FACULTY OF ARTS
AND PHILOSOPHY

FILLING IN THE GAPS:
EFFICIENT EVENT COREFERENCE RESOLUTION USING GRAPH AUTOENCODER NETWORKS

Loic De Langhe, Orphée De Clercq, Véronique Hoste
CRAC 2023 -6/12/2023

GHENT
UNIVERSITY

language and
translation
technology
team

INTRODUCTION: CD EVENT COREFERENCE RESOLUTION

- Link textual events (either real or fictional) if they refer to the same conceptual event (same time, same place, same participants)
- Events can be either noun phrases (NPs) or verb phrases (VPs)

DOCUMENT 1

The 2022 Fifa World Cup started on November 20th and lasted until December 18th. Argentina captain Lionel Messi lifted the cup in what was Argentina's third victory on the world's biggest stage

DOCUMENT 2

The 2022 football world championship lasted a total of 28 days from start to finish. Argentina won its 3rd championship in a fiery final against the French national team.

GHENT UNIVERSITY

SCALING CROSS-DOCUMENT COREFERENCE

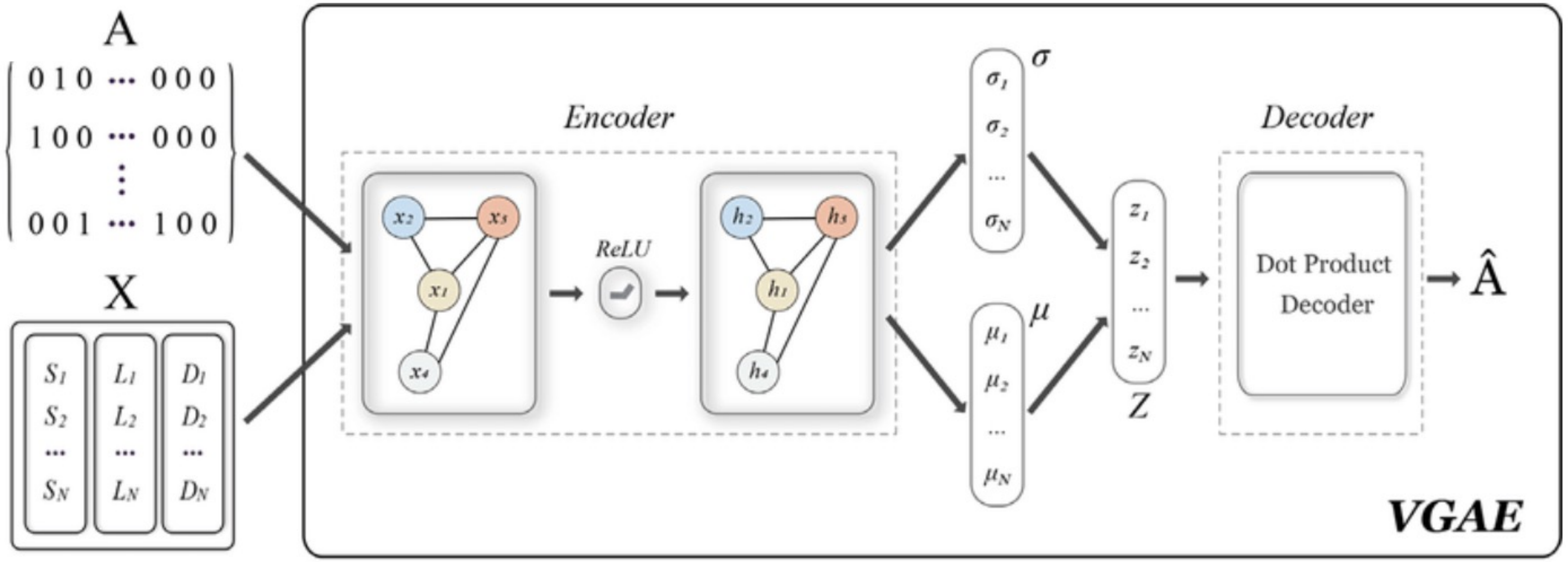
- Most present-day coreference models rely on pairwise computation between anaphora and candidate antecedents
- In cross-document settings this means that the number of needed computations grows exponentially
- Some mitigation strategies in place, but these often rely on document-level clustering Within-document models' output

GHENT UNIVERSITY

EXPERIMENTS

A GRAPH-BASED COMPLETION MODEL

- Graph Auto-encoders
  - Introduced by Kipf and Welling (2016)
  - Reconstructs incomplete graph data based on a set of given edges
  - Mostly used in citation network prediction or molecule completion, not many applications in NLP
  - Probabilistic (VGAE) and non-probabilistic (GAE) settings



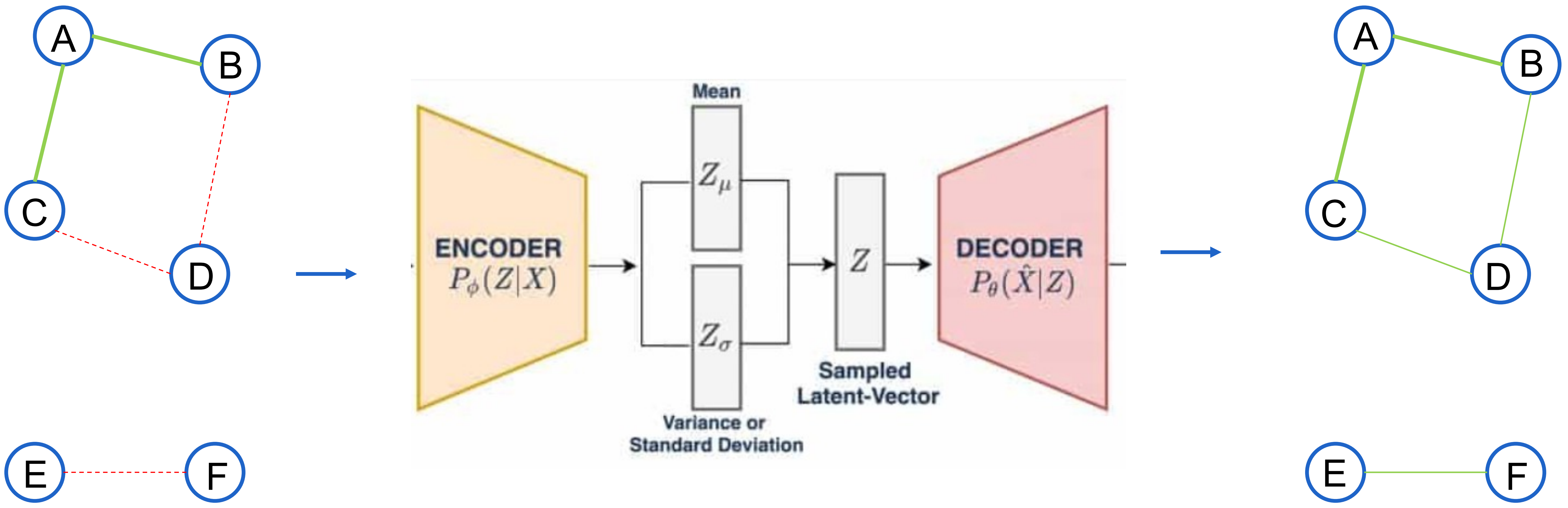
GHENT UNIVERSITY

EXPERIMENTAL SETUP WITH GVAE

- Training Setup
  - Assume that coreference chains can be modelled as undirected, unweighed graphs, where each node represents an event and each edge represents a coreferential relation between two events
  - Mask 15% of edges, 5% to be used as validation data and 10% as test data
  - Experiments using both the Probabilistic (VGAE) and non-probabilistic (GAE) setting

- Input Features
  - Average-pooled BERT embeddings of the tokens in the mention span based on final (768-dim) or 4 final (3072-dim) layer(s)

EXPERIMENTAL SETUP



GHENT
UNIVERSITY

DATA AND BASELINE MODELS

- Dutch ENCORE corpus
  - 1015 documents, 12875 annotated events
  - Cross-document coreference annotations between the events

- Baseline model
  - Encoder-based mention-pair model which creates pairwise representations and scores them through a coreference scoring layer
  - No SpanBERT encoder available for Dutch, use of standard BERT-based encoders instead (BERTje, RobBERT, RobBERTje)

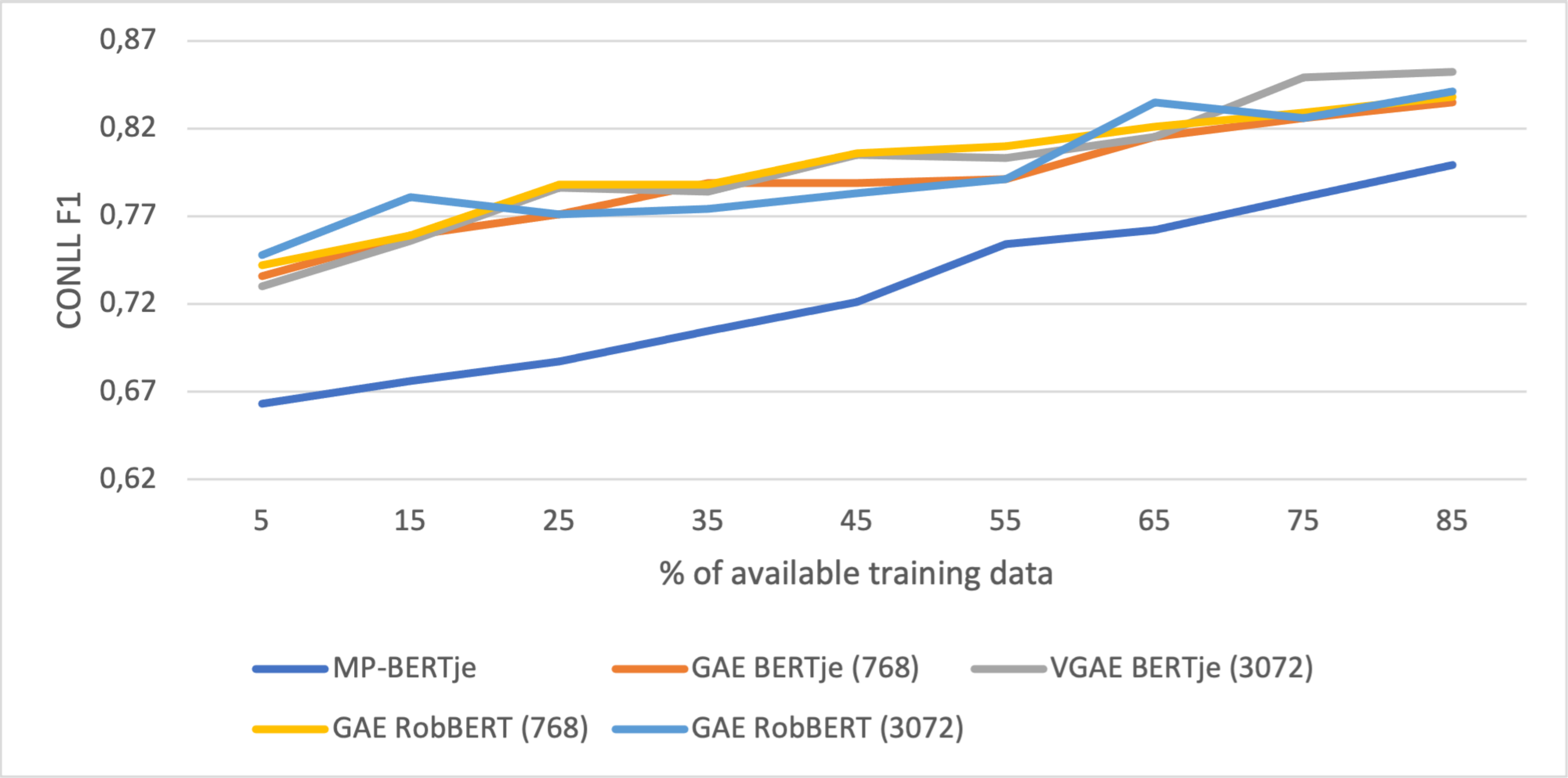
RESULTS AND ABLATION STUDY

RESULTS

Model	CONLL F1	Training Runtime (s)	Inference Runtime (s)	Trainable Parameters	Disk Space (MB)
MP RobBERTje	0.767	7962	16.31	74M	297
MP BERTje$_{ADPT}$	0.780	12 206	20.61	0.9M	3.5
MP BERTje	0.799	9737	21.78	110M	426
GAE BERTje$_{768}$	0.835 ± 0.010	975	0.263	51200	0.204
GAE BERTje$_{3072}$	**0.852 ± 0.006**	1055	0.294	198656	0.780
GAE RobBERT$_{768}$	0.838 ± 0.004	1006	0.273	51200	0.204
GAE RobBERT$_{3072}$	0.841 ± 0.007	1204	0.292	198656	0.780
VGAE BERTje$_{768}$	0.822 ± 0.011	1233	0.282	53248	0.212
VGAE BERTje$_{3072}$	0.842 ± 0.009	1146	0.324	200704	0.788
VGAE RobBERT$_{768}$	0.828 ± 0.0021	1141	0.288	53248	0.212
VGAE RobBERT$_{3072}$	0.831 ± 0.004	1209	0.301	200704	0.788

ABLATION STUDY: DATA AVAILABILITY

- Reduce the amount of available training data by increments of 5%
- Interestingly, the drop in performance for the traditional MP model is larger than for the graph-based models



GHENT UNIVERSITY

FURTHER THOUGHTS AND CONCLUSION

CONCLUSION

- Graph Auto-encoders can be a fast and efficient alternative to traditional mention-pair coreference setups (**when some links are known!**)
- (V)GAE models do not need a lot of data to provide a satisfactory reconstruction

- The setting in which edge data is given is still quite artificial

  ➔ In the future, evaluate graph-based models for graphs were no edges are given

  ➔ Integrate VGAE-based method in end-to-end settings