ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

HAEE: Low-Resource Event Detection with Hierarchy-Aware Event Graph Embeddings

Guoxuan Ding[1,2], Xiaobo Guo[1(✉)], Gaode Chen[1,2],
Lei Wang[1], and Daren Zha[1]

[1] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

1

中国科学院信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING,CAS

2

中国科学院大学
University of Chinese Academy of Sciences

November 2023

中国科学院大学
University of Chinese Academy of Sciences

ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

CONTENTS

中国科学院大学
University of Chinese Academy of Sciences

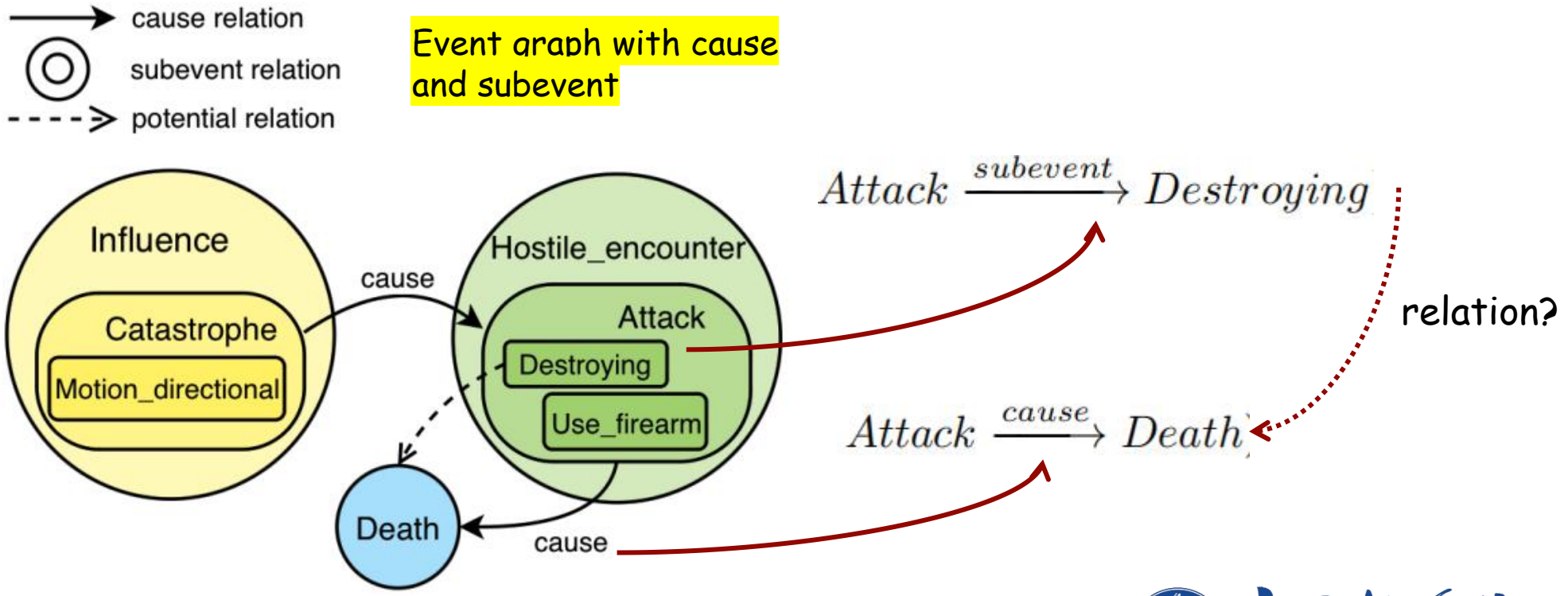
ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

CONTENTS

Events and Event Relations

➢**More abstract connections**:

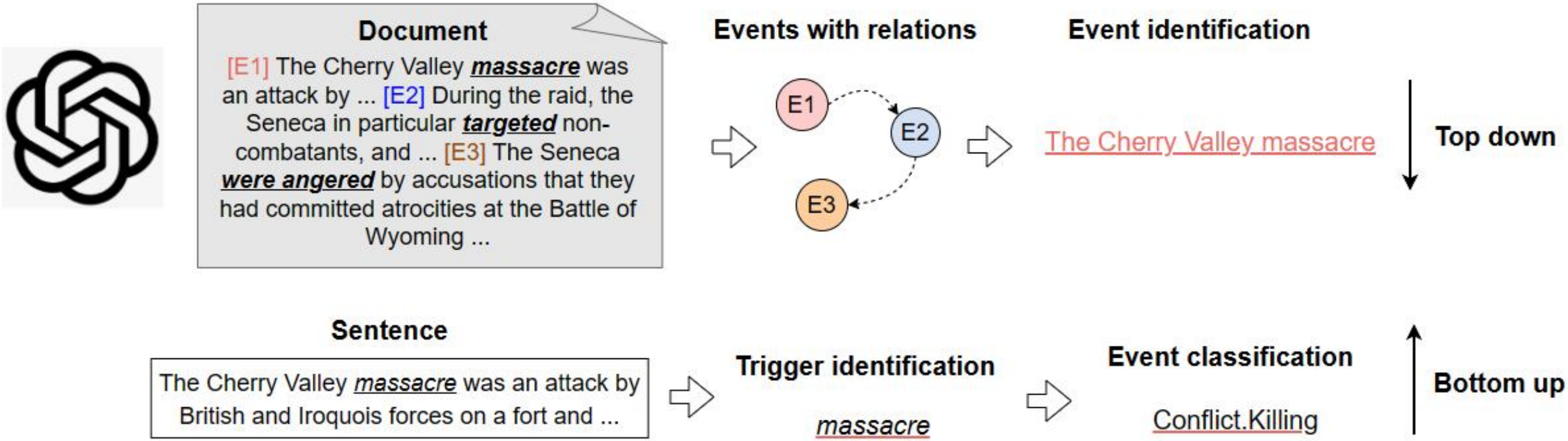
Event relations connect individual events together to form a logical group of events.



Maven-ere: A unified large-scale dataset for event coreference, temporal, causal, and subevent relation extraction. (2022)

Different Direction in Event Detection

➢**LLM:** Events are connected and dynamic



➢**Before LLM:** Events are isolated and static

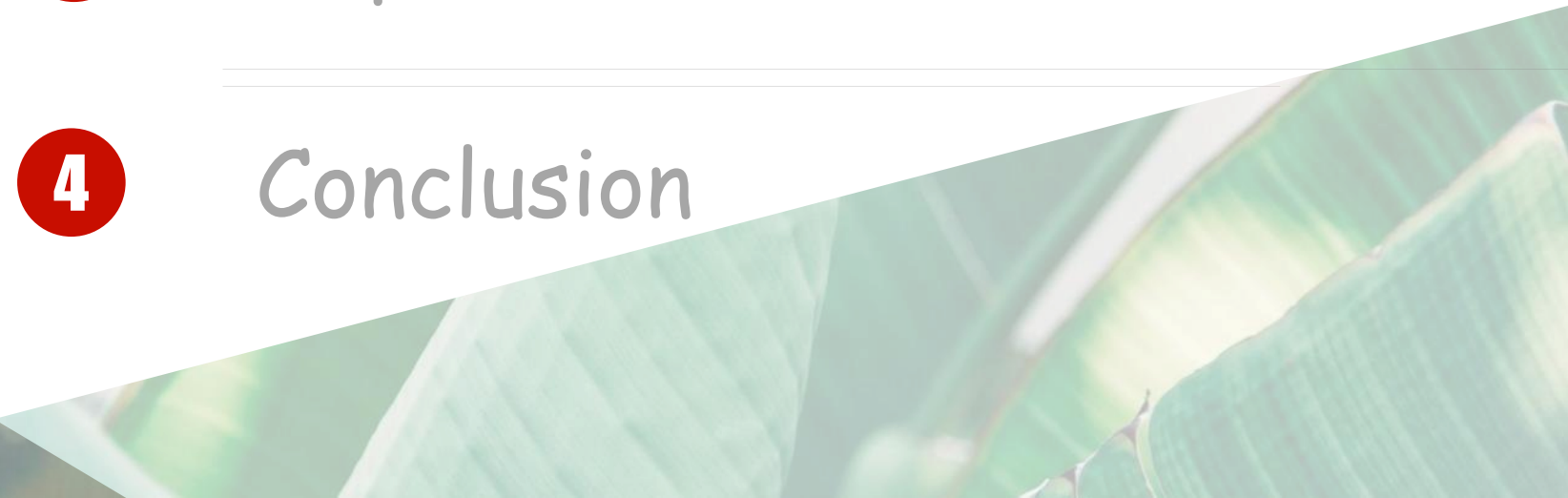
missing event relations?

University of Chinese Academy of Sciences

中国科学院大学
University of Chinese Academy of Sciences

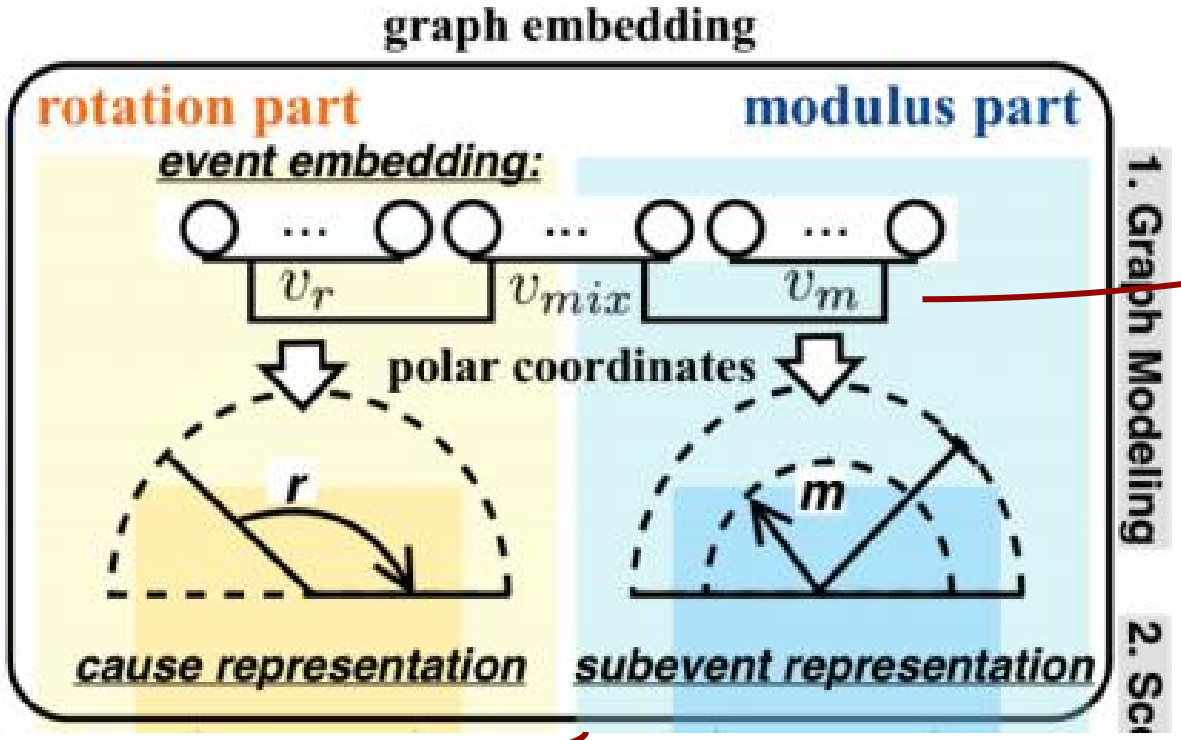
ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

CONTENTS



Embedding with Rotation and Modulus

➢**Polar coordinates consisting of rotation and modulus**



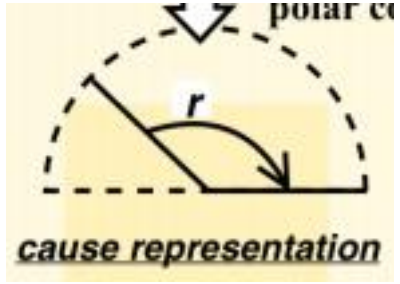
Embeddings are divided into **Rotaion** part and **Modulus** part

Rotation part presents ***cause*** relation and **Modulus** part presents ***subevent*** relation

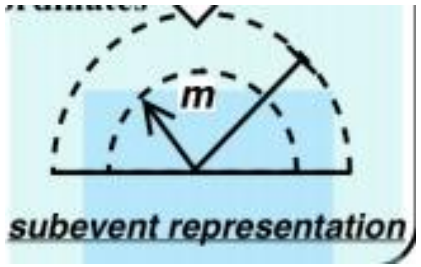
University of Chinese Academy of Sciences

The Calculation of Rotation and Modulus

- **Rotation Part -** rotation-based approach
  - The event pairs are set at **opposite positions** on the circle: $(\boldsymbol{v} - \boldsymbol{v}_c)\, mod\, 2\pi = \pi.$
  - The **distance** between events: $D(e_h, e_t) = \|sin((\boldsymbol{v}_h - \boldsymbol{v}_t + \pi)/2)\|_1$
  - The rotation loss: $\mathcal{L}_r = \sum_{e_c \in \mathcal{E}_c} -log\sigma(\lambda - D(e, e_c))$



- **Modulus Part -** contrastive learning
  - The distance **between child events** is as close as possible compare to the distance **from child events to their parents**
  - The **distance** between events: $D(e_h, e_t) = \|\boldsymbol{v}_h - \boldsymbol{v}_t\|_2$
  - The modulus loss: $\mathcal{L}_m = \sum_{(e,e_i)\in\mathcal{P}_d} \sum_{(e_i,e_j)\in\mathcal{P}_s} max(D(e, e_i) - D(e_i, e_j) + \gamma, 0).$

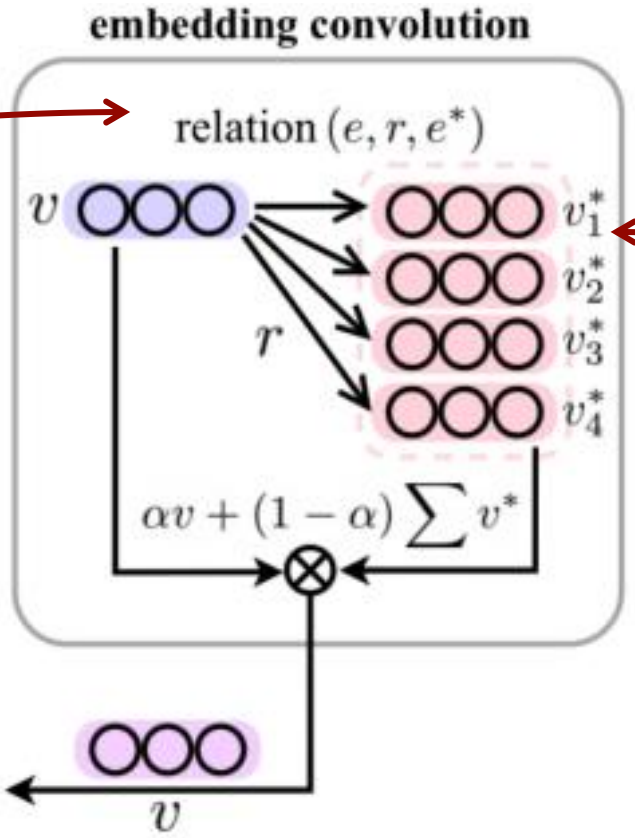


Embedding Convolution

- **Enhance the representation of events**

  further and hierarchical relation knowledge between events can be learned

For the given embedding e and relation r



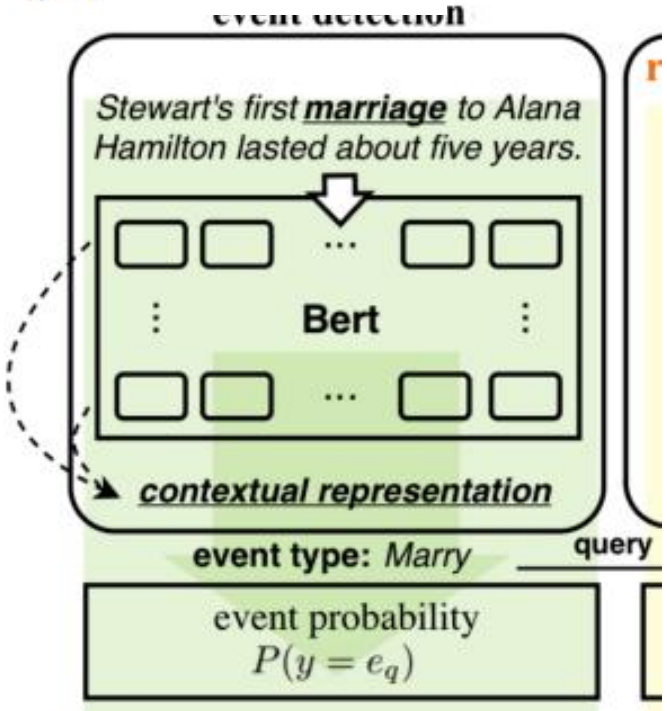
Find all e^* that satisfy (e, r, e^*)

Form the event embedding

Event Detection

- Event **probability**: $P(y = e_k) = \frac{exp(-||\boldsymbol{X}_i - \boldsymbol{v}_k||)}{\sum_{j=1}^{|\mathcal{E}|} exp(-||\boldsymbol{X}_i - \boldsymbol{v}_j||)}$
- Event detection loss: $\mathcal{L}_e = -\sum_{k=1}^{|\mathcal{E}|} ylogP(y = e_q)$

The first and last layers of BERT's transformer



Loss Function

- **Uncertainty to Weigh Losses multi-task optimization strategy (UWL)**

  Combine three loss function

Rotation loss

$$\mathcal{L} = \frac{1}{2\sigma_1^2}\mathcal{L}_e + \frac{1}{2\sigma_2^2}\mathcal{L}_r + \frac{1}{2\sigma_3^2}\mathcal{L}_m + log\sigma_1\sigma_2\sigma_3$$

Event loss

Modulus loss

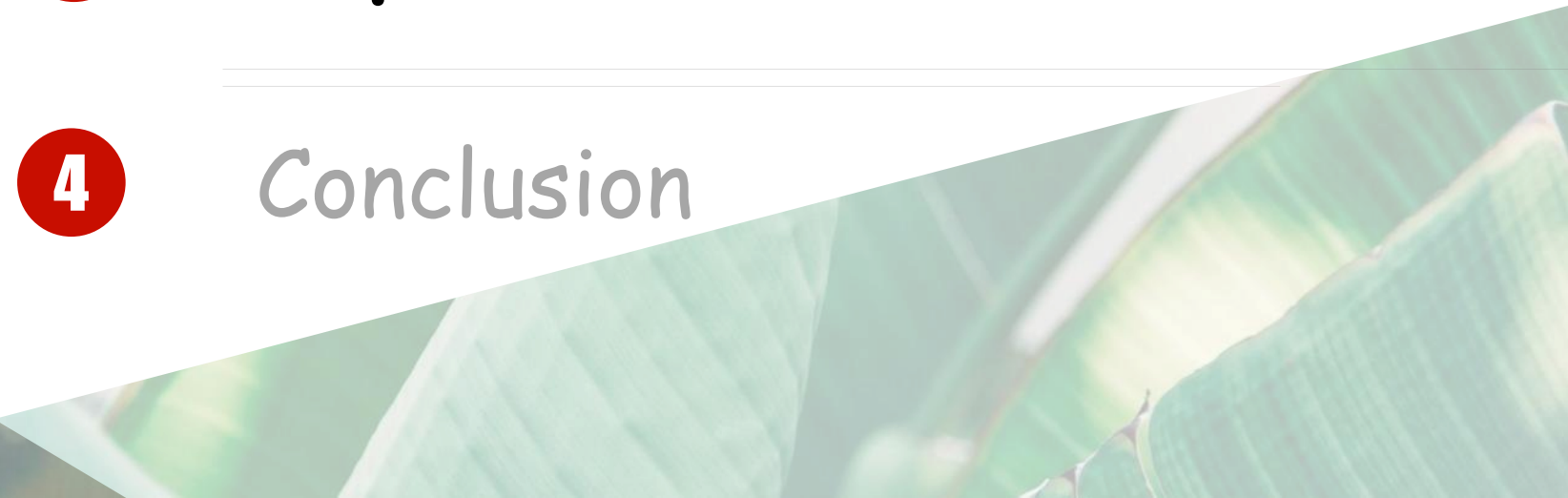
Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. (2018)

University of Chinese Academy of Sciences

中国科学院大学
University of Chinese Academy of Sciences

ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

CONTENTS



Datasets and Baselines

- **Datasets**
  - OntoEvent
  - MAVEN-Few

Dataset	#Doc	#Train	#Valid	#Test	#Class	#Caus-Rel	#Sub-Rel
OntoEvent [5]	4115	48436	6055	6055	100	9	-
MAVEN-Few	-	4416	552	551	71	277	83

-Selected from **MAVEN** datasets according to event types in OntoEvent.

- **Baselines**
  - Event Classification: AD-DMBERT, OneIE, PathLM, OntoED
  - Low-resource event classification: OntoED

OntoED: Low-resource Event Detection with Ontology Embedding (ACL 2021)
Maven-ere: A unified large-scale dataset for event coreference, temporal, causal, and subevent relation extraction. (2022)

University of Chinese Academy of Sciences

Event Classification

Model	Precision	Recall	F1 Score
AD-DMBERT† [25]	0.6735	0.7346	0.7189
OneIE† [15]	0.7194	0.6852	0.7177
PathLM† [13]	0.7351	0.6874	0.7283
OntoED‡ [5]	0.7756	0.7844	0.78
HAEE	**0.8882**	**0.8868**	**0.8875**

Use event relations as ontology learning

Model	OntoEvent				MAVEN-Few			
	Full	50%	25%	10%	Full	50%	25%	10%
OntoED [5]	0.78	0.7154	0.6198	0.4989	0.7725	0.6034	0.5195	0.2534
HAEE	**0.8875**	**0.8747**	**0.8634**	**0.831**	**0.8722**	**0.8577**	**0.8165**	**0.5993**

Hierarchy of Rotation

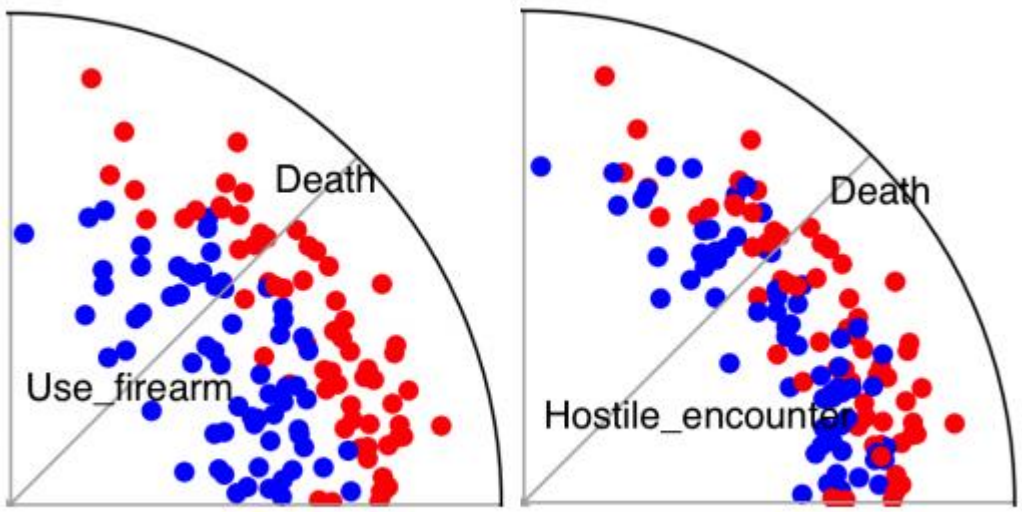
Table 4: The average rotation distance from each event to different groups. (p.g: positive group, b.g: blank control group, n.g:negative group. The bold numbers represent groups with a greater distance.)

Positive Events	Distance			Negative Events	Distance		
	p.g	b.g	n.g		p.g	b.g	n.g
come_together	0.45	0.55	**0.59**	destroying	**0.34**	0.23	0.24
elect	0.34	0.42	**0.47**	kidnapping	**0.30**	0.22	0.23
committing_crime	0.19	0.21	**0.28**	violence	**0.28**	0.21	0.22
employment	0.19	0.19	**0.23**	theft	**0.45**	0.32	0.33
award	**0.27**	0.20	0.21	robbery	0.18	0.19	**0.24**
arriving	0.31	0.38	**0.44**	hostile_encounter	0.18	0.20	**0.26**
contact	**0.51**	0.36	0.35	killing	0.23	0.29	**0.37**
recovering	0.21	0.19	**0.22**	terrorism	0.28	0.37	**0.45**
commerce_sell	0.28	0.34	**0.40**	conquering	**0.26**	0.20	0.22
exchange	0.19	0.20	**0.26**	arrest	0.19	0.19	**0.23**
marry	0.37	0.46	**0.51**	divorce	0.18	0.19	**0.25**

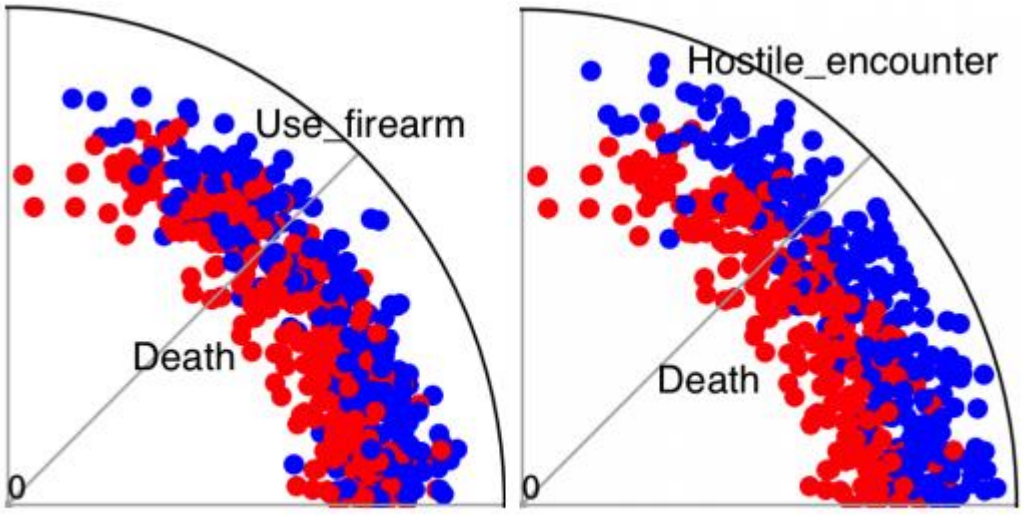
Tend to away from negative events

The clustering effect of events is more significant

Hierarchy of Modulus



(a) 100 dimension.

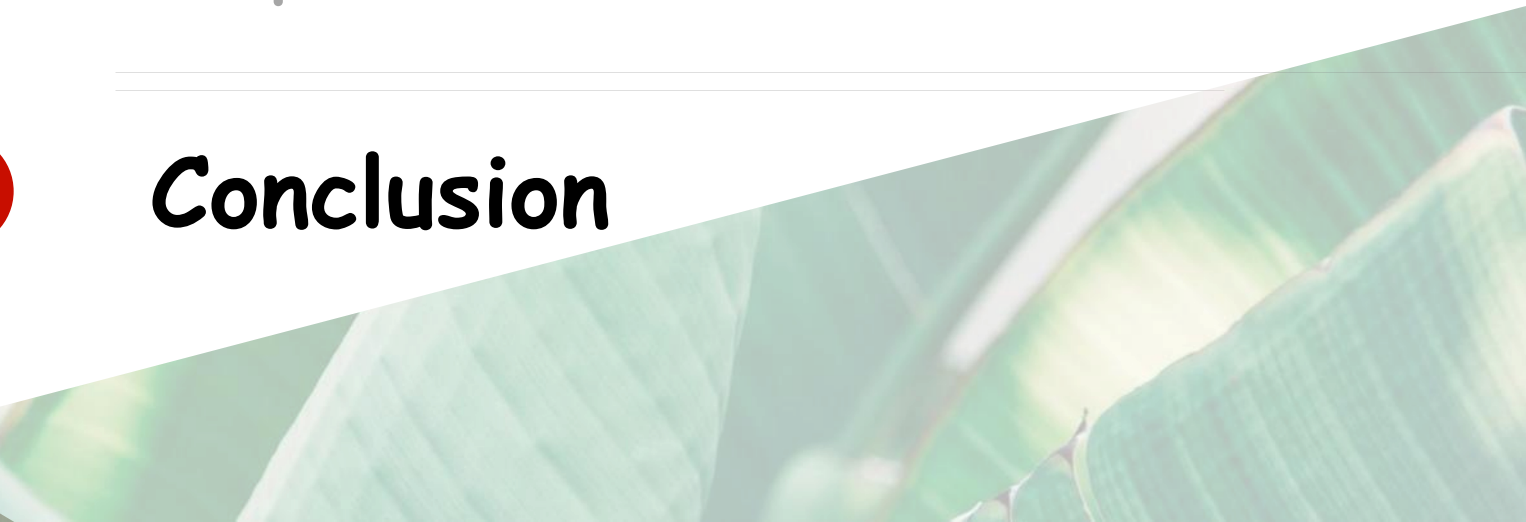


(b) 500 dimension.

中国科学院大学
University of Chinese Academy of Sciences

ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

CONTENTS



Conclusion

- **Summary**
  - We propose a new hierarchy-aware model HAEE, allowing knowledge to flow from high-resource events into the low-resource events.
  - Experimental results demonstrate that HAEE model can achieve better performance in low-resource ED task.
- **Outlook**
  - Exploring better graph embedding methods to introduce more event relations into event graphs
  - Applying event relations to other fields such as sentiment analysis and event reasoning.

University of Chinese Academy of Sciences

ISWC 2023
NOVEMBER 6-10
ATHENS-GREECE

中国科学院信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING,CAS

中国科学院大学
University of Chinese Academy of Sciences

Thank You

Speaker: Guoxuan DING

Contact: dingguoxuan@iie.ac.cn

Project: https://github.com/cdmelon/HAEE

November 2023