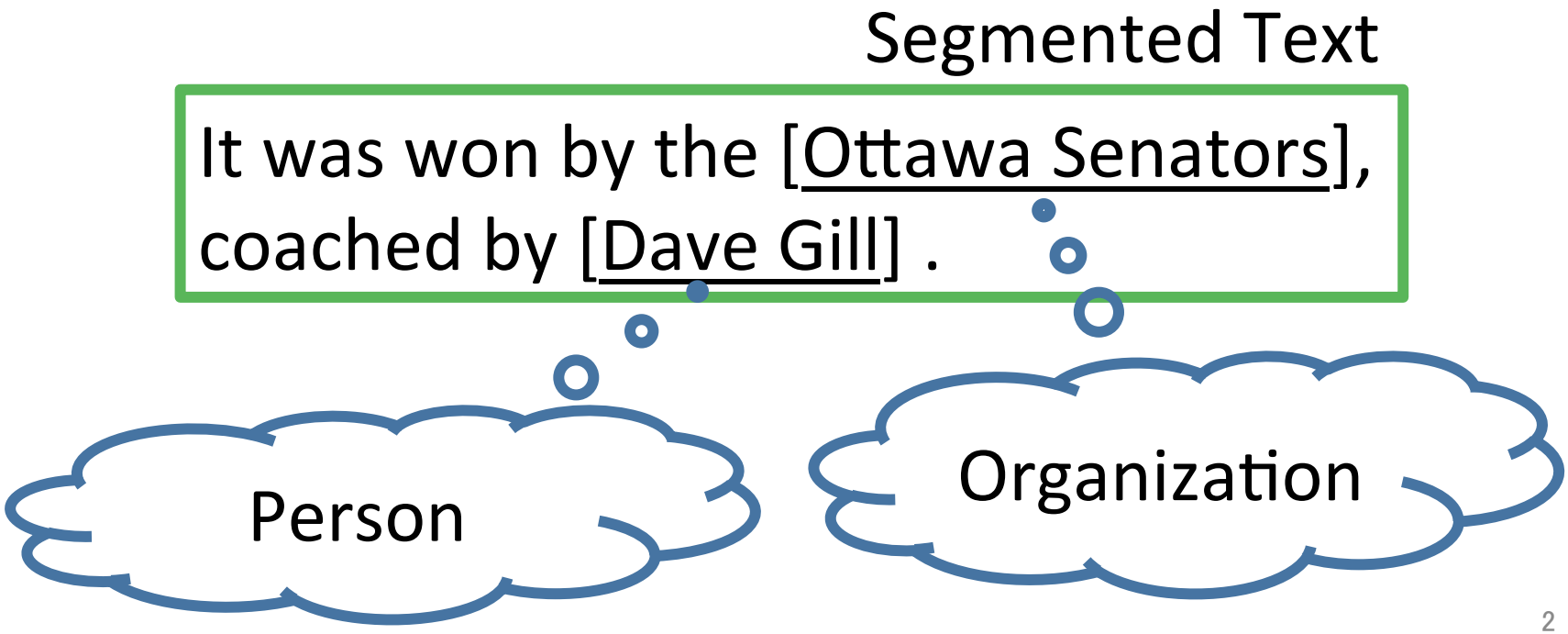
An Attentive Neural Architecture for Fine-grained Entity Type Classification

Sonse Shimaoka Pontus Stenetorp

Kentaro Inui Sebastian Riedel

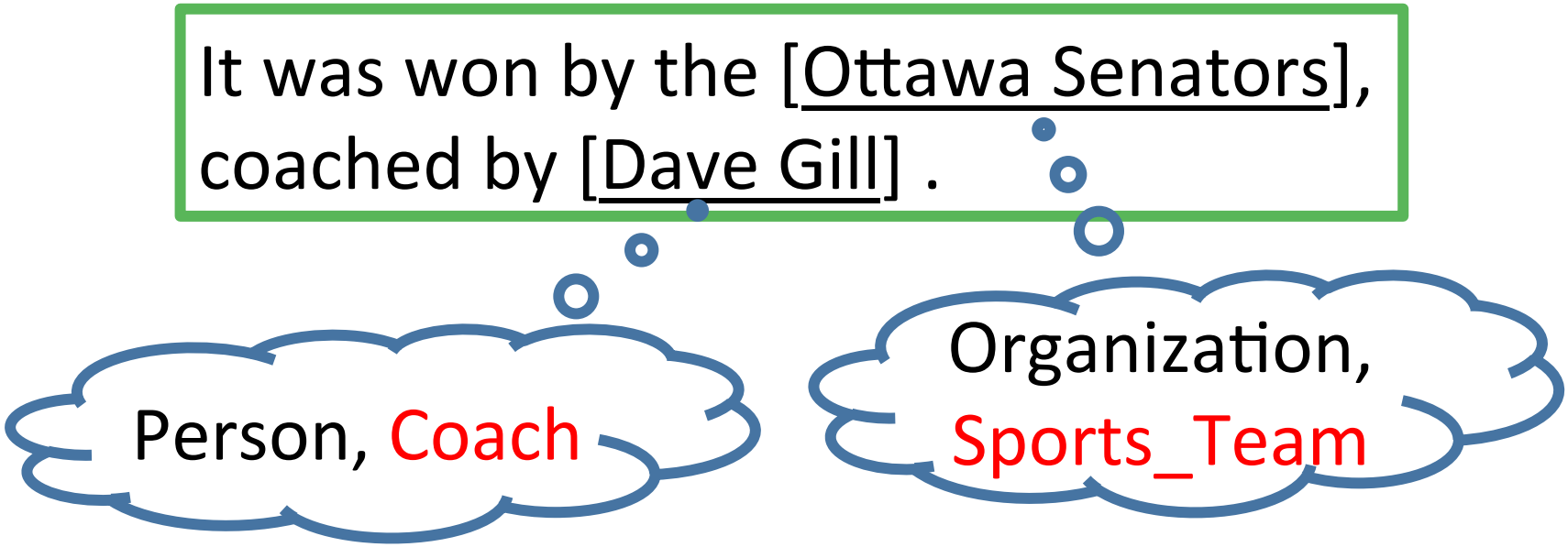
Entity Type Classification

- Assigning semantic types to entity mentions
- Sentences are segmented beforehand



Fine-grained entity type classification

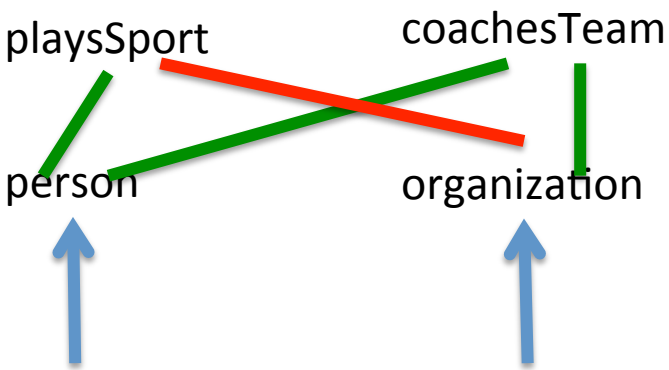
- A fine-grained set of types
  - [Ling and Weld AAAI'12], [Yosef+ ACL'12], [Gillick+ arXiv'14], [Yogatama+ ACL'15]



Applications

- Useful for Relation Extraction

Imposing constraints based on entity types [Carlson + , AAAI'10]



Using entity types as input features of a relation classifier [Ling and Weld , AAAI'12]

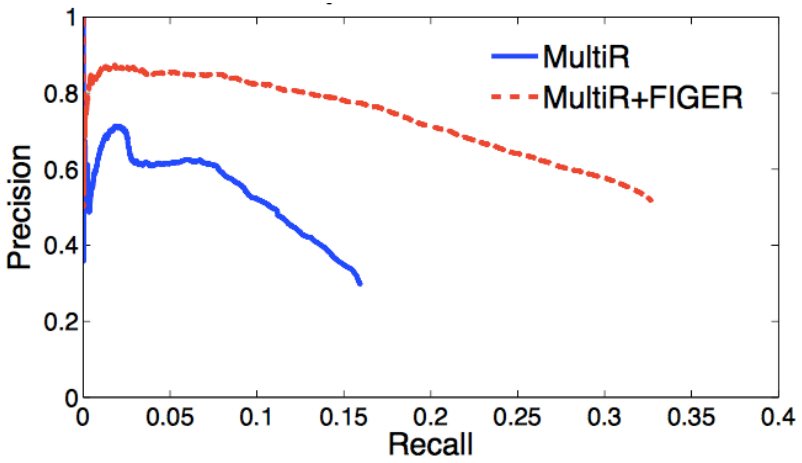


Figure 3: Precision / Recall curves for relation extraction.

Importance of Context

- Some labels are impossible to correctly predict *without context* [Gillick+ arXiv'14]

A match against [New Zealand] is held on Monday

Claim

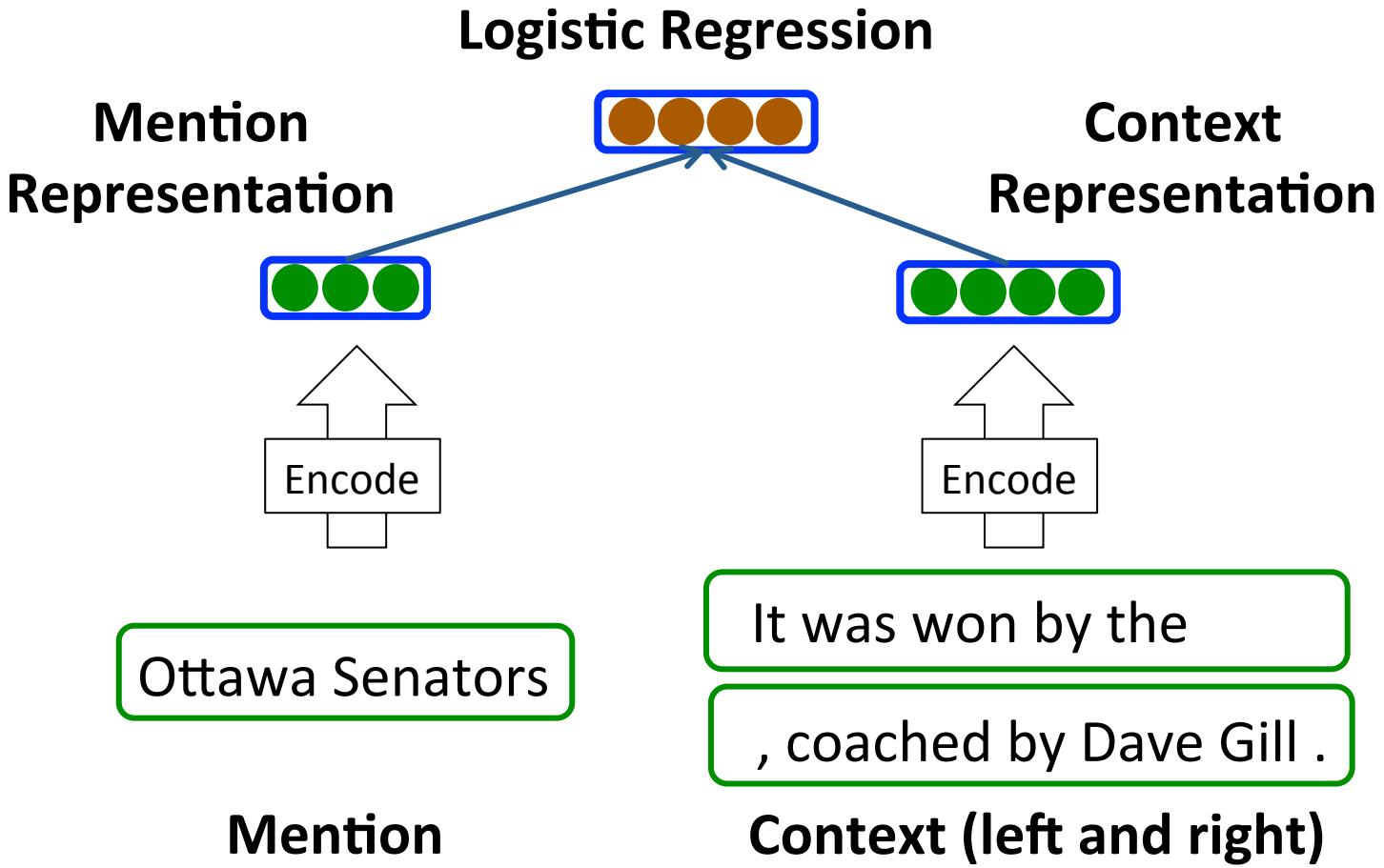
- Previous Models have been too simplistic when handling contextual information
  - Linear Classifiers with sparse, hand-crafted features [Ling+12], [Yosef+12],[Gillick+14], [Yogatama+15]

Claim

- Previous Models have been too simplistic when handling contextual information
  - Linear Classifiers with sparse, hand-crafted features [Ling+12], [Yosef+12],[Gillick+14], [Yogatama+15]
- To address this limitation,
  - We apply RNNs to model context
  - We introduce a novel attention mechanism to encourage the model to focus on important expressions

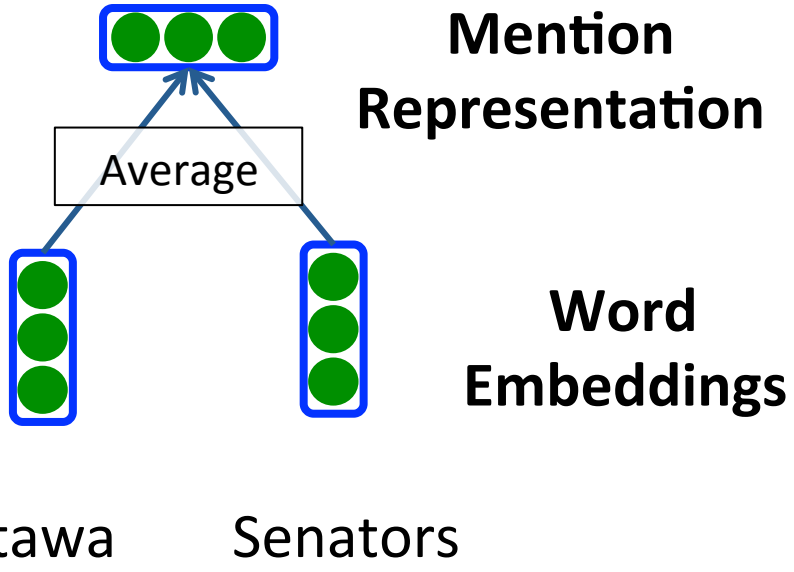
General Model

- Separately Encoding Mention and Context



Mention Representation

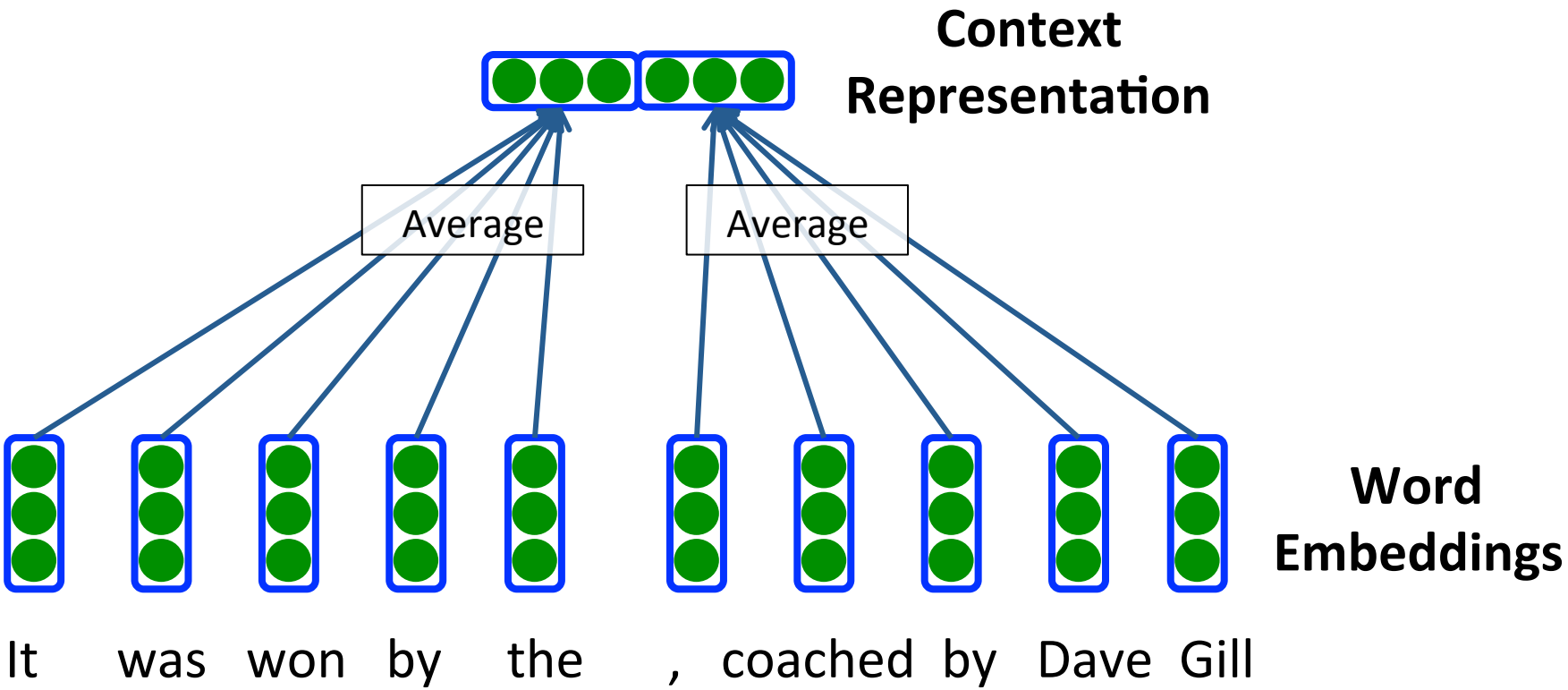
- Averaging embeddings of mention words
- More complex models tend to overfit



Context Representation①

Averaging Encoder

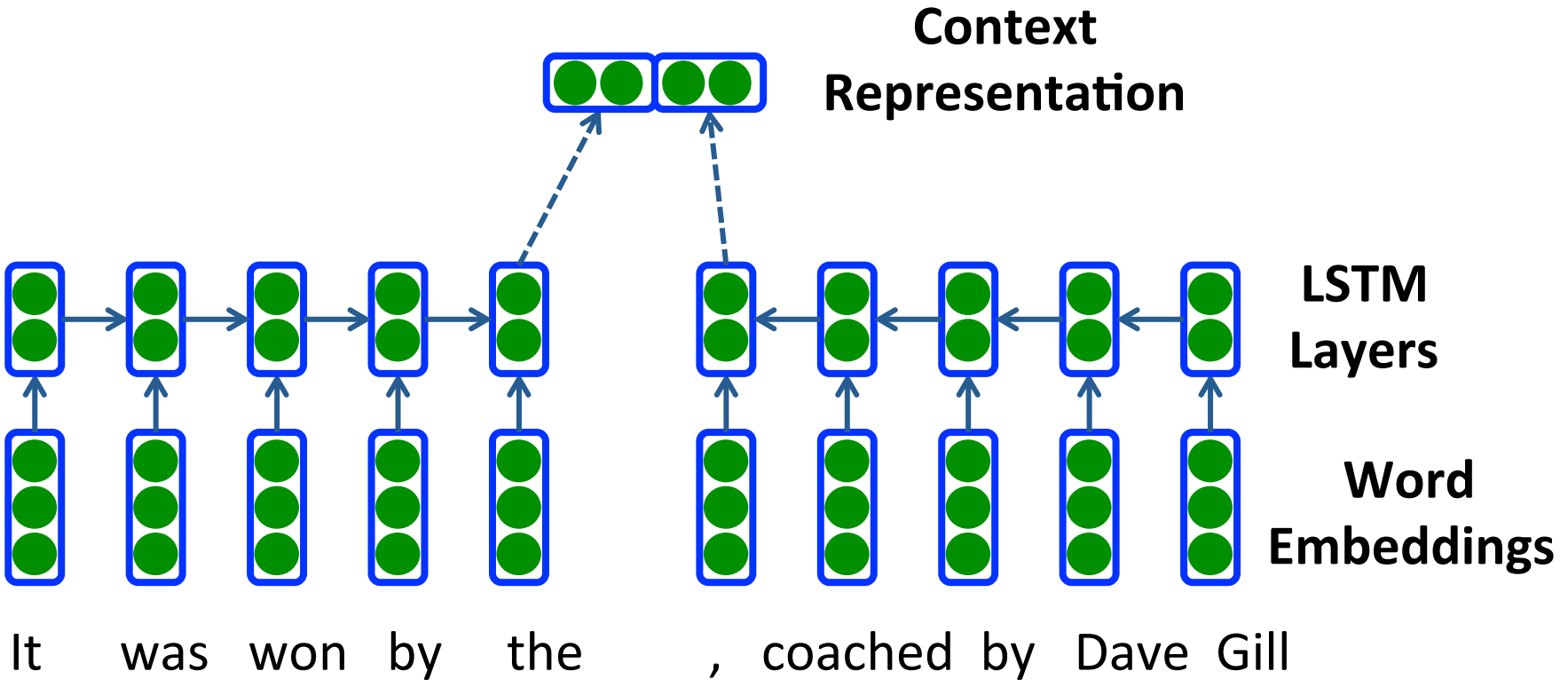
- Averaging left and right context embeddings



Context Representation②

LSTM Encoder

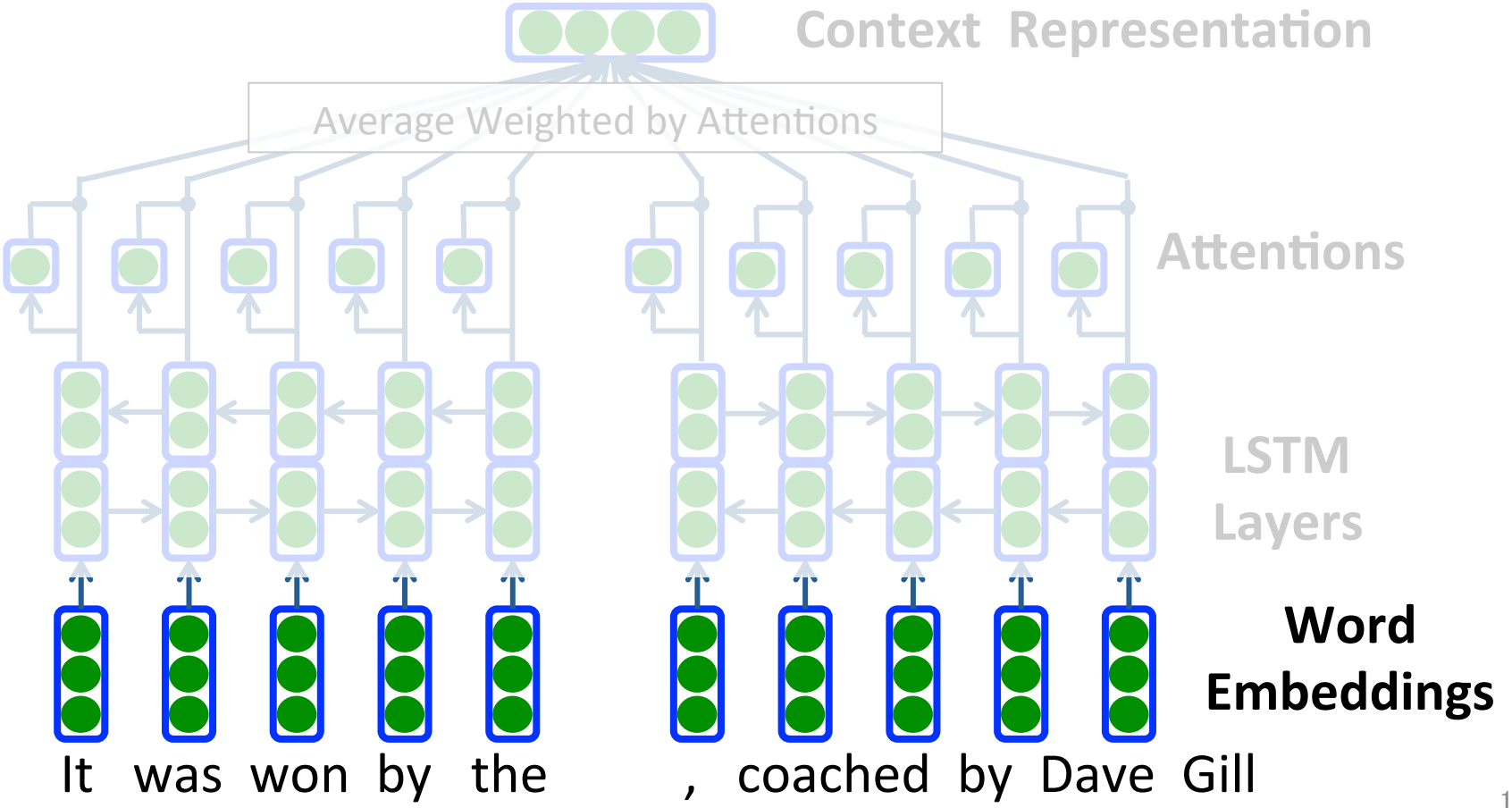
- Recurrently compose left and right context



Context Representation③

Attentive Encoder

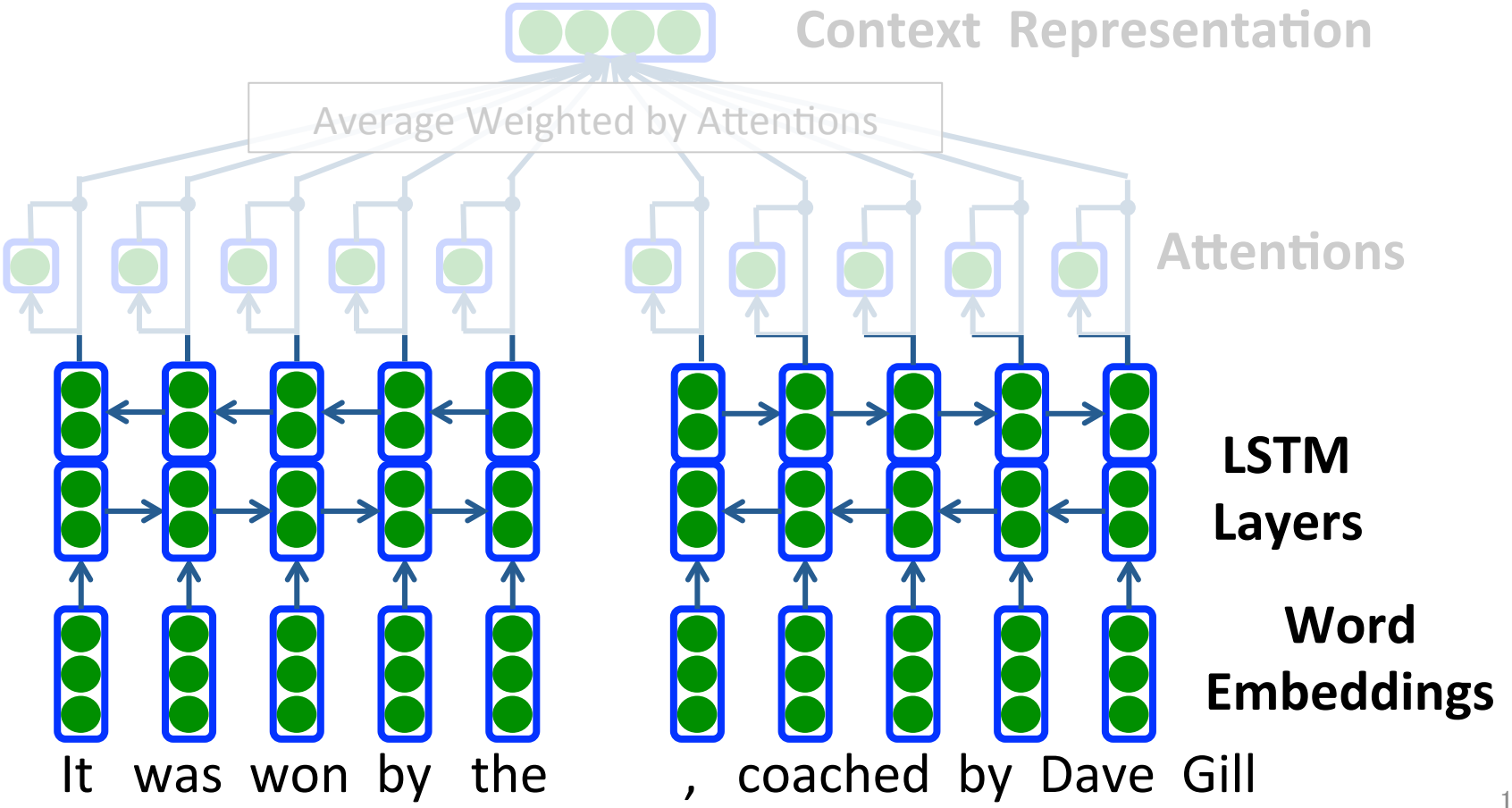
- Attention modules on top of bi-LSTM layers



Context Representation③

Attentive Encoder

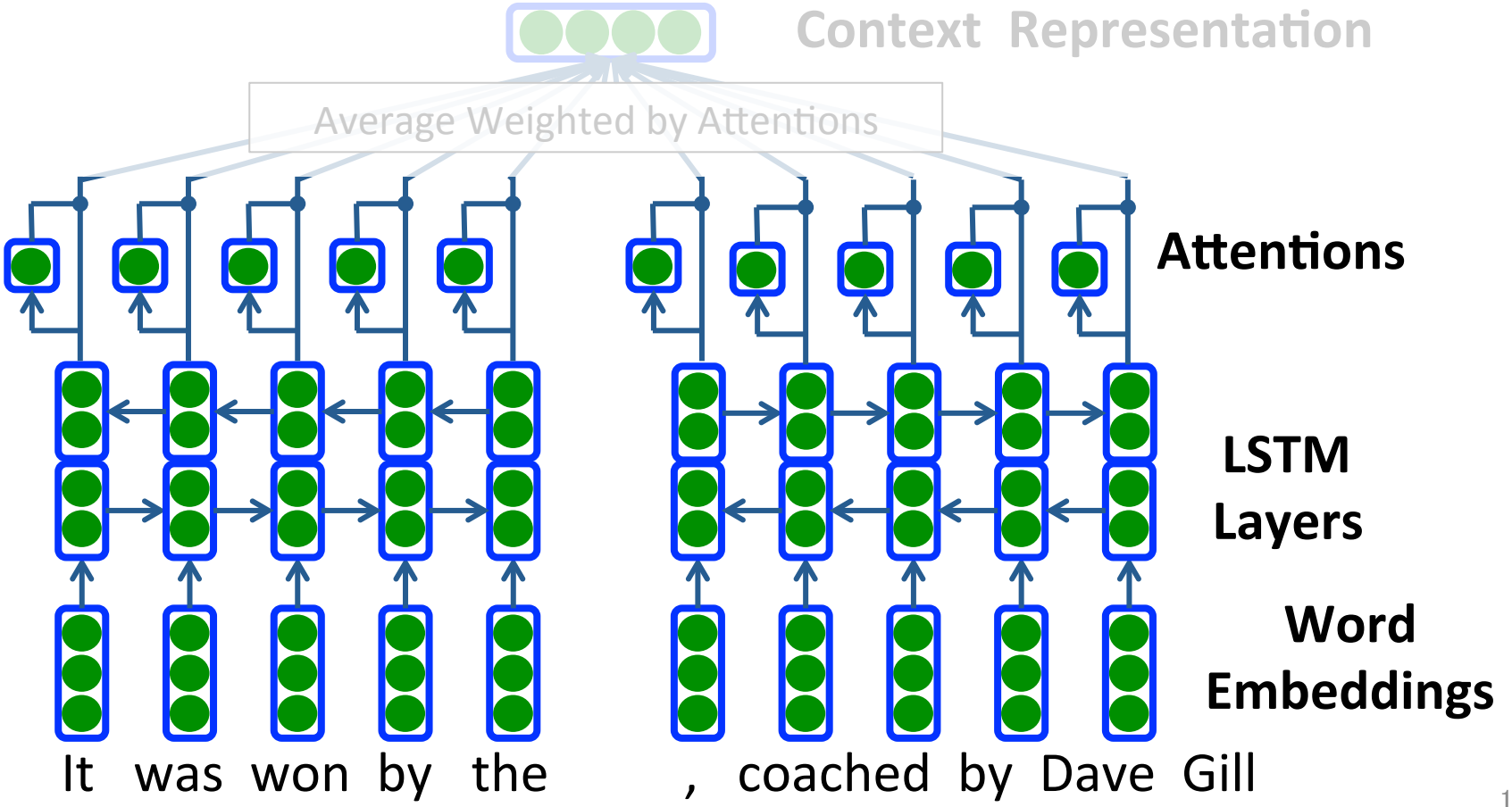
- Attention modules on top of bi-LSTM layers



Context Representation③

Attentive Encoder

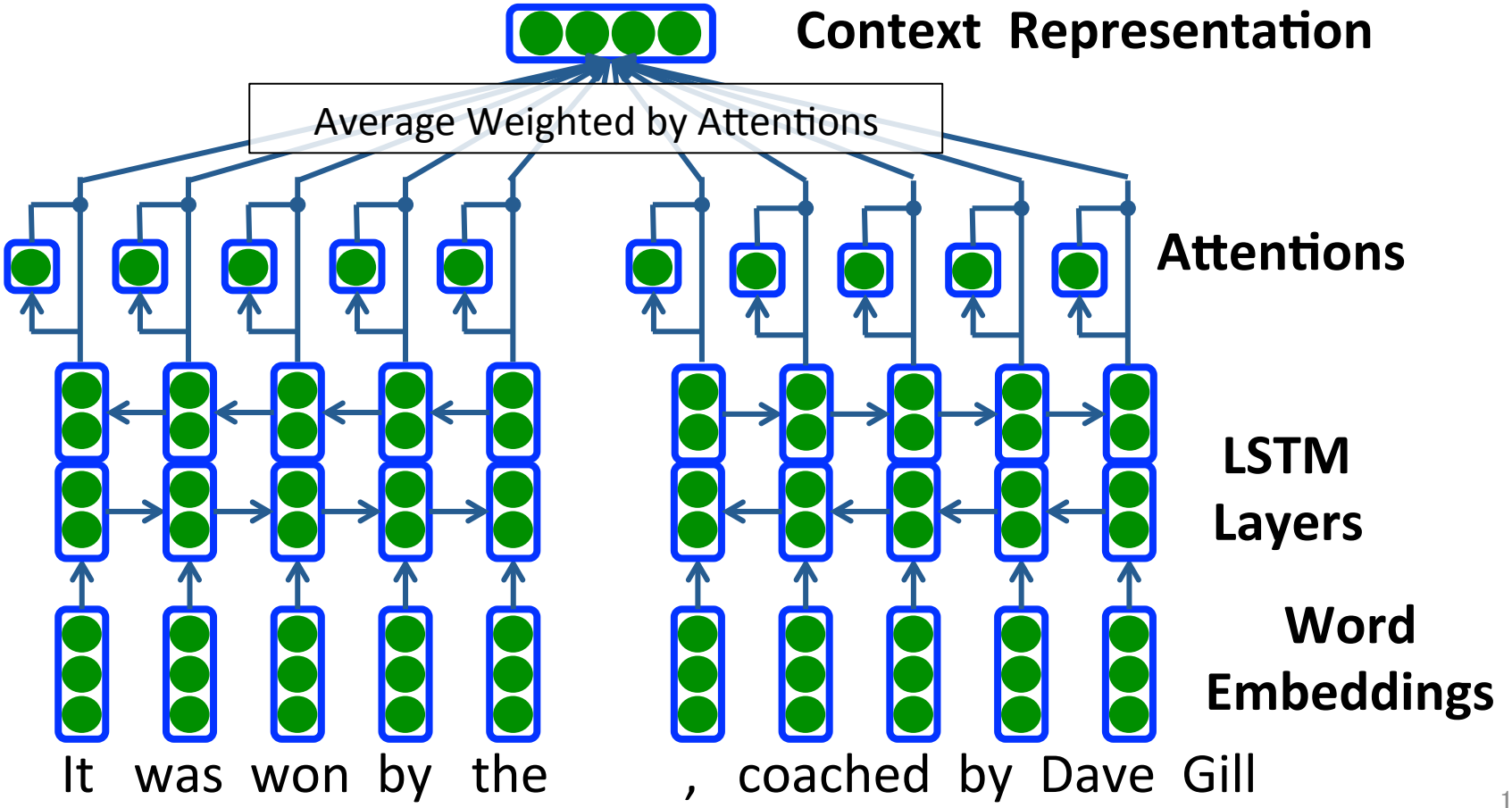
- Attention modules on top of bi-LSTM layers



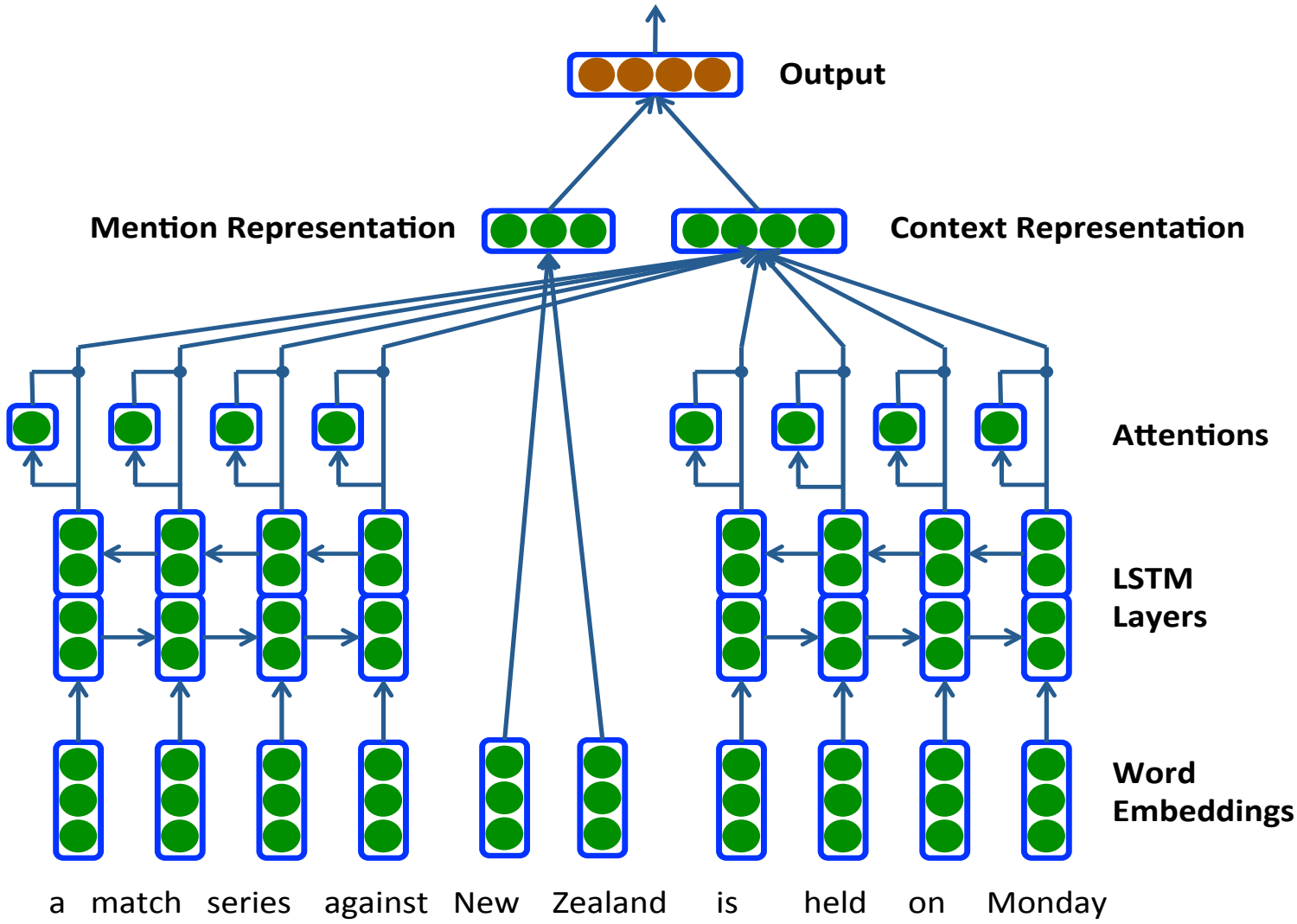
Context Representation③

Attentive Encoder

- Attention modules on top of bi-LSTM layers



Proposed Architecture

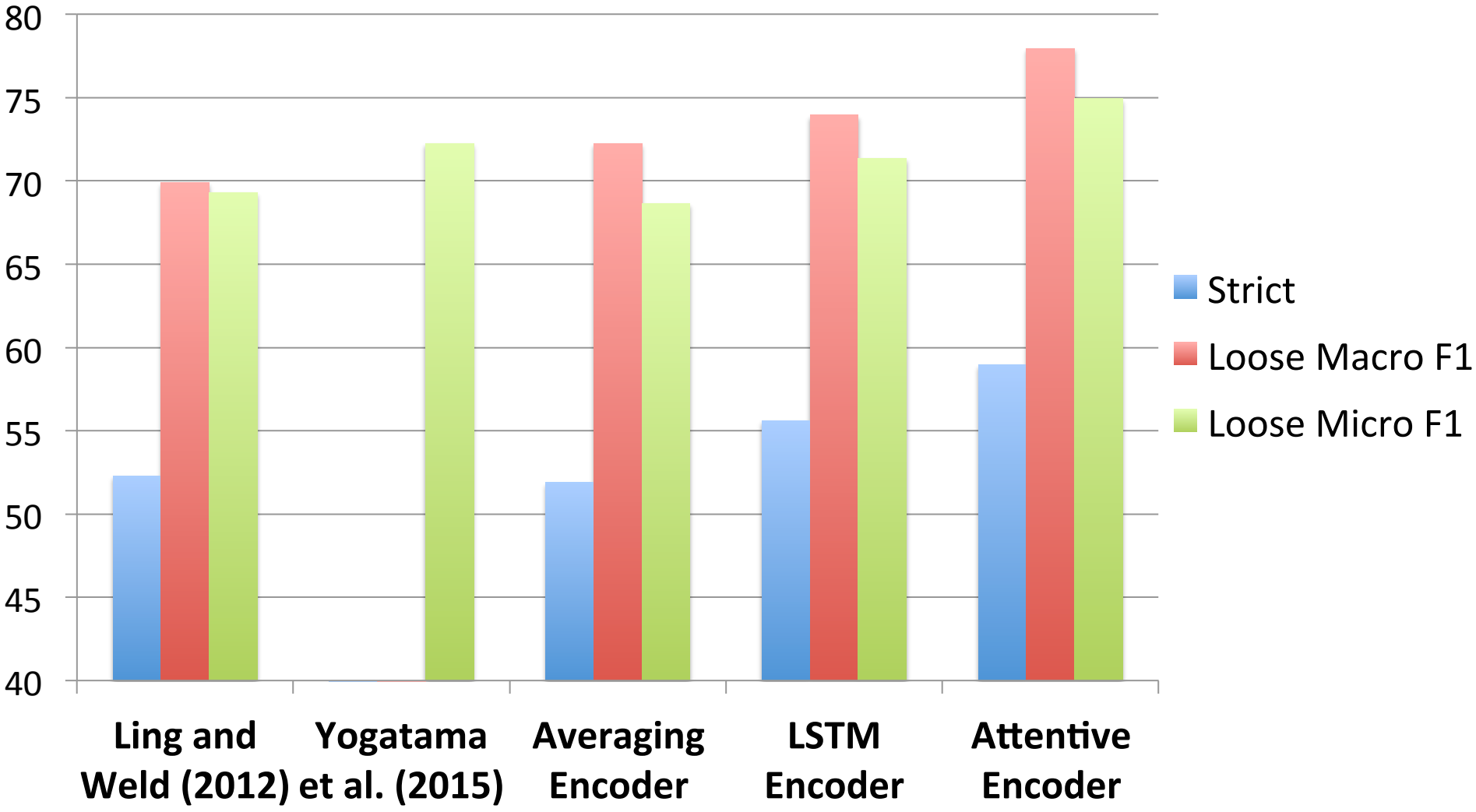


Experiment

- FIGER Type System [Ling and Weld AAAI'12]
  - 112 types from Freebase

- FIGER dataset [Ling and Weld AAAI'12]
  - Train: 2.6 Million mentions from Wikipedia
  - Test: 563 manually annotated mentions

Results


80
75
70
65
60
55
50
45
40
Strict
Loose Macro F1
Loose Micro F1
Ling and Weld (2012)
Yogatama et al. (2015)
Averaging Encoder
LSTM Encoder
Attentive Encoder

Attention Visualization

British comedy film starring Googie Withers , Tyrell Davis and [Rex Harrison] .

➡ /person, /person/actor

He returned to the Riverina in 1913 and died of [endocarditis] in Hay , survived by his daughter .

➡ /disease

Hall in Amsterdam , Netherlands for a live DVD titled [Live from Amsterdam] .

➡ /music

the 1947-48 season to play a five-match Test series against [Australia] .

➡ /organization, /organization/sports_team

Conclusion

- **State of the art** with 74.94% loose micro F1

  2.59% relative improvement

- First work to apply RNN to model context

- A novel attention mechanism
  - Success in learning to attend over informative expressions

Extensions

- Recently, we have published the extension of this work [**arXiv:1604.05525]** !

  Neural Architectures for Fine-grained Entity Type Classification

  - **New state of the art**
  - **Hierarchical label encoding**
  - **Quantitative analysis of the attention mechanism, relating it to hand-crafted features**