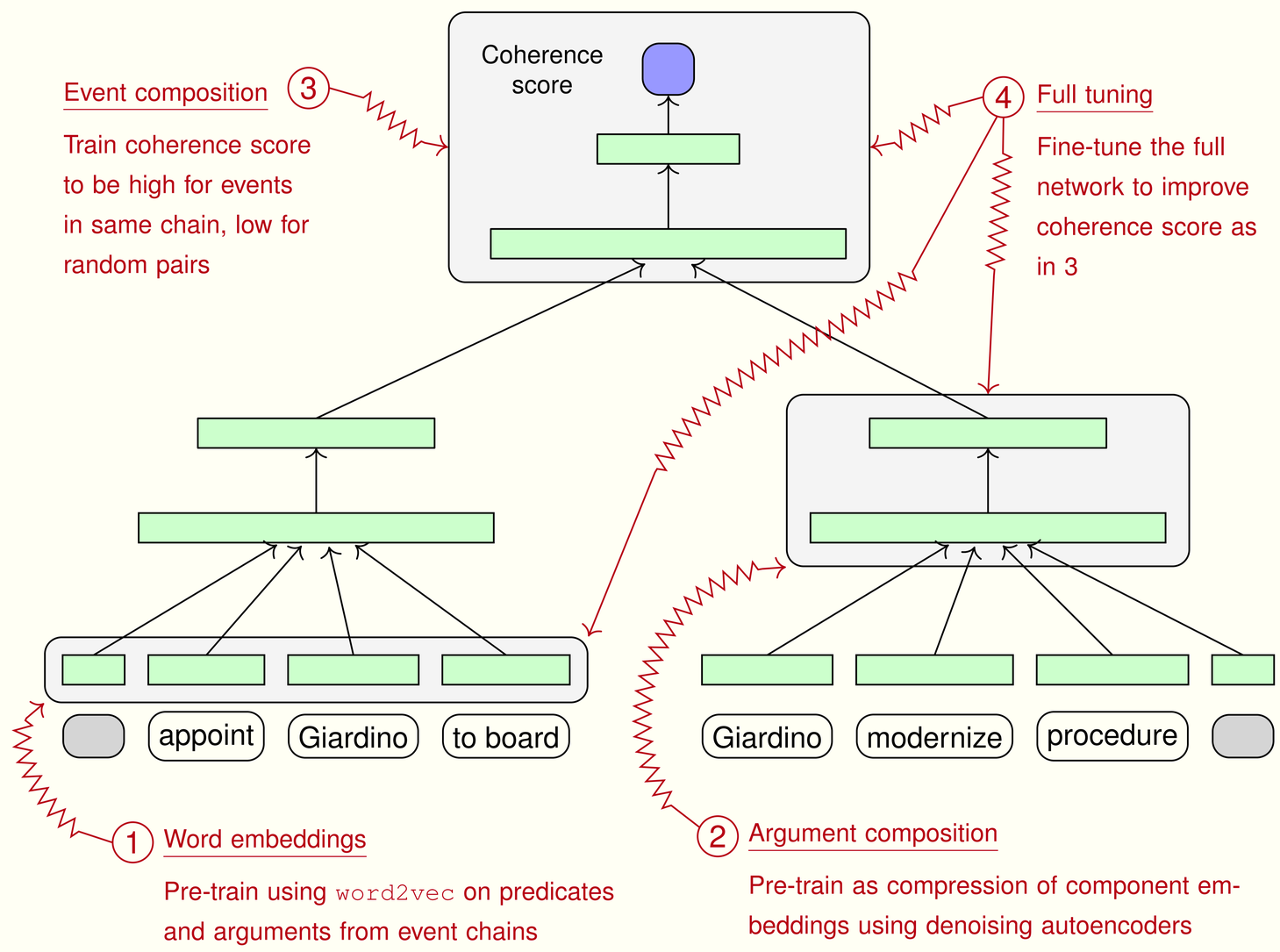
WORD2VEC-PRED:
Learn vectors for predicates so that those that appear in similar chains are close together.
Coherence = vector similarity

WORD2VEC-PRED+ARGS:
Learn vectors for predicates and arguments from contexts of predicates and arguments in chain. Sum components to get event vector.
Coherence = similarity of event vectors

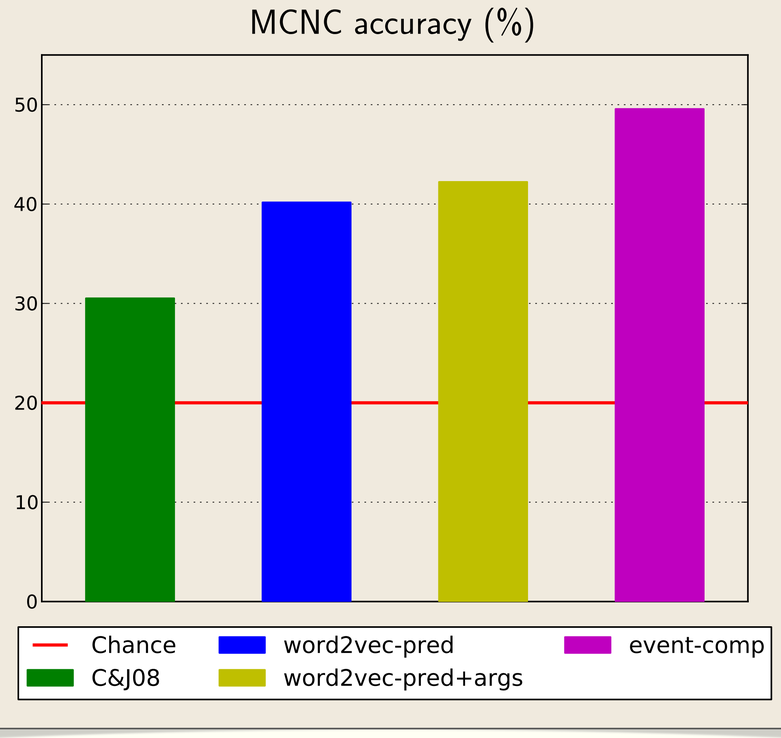
EVENT-COMP:
Simultaneously learn vectors for predicates and arguments, function to compose them, and coherence function as neural network.
See right ⇒

EVENT-COMP neural network

Word embeddings: represent predicate and argument words as vectors
Argument composition: compose embeddings into event vector
Event composition: predict whether two event vectors come from the same chain



Results



Conclusions

- Simple C&J08 event association measure performs well above chance on MCNC
- Learning vector representations of predicates gives a more robust model
- Simple inclusion of argument information (WORD2VEC-PRED+ARGS) helps on MCNC
- EVENT-COMP learns better event association model

Chambers, N., and Jurafsky, D. 2008. Unsupervised learning of narrative event chains. In *Proceedings of ACL*, 789–797. Association for Computational Linguistics.