

Embedding time expressions for deep temporal models



Tanya Goyal and Greg Durrett

ACL

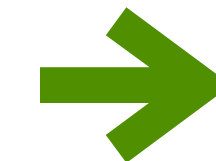
July 31, 2019



Temporal Ordering

Former Singapore premier Lee Kuan Yew **stepped** down in **1990** in favour of the incumbent, Prime Minister Goh Chok Tong, but **remains** influential in the new government. Lee **came** to power in **November 1959**...

November 1959



1990

came



stepped

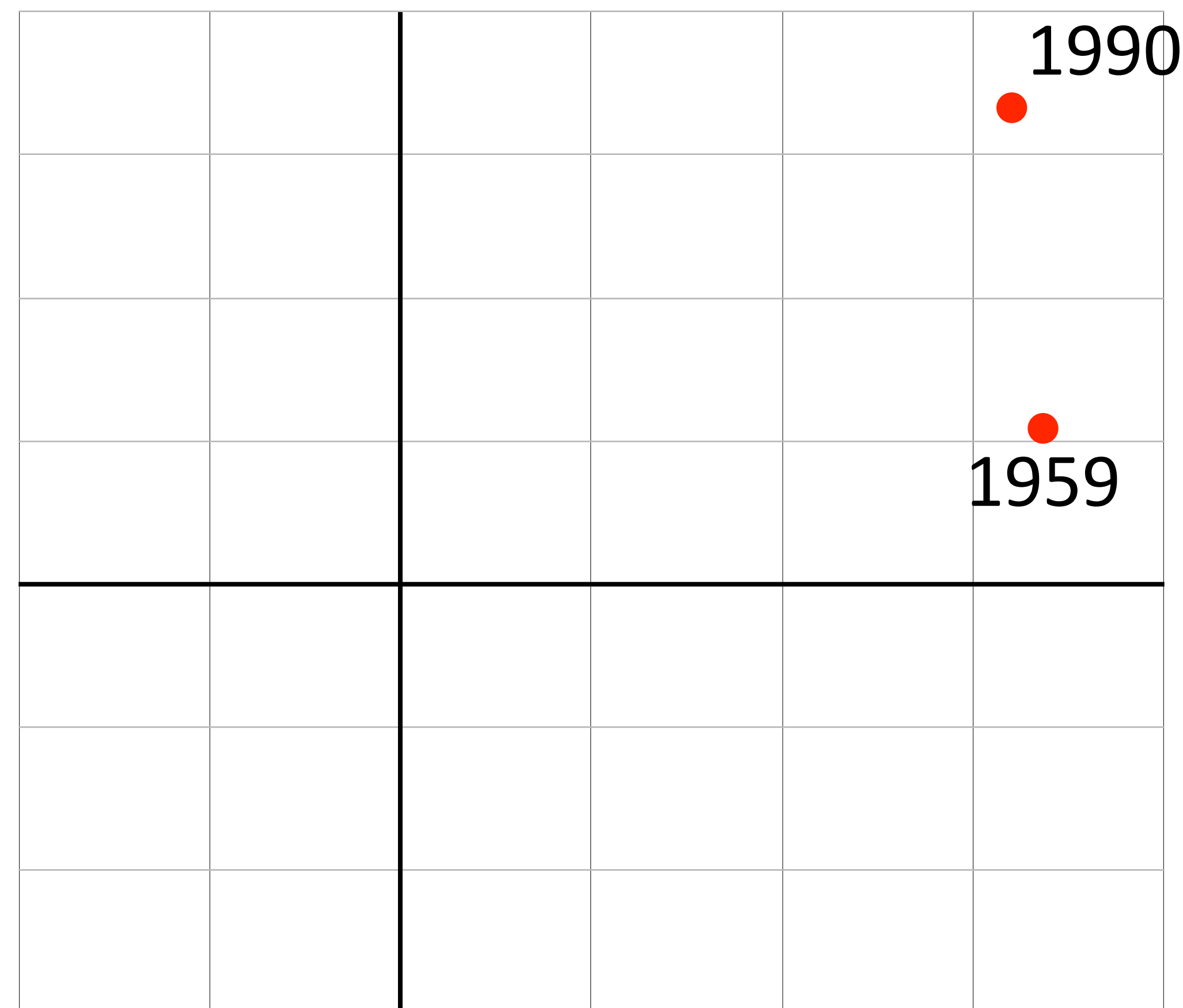


remains



Pre-trained embeddings encode time?

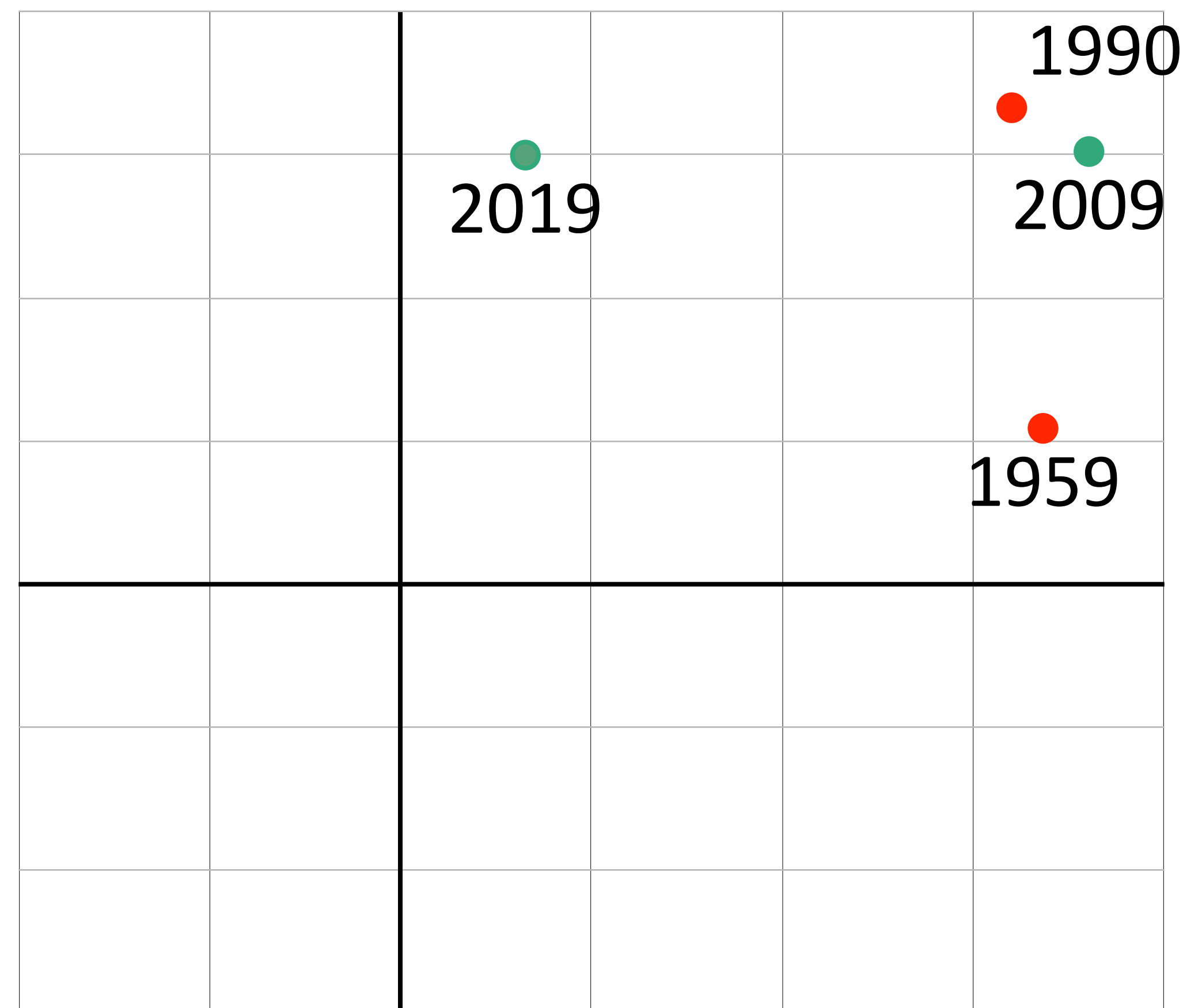
Former Singapore premier Lee Kuan Yew stepped down in 1990 in favour of the incumbent, Prime Minister Got Chok Tong, but remains influential in the new government. He came to power in November 1959...





Pre-trained embeddings encode time?

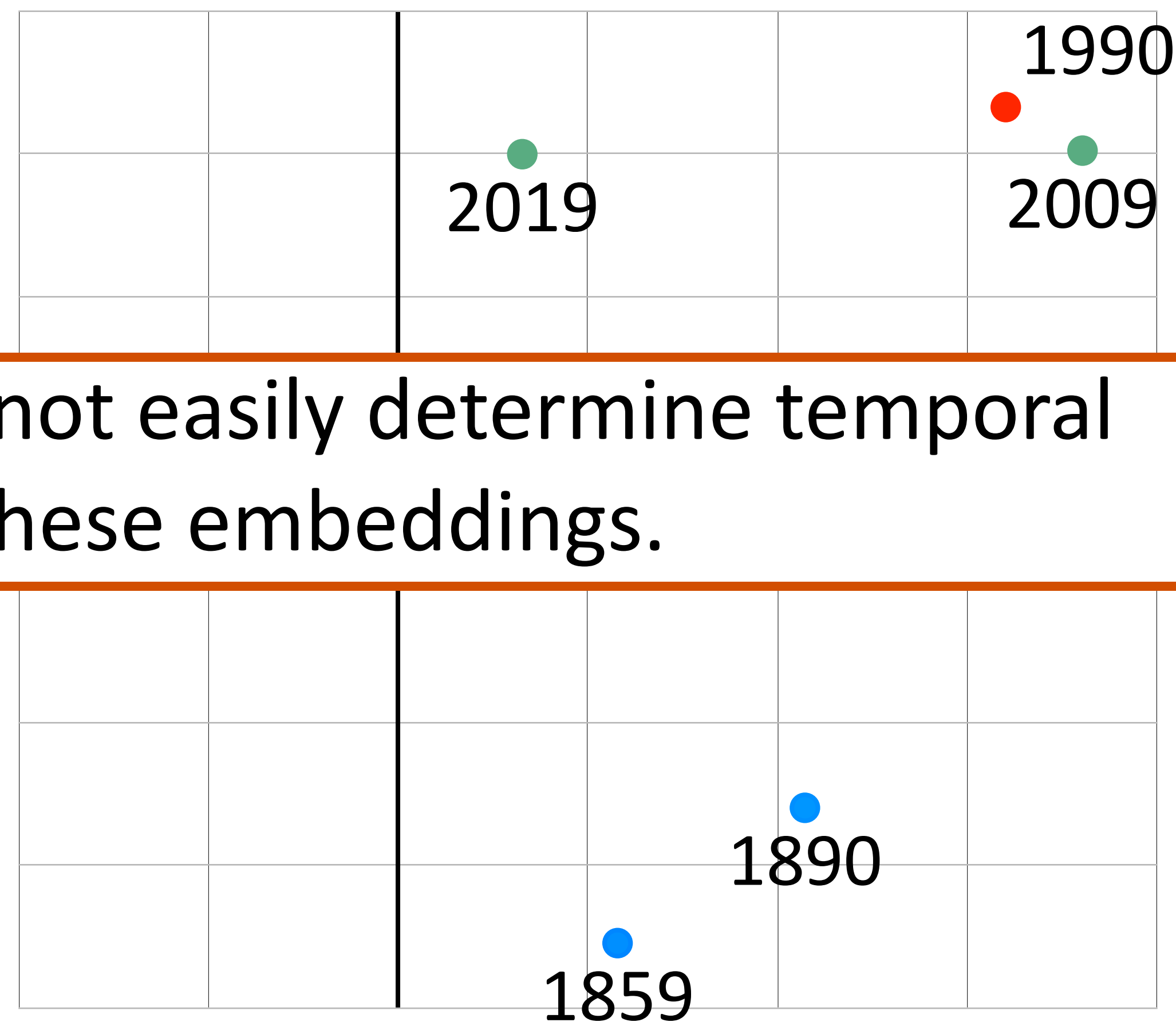
Former Singapore premier Lee Kuan Yew stepped down in 2019 in favour of the incumbent, Prime Minister Got Chok Tong, but remains influential in the new government. He came to power in November 2009...





Pre-trained embeddings encode time?

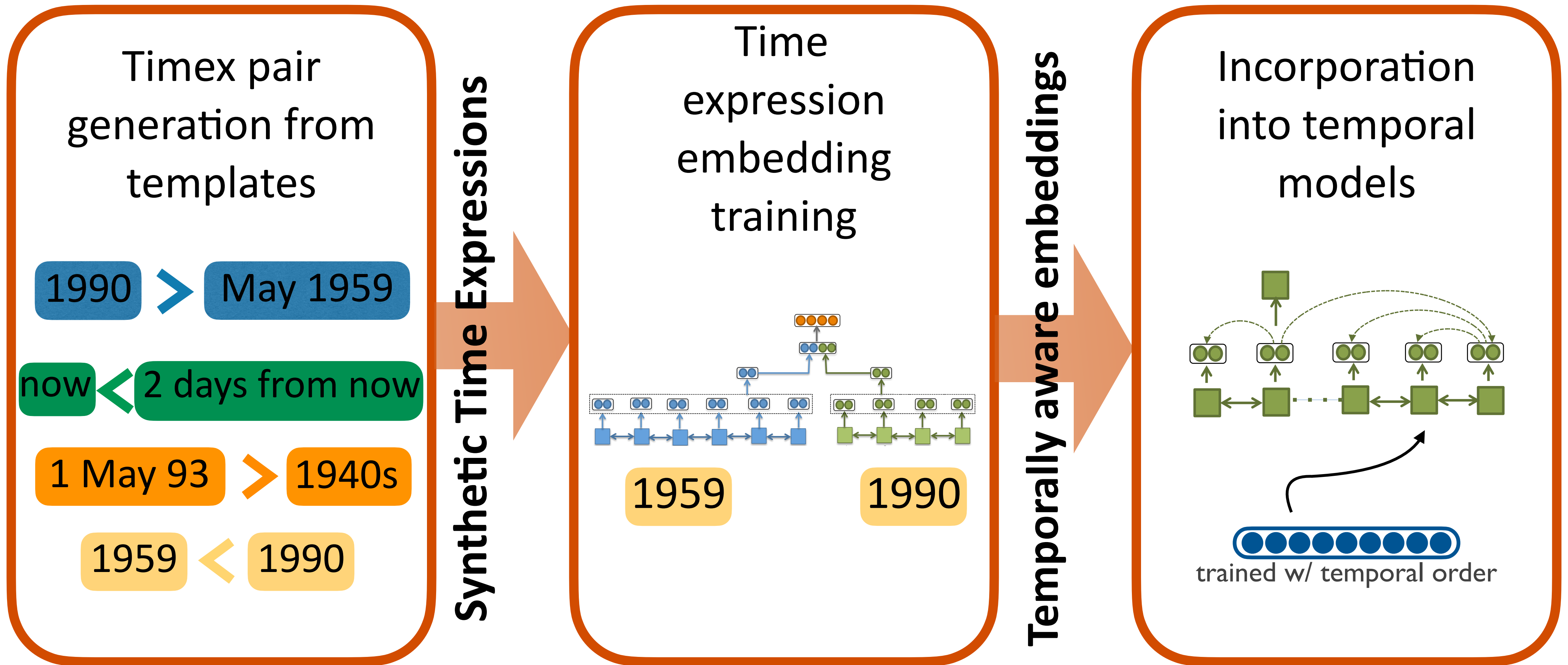
Former Singapore premier
Lee Kuan Yew stepped down
in 1890 in favour of the
incumbent.
Got C
influential in the new
government. He came to
power in November 1859...



Downstream models cannot easily determine temporal status from these embeddings.



Solution: Temporally-aware embeddings





Timex Pair Generation

Goal: Construct embeddings for time expressions that are capable of inferring temporal relation between them.

Synthetic Data Generation:

- Timex Templates
 - Explicit datetime templates
[yyyy], ['yy], [mm dd yy], [mm yy] ...
 - Natural language indicators
[xx units later], [xx units from now], [now]...
- Sample valid values for slots
- Assign temporal label (before, after, simultaneous)

mmm yy

yyyy

May 93

1995

BEFORE

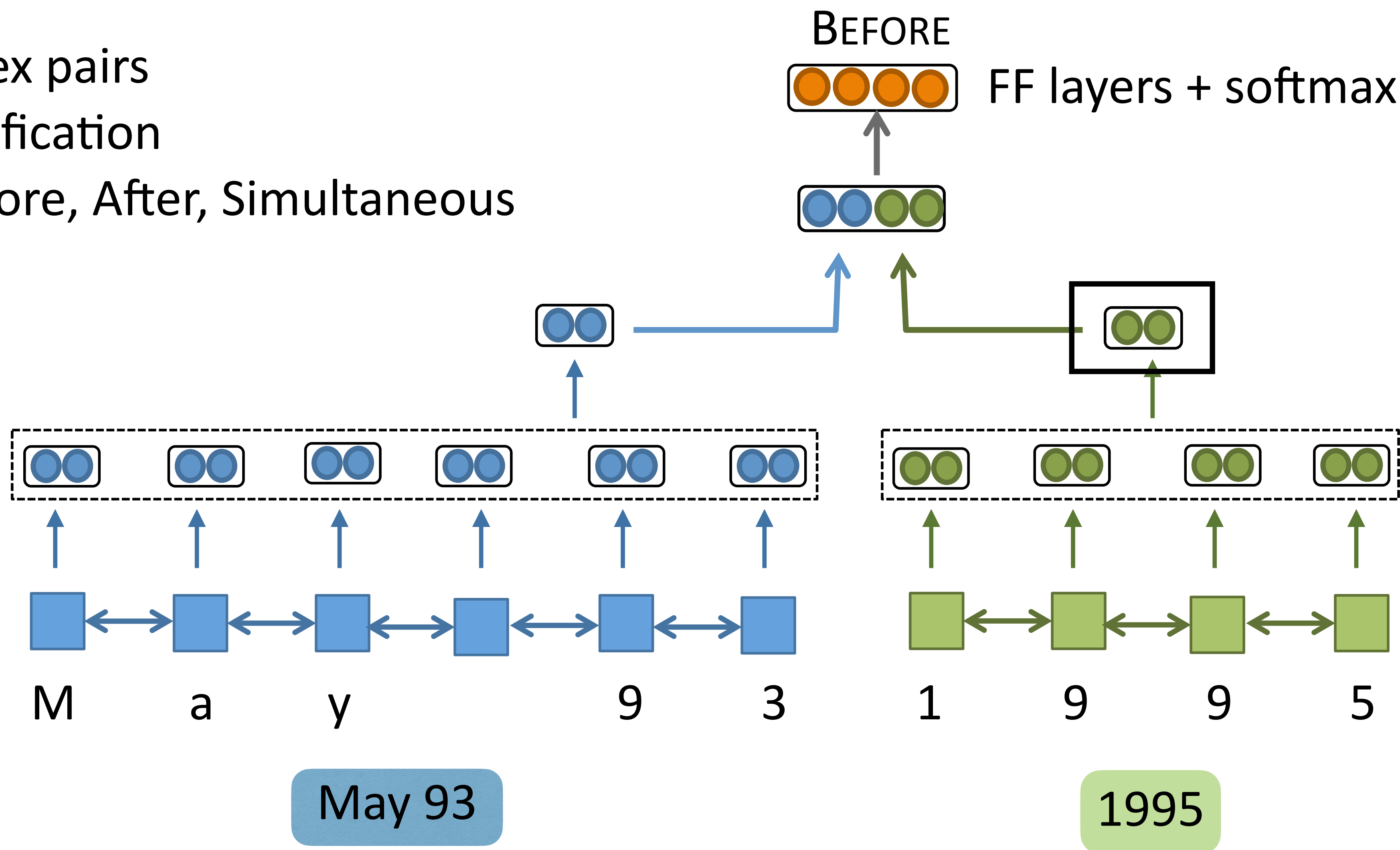


Temporally-aware embeddings

Input: Timex pairs

Task: Classification

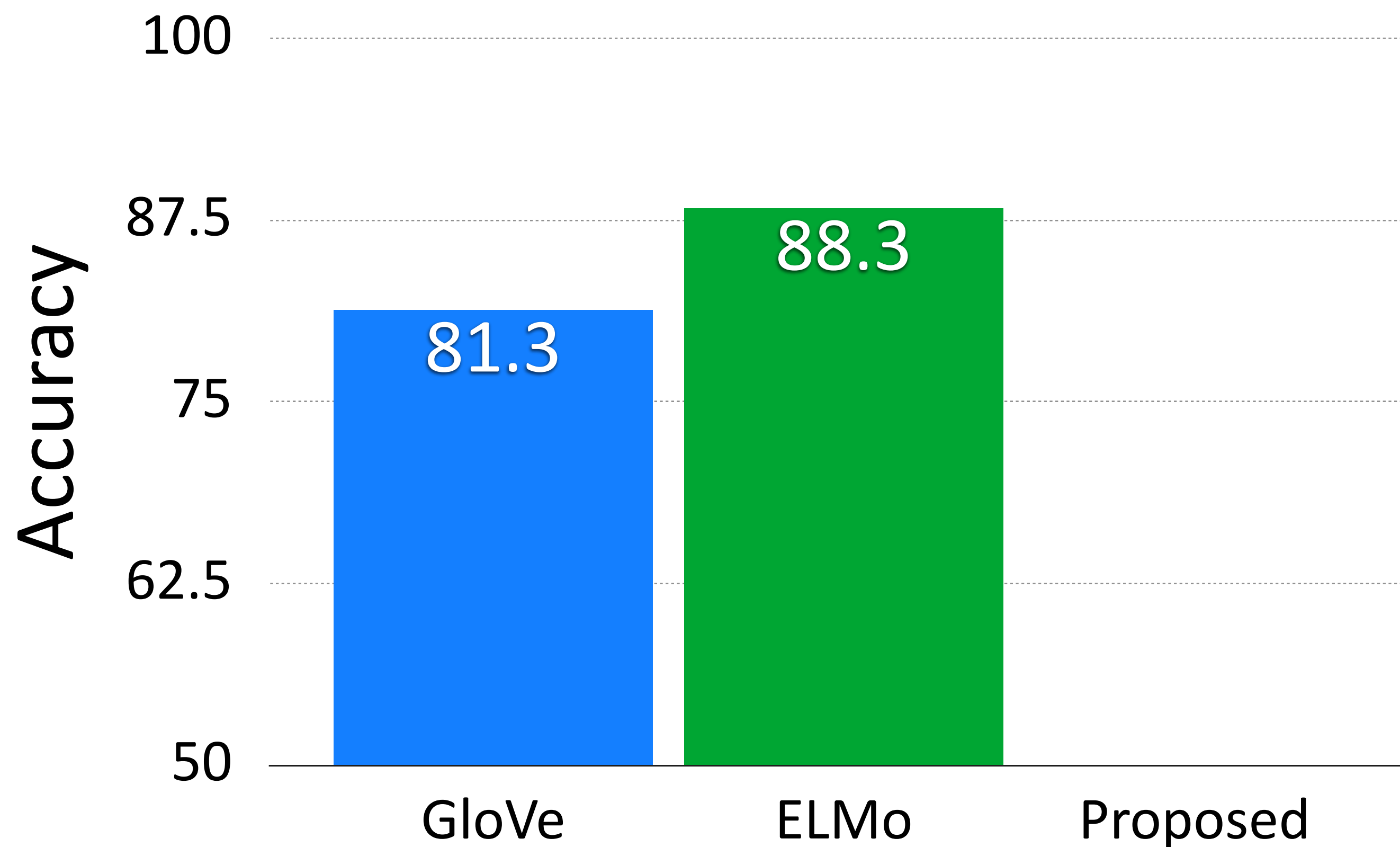
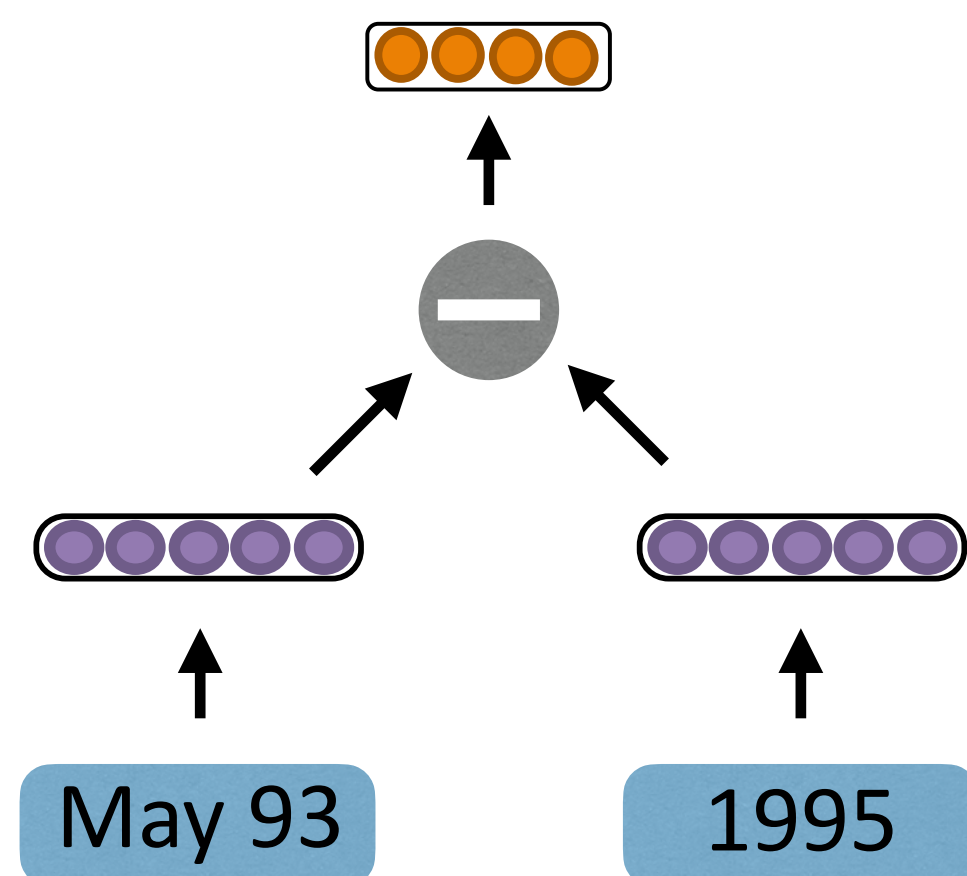
Labels: Before, After, Simultaneous





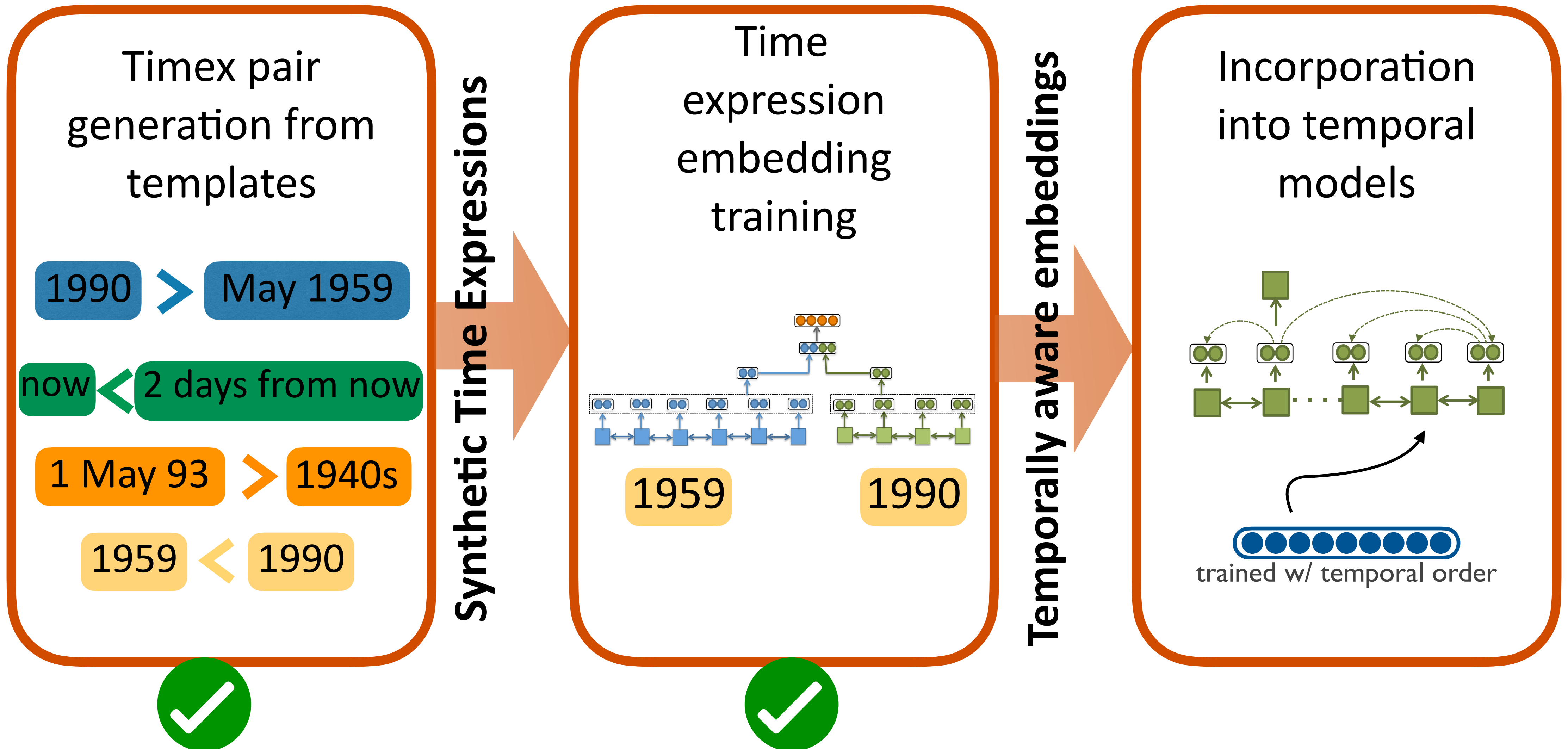
Evaluation: Timex Ordering

- ▶ Do pre-trained embeddings capture temporal info?
- ▶ Do proposed timex embeddings capture temporal info?



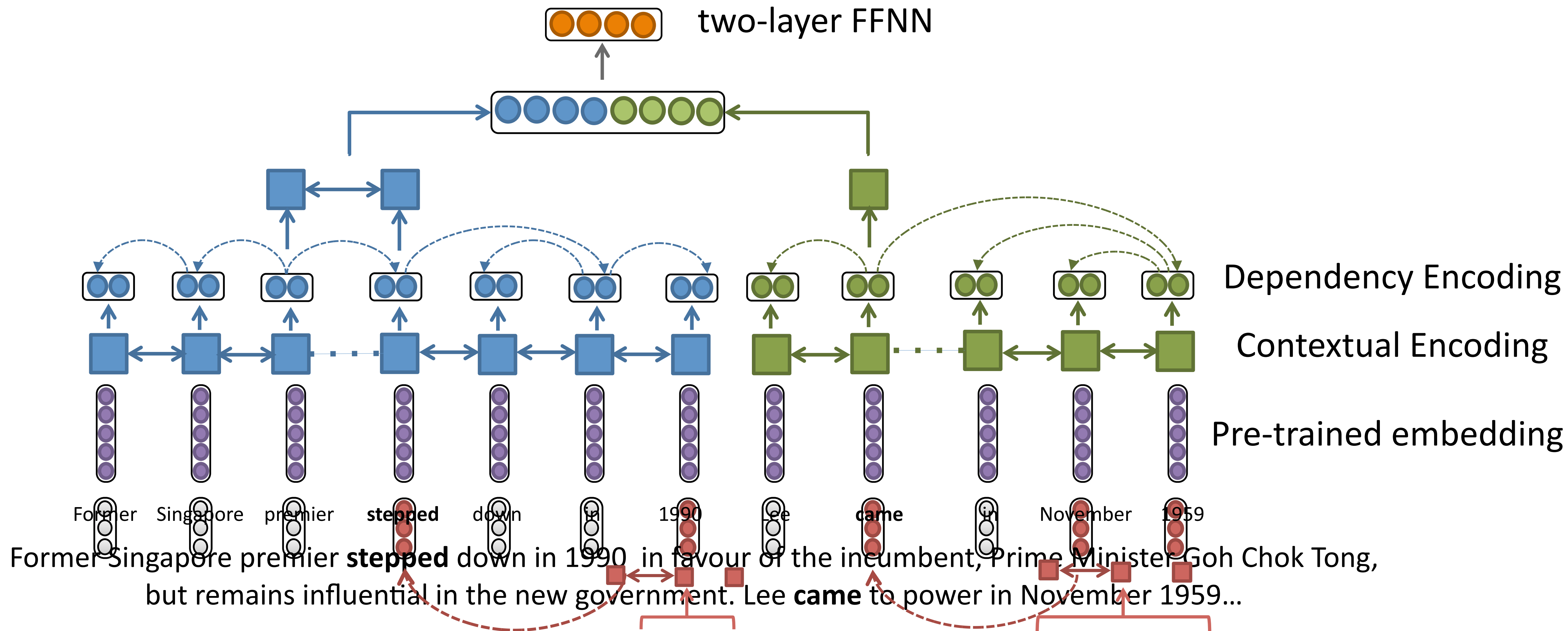


Solution: Temporally-aware embeddings





Incorporation in temporal models





Evaluation: MATRES

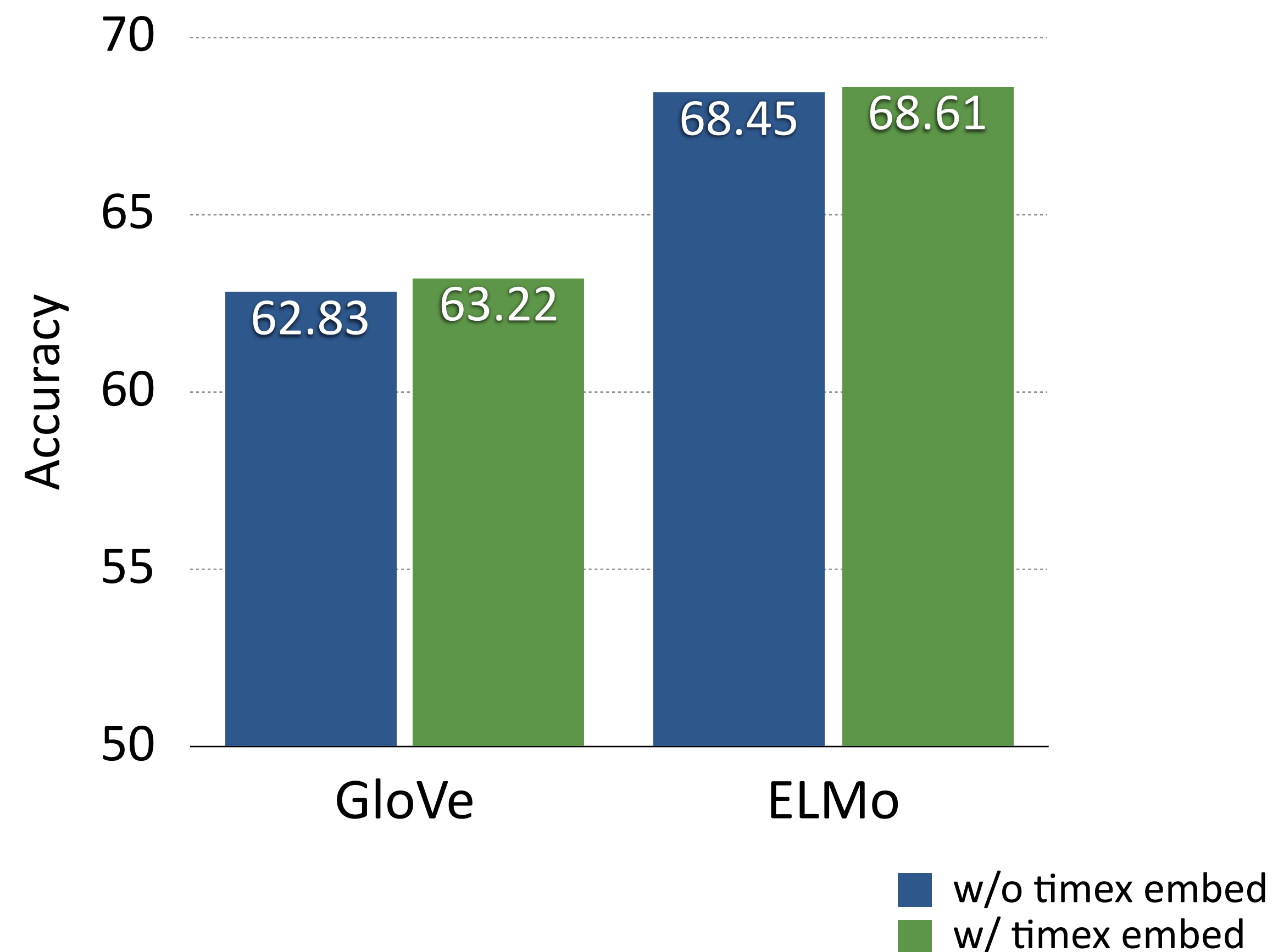
Data:

256 newswire articles

Event pairs in adjacent sentences are assigned temporal label.

Labels: Before, After, Simultaneous, Vague

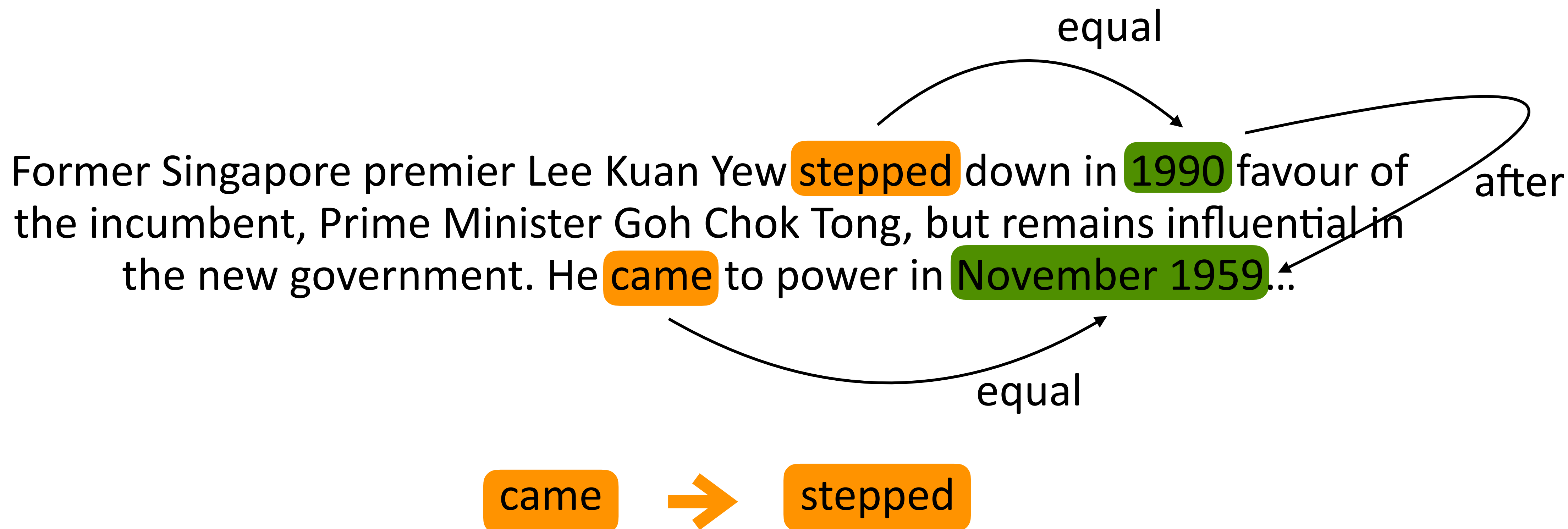
- ▶ Only a fraction of examples in the dataset contain distinct timex expressions pairs.
- ▶ Temporal relation resolved by the baseline model without the help of time expressions.





Evaluation: Distant Data

Data Collection through rule based classifiers

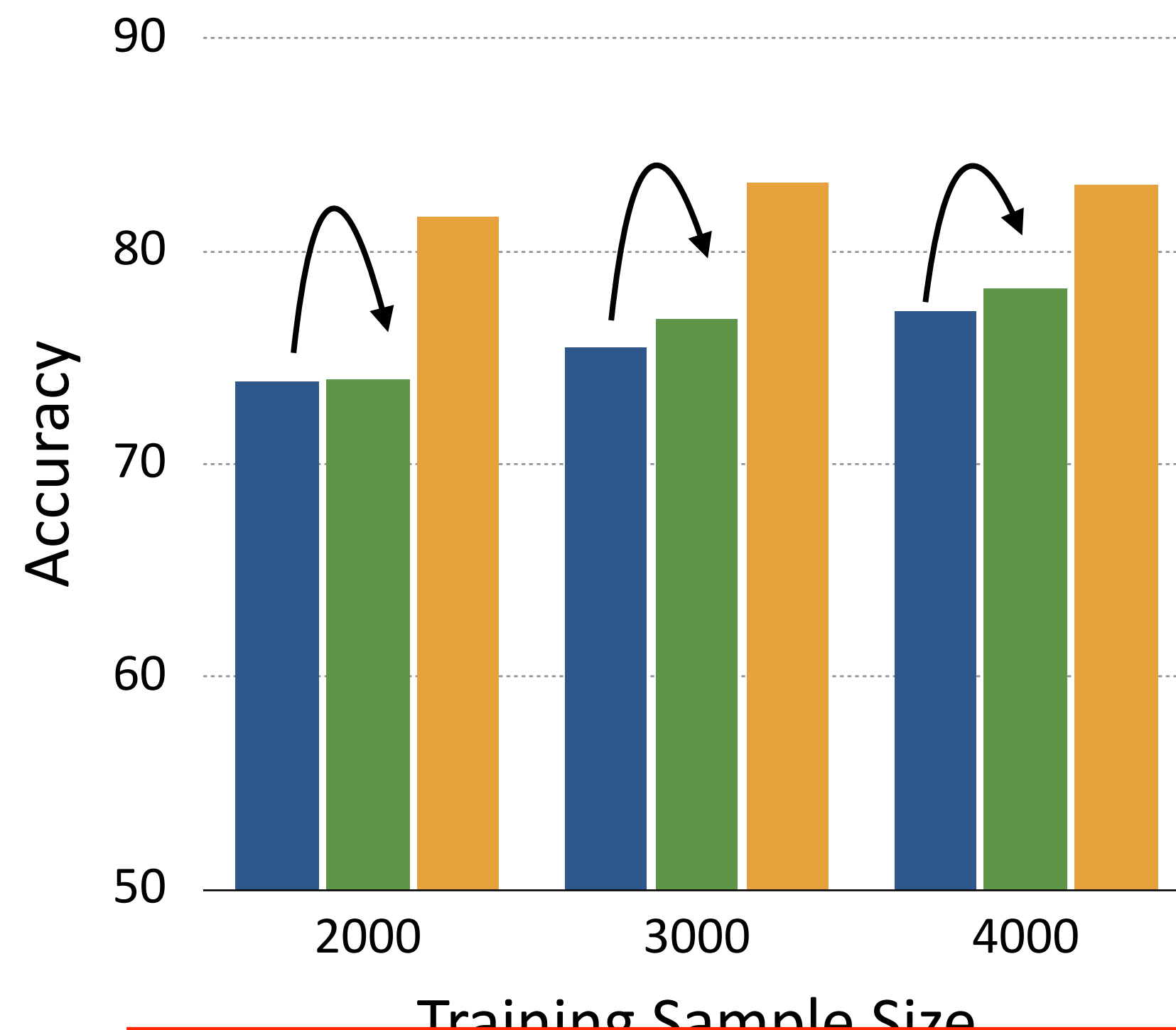




Evaluation: Distant Data

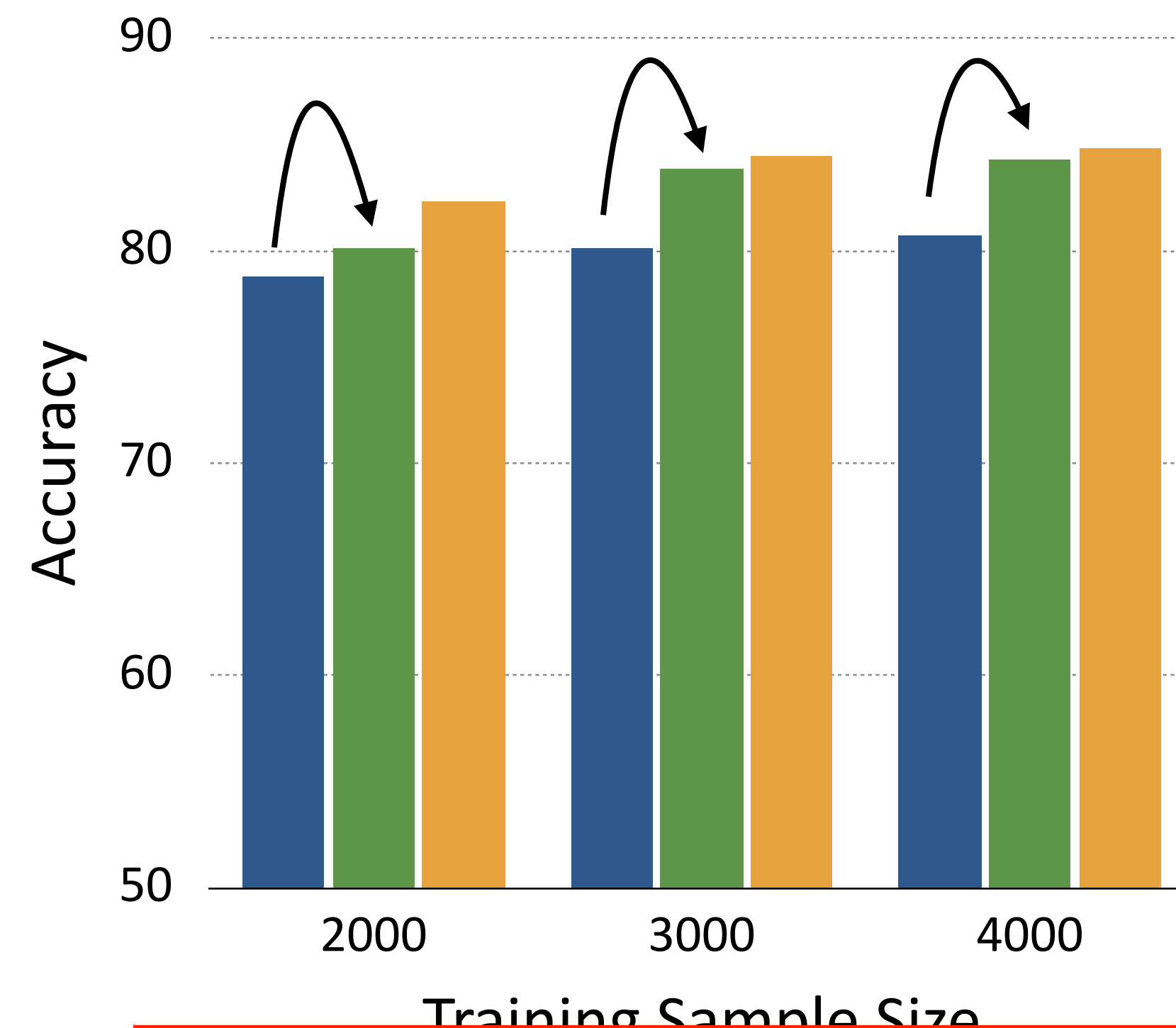
- Masked Timex
stepped down in xxxx
 - Baseline
stepped down in 1959
 - w/ Timex Embed
stepped down in 1959
- +
- (Heuristically labelled,
performance should be
close to 100%)

Glove Embeddings



Timex embeddings
improve temporal relation
extraction performance
over pre-trained GloVe.

ELMo Embeddings



For small training samples,
timex embeddings contain
useful temporal
information over ELMo.



Takeaways

- ▶ Synthetic data can be used to encode temporal knowledge into embeddings.
- ▶ Useful to encode rule-based knowledge in low-data settings.
- ▶ Task-specific embeddings can be leveraged to improve the performance of downstream tasks.

Thank You!