Web Conference 2020, Taipei, Taiwan

Identifying Referential Intention with Heterogeneous Contexts

Wenhao Yu, Mengxia Yu, Tong Zhao, Meng Jiang

Department of Computer Science and Engineering,
University of Notre Dame, Indiana, USA

Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Referential Behavior

- It is a type of behavior when people talk about or write about something to express their opinions or ideas.
  - **Social media:** Retweeting someone's message to community
  - **Academic writing:** Citing someone's paper in some sentences
  - . . .

Author

write Generated content (GC)

refer to Referred content (RC)

Author

GC

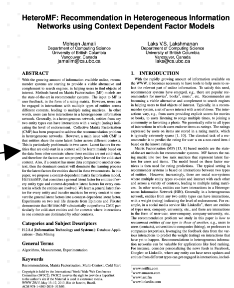
RC

Author:

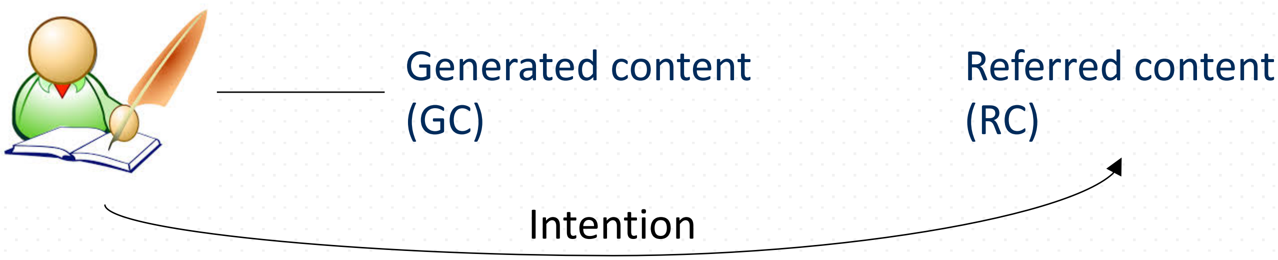
Contextual behavior modeling. There has been a wide line of research on learning latent representations of context items in behavior data towards various applications [3, 21, 34]. Agarwal et al. proposed localized factor models combining multi-context information to improve predictive accuracy in recommender systems [2]. Jamali et al. proposed context-dependent factor models to learn latent factors of users and items for recommendation [19]. Besides factor models, tensor decompositions have been widely used for modeling multi-contextual data [33]. Jiang et al. proposed a tensor-sequence decomposition approach for discovering multi-faceted behavioral patterns [20]. Ermiş et al. studied various al-

GC

RC



Identifying Referential Intention



Problem definition in existing work:

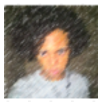
$$f: \text{GC}, \text{RC} \rightarrow \text{Intention type}$$

Related work 1: **Stance detection**: Identifying the author's stance in GC.
Schema: Positive, (Neutral,) Negative

Related work 2: **Sentiment analysis**: Identifying the author's sentiment in GC.
Schema: Polarity: Positive, Negative
Emotion: Happy, Sad

"Topic Modeling with Sentiment Clues and Relaxed Labeling Schema", ACL 2013.

Referential Intention ≠ Author's Sentiment/Stance



RT: So True. @Obama is truly amazing! Glad you're giving credit where credit is due!

GC

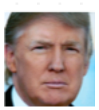
{true, truly, amazing, glad, credit}

Positive √ ~~Negative~~

Happy √ ~~Sad~~

Sentiment analysis

Then what if RC =

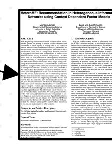


Stock Market is heading for one of the best months (June) in the history of our Country. Thank you Mr. President!

author of *node2vec*

Positive stance: **When describing the impact of network representation learning**

Negative stance: **When describing the uniqueness and advantage of the proposed work** (vs. *LINE*)

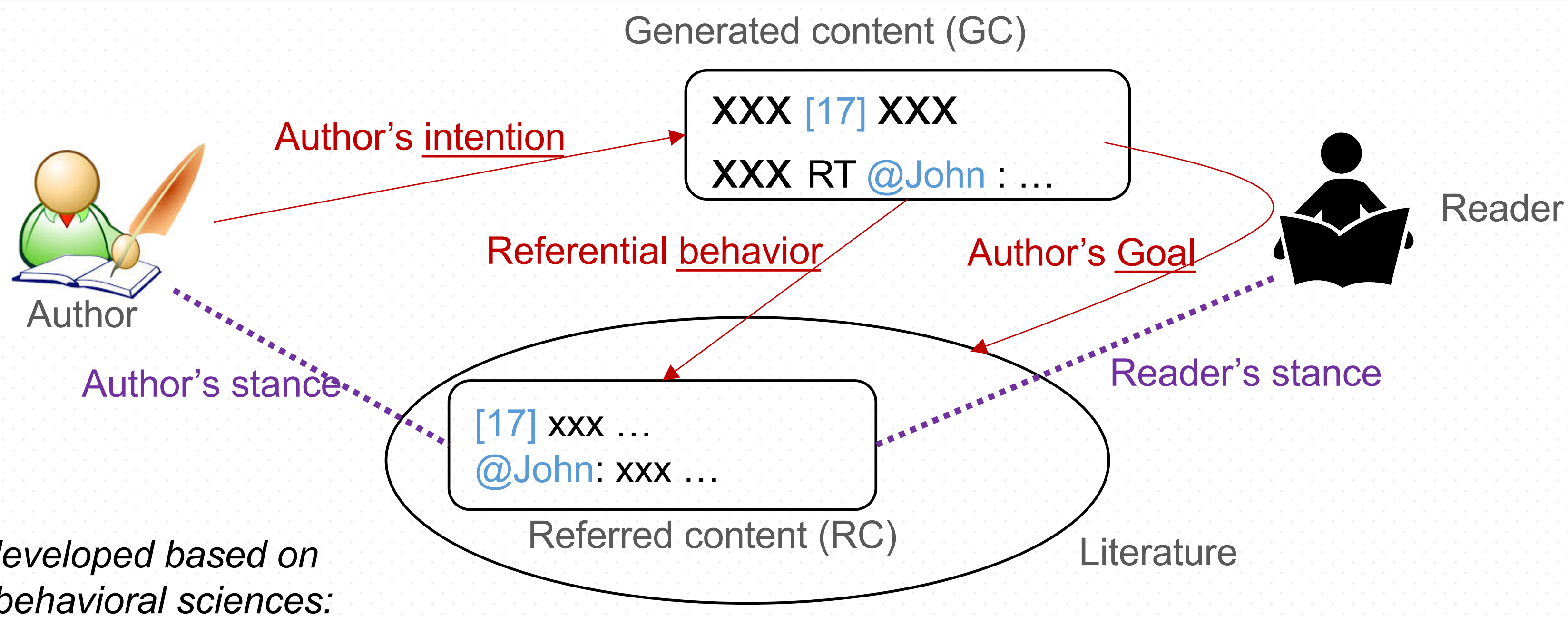


RC = *LINE*

(*node2vec* [KDD'16] and *LINE* [WWW'15] are two network representation learning algorithms)

What if (and why) an author has different stances to the same thing?

A Novel Intention Type Schema



Schema developed based on a work in behavioral sciences:

Referential Intention		Reader's Stance		
		Positive	**Neutral**	**Negative**
Author's Stance	**Positive**	Accept	Strong Accept	Strong Accept
	Neutral	Background	Background	Background
	Negative	Strong Reject	Strong Reject	(Strong) Reject

Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Challenges

- Identifying referential intention in the **new schema**

  Strong Accept, Accept, Background, Strong Reject

- **Needs contextual information of the referential behavior:** Not only GC and RC but also the former sentences, latter sentences, and author/venue/keyword information.

- **Needs to learn from heterogeneous data:** textual content and network (social network, academic network, etc.)

Heterogeneity of Referential Contexts

Generated content

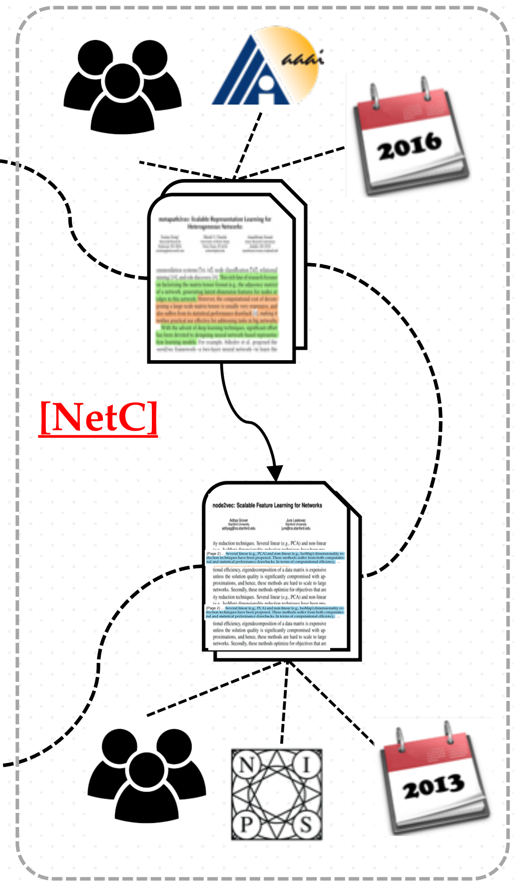
(Page 1) … [NC-former] This step can be understood as a process for converting a general graph structure into a large collection of linear structures. [LC (Local Context)] Next, the skip-gram model proposed by [22] to learn low-dimensional representations for vertices from such linear structures. The learned vertex representations were shown to be effective across a few tasks, outperforming several previous approaches such as … [NC-latter] While such an approach for learning vertex representations of a unweighted graph is effective …

[22] Mikolov et al. (NeurIPS 2013) Distributed representations of words and phrases and their compositionality.

Referred content

(Page 7) … [RC] We demonstrated that the word and phrase representations learned by the Skip-gram model exhibit a linear structure that makes it possible to perform precise analogical reasoning using simple vector arithmetics. …

Network context



Problem Definition

Intention $S = \{“SA”, “A”, “BG”, “SR”\}$

$g(b): B \rightarrow S$

NC-former

This step can be understood as a process for converting a general graph …

NetC

LC

Next, the skip-gram model proposed by [22] to learn low-dimensional …

NC-latter

While such an approach for learning vertex representations of a unweighted graph is …

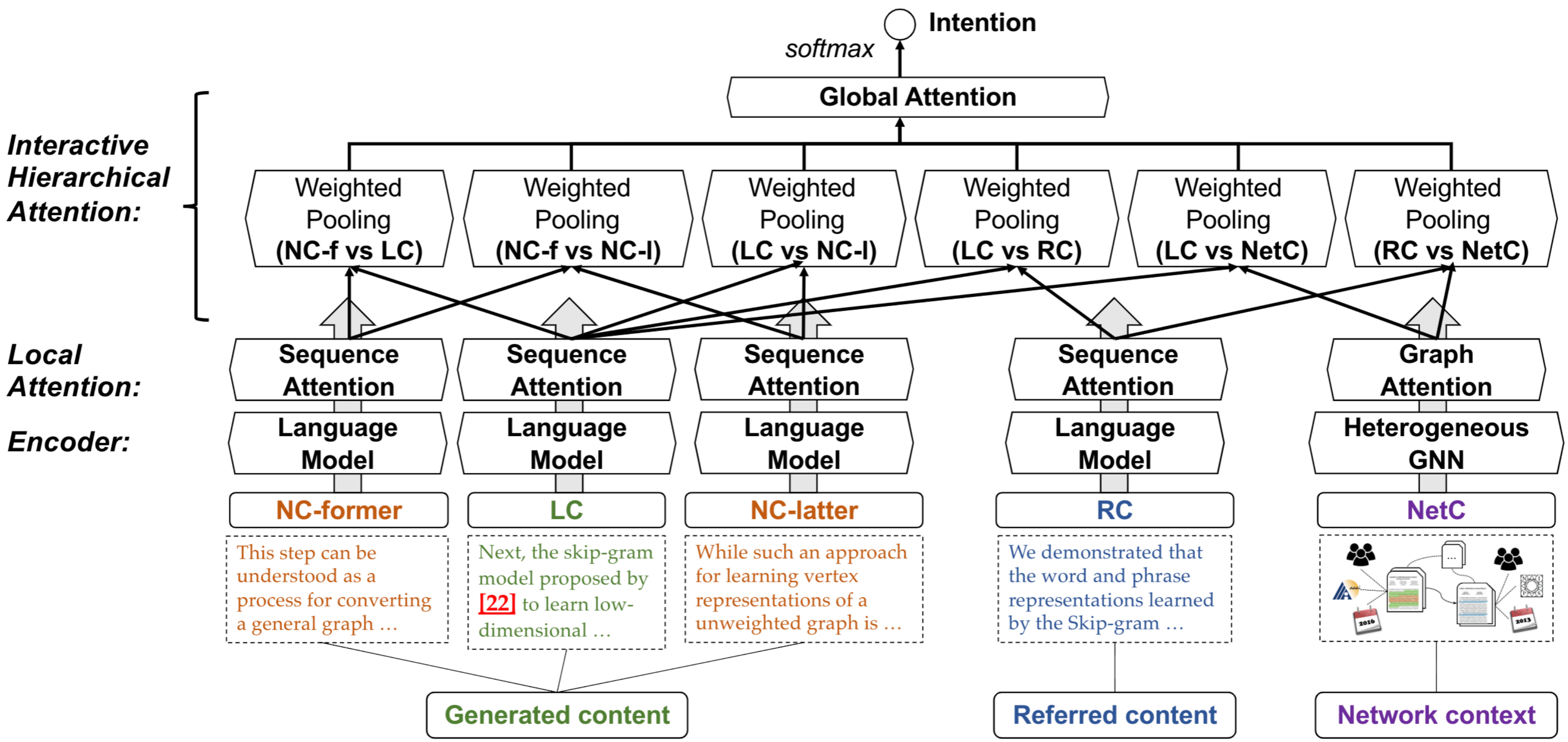
RC

We demonstrated that the word and phrase representations learned by the Skip-gram …

Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Proposed Framework



Framework Components

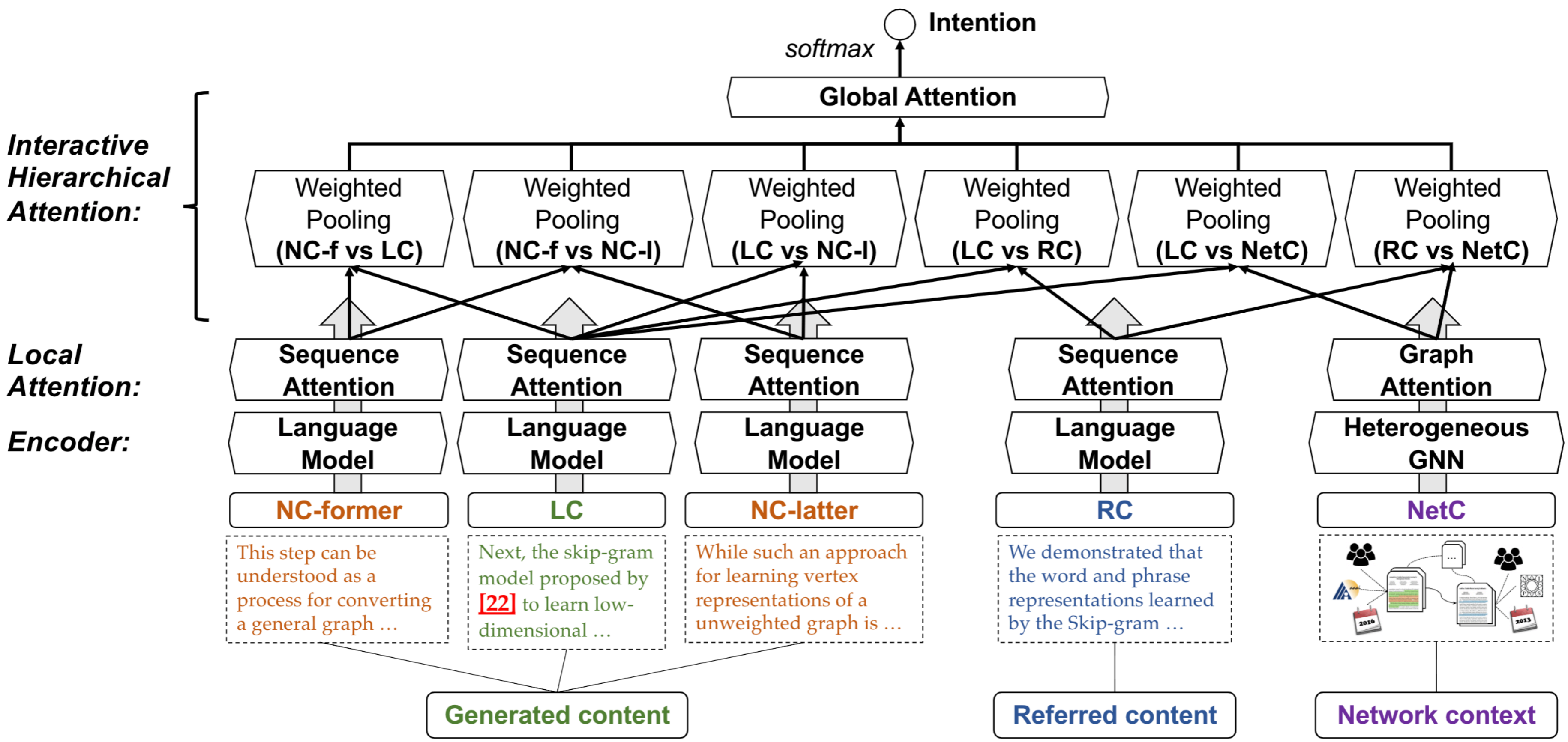
- Modeling textual content (i.e., NC-former, LC, NC-latter, RC)
  - Language model
  - Sequential attention
- Modeling network contexts (i.e., NC)
  - Heterogeneous GNN
  - Graph attention
- Interactive hierarchical attention
- Global attention

Interactive Hierarchical Attention

Multiple types of contexts have interaction with each other

- (1) LC and NC came originally from the same document and form semantic ordering;
- (2) RC and LC were tightly related to the same topic;
- (3) RC and LC were produced by a pair of linked nodes in NetC.

Review: Proposed Framework



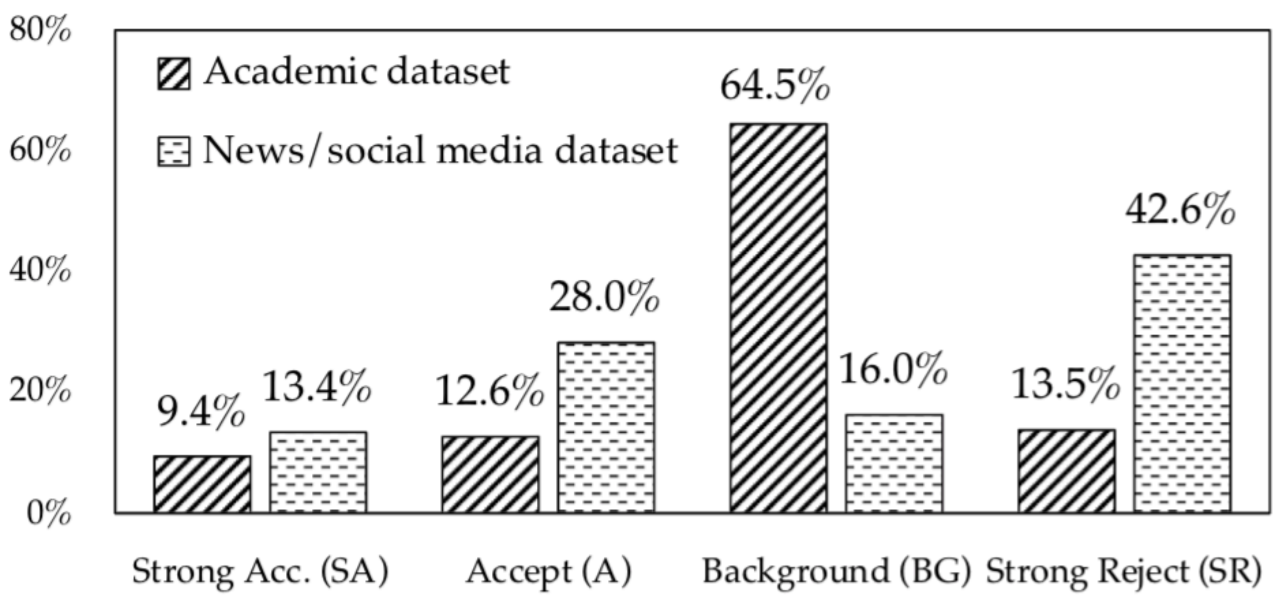
Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Data Description and Experimental Settings

Academic dataset. (1) 40 papers from 2 popular research topics (network embedding and language model); (2) published in *WWW, KDD, ICDM, AAAI, NeurIPS,* and *ACL*; (3) labelled 1,565 citations in total.

News/social media dataset. (1) 297 news articles and retweets about 20 political events in the US in 2019. (2) labelled 401 referential behaviors.



Experimental Results

	Academic dataset (%)					News/social media dataset (%)				
	Acc.	**Prec.**	**Rec.**	**F1**	**AUC**	**Acc.**	**Prec.**	**Rec.**	**F1**	**AUC**
CiteFrame (2018) [14]	63.59	54.95	52.66	52.72	71.61	-	-	-	-	-
CiteFunc (2019) [19]	71.74	63.22	53.61	57.12	74.83	56.25	57.01	56.11	53.67	73.64
SecClass (2019) [6]	79.35	67.71	68.43	67.86	83.04	61.25	61.87	64.07	62.72	78.57
Ours (**M8-IHA**)	**82.21**	**77.01**	**70.14**	**73.14**	**89.66**	**68.75**	**70.36**	**72.10**	**70.72**	**82.31**

Observations:

- A: Our purposed methods outperform all baseline methods.
- B: The baselines cannot extract and learn the important information.

Ablation Study

	Contexts				Academic dataset (%)					News/social media dataset (%)				
	LC	NC	RC	NetC	Acc.	Prec.	Rec.	F1	AUC	Acc.	Prec.	Rec.	F1	AUC
M1	✔				77.16	66.94	63.39	64.77	84.42	61.00	60.30	62.05	62.08	74.45
M2	✔		✔		79.35	71.35	62.10	65.59	84.60	63.75	64.51	65.24	64.46	81.40
M3	✔			✔	79.62	75.34	60.44	65.33	84.27	65.00	67.26	66.83	66.79	80.02
M4	✔		✔	✔	79.89	74.11	64.90	68.45	84.46	67.75	67.33	69.97	67.94	81.75

Observations:

- A: adding RC (M2) and NetC (M3) improves performance.
- B: add both RC and NetC could further improves performance.

Ablation Study (cont'd)

	Contexts				Academic dataset (%)					News/social media dataset (%)				
	LC	NC	RC	NetC	Acc.	Prec.	Rec.	F1	AUC	Acc.	Prec.	Rec.	F1	AUC
M8(LA)	✔	✔	✔	✔	81.01	74.26	68.35	70.84	89.08	67.75	67.33	69.97	67.94	81.75
M8-HA	✔	✔	✔	✔	81.79	75.06	69.08	71.57	89.35	67.50	69.91	68.83	68.73	82.25
M8-IHA	✔	✔	✔	✔	**82.21**	**77.01**	**70.14**	**73.14**	**89.66**	**68.75**	**70.36**	**72.10**	**70.72**	**82.31**

Observations:

- A: flat local attention or pure hierarchical attention cannot perform well.
- B: interactive attention can significantly improve performance, Each pooling module integrates outputs of two types of contexts with local attention.

Ablation Study (cont'd)

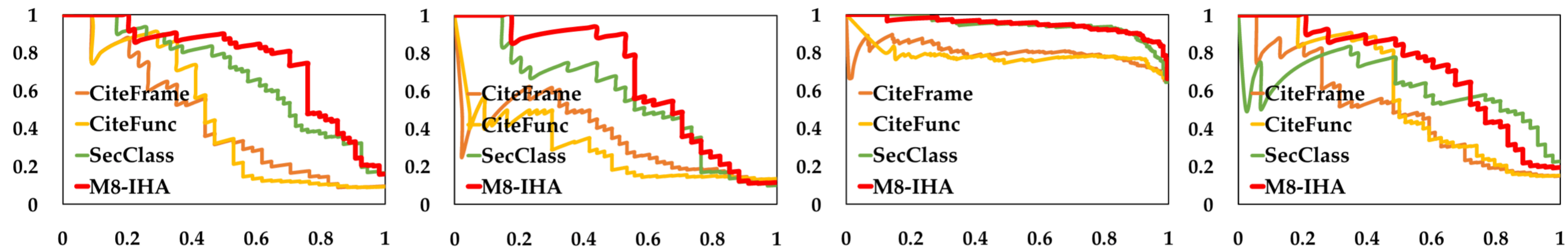
	Contexts				Academic dataset (%)				
	LC	NC	RC	NetC	Acc.	Prec.	Rec.	F1.	AUC
M4	✔		✔	✔	79.89	74.11	64.90	68.45	84.46
M5	✔	✔			78.53	69.14	69.21	68.46	87.80
M6	✔	✔		✔	80.77	73.51	67.92	70.17	88.76
M7	✔	✔	✔		**81.52**	72.35	**69.94**	70.54	**89.50**
M8	✔	✔	✔	✔	81.01	**74.26**	68.35	**70.84**	89.08

	Contexts			Academic dataset (%)				
	LC	NC-f	NC-l	Acc.	Pre.	Rec.	F1.	AUC
M5	✔	✔	✔	78.52	69.14	**69.21**	**68.46**	**87.80**
M9	✔	✔		78.26	**73.22**	65.14	68.19	87.49
M10	✔		✔	**78.53**	69.16	67.98	68.11	87.56

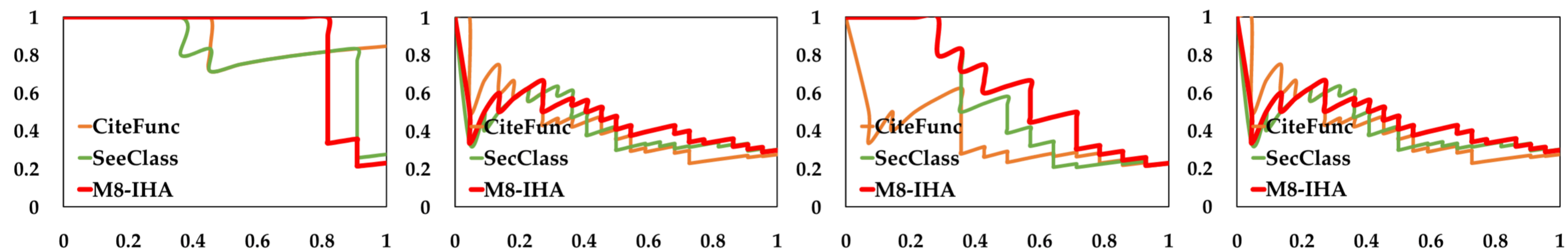
Observations:

- A: Both neighboring contexts(NC-f & NC-l) can improve model effectiveness
- B: combined with RC and NetC (M6–8); it does not have much improvement.

Results on Particular Types



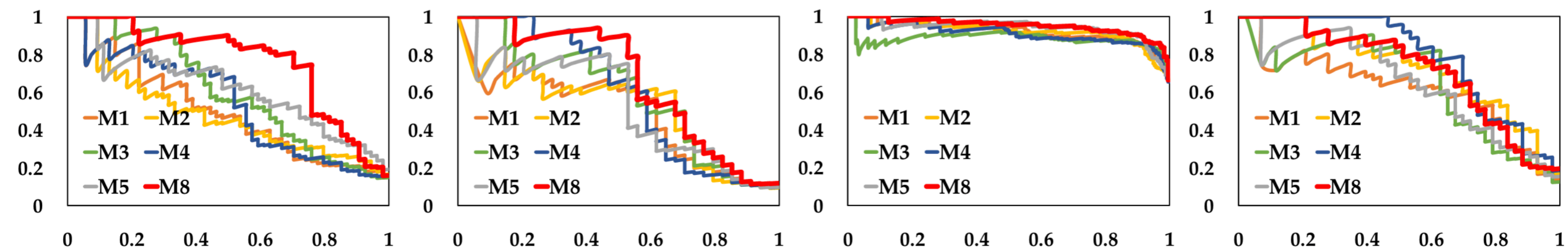
(a) Strong Accept (SA) (b) Accept (A) (c) Background (BG) (d) Strong Reject (SR)



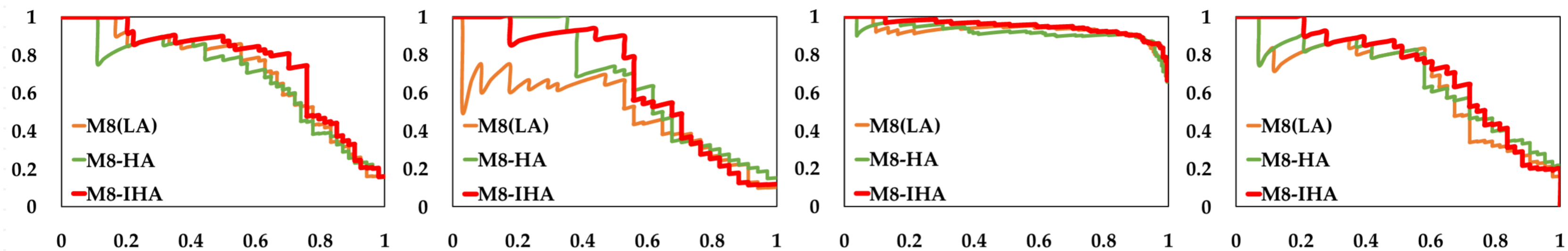
(a) Strong Accept (SA) (b) Accept (A) (c) Background (BG) (d) Strong Reject (SR)

- A: our method consistently performs better than baselines on polar intentions
- B: Compared to polar intentions, BG is relatively easy to be identified.
- C: Strong Accept usually contains literal praise, so predicting SA achieves the highest accuracy.

Results on Particular Types (cont'd)



(a) Strong Accept (SA) (b) Accept (A) (c) Background (BG) (d) Strong Reject (SR)



(a) Strong Accept (SA) (b) Accept (A) (c) Background (BG) (d) Strong Reject (SR)

- A: On the referential intentions of Strong Accept and Accept, M8 performs the best; on Background, all methods are comparable good.
- B: The IHA mechanism is more effective on Accept and Strong Accept.

Case Study (Social Media)

(2) [Tweet AUG.26th 2019] The story by Axios that President Trump wanted to blow up large hurricanes with nuclear weapons prior to reaching shore is **ridiculous**. Just more **FAKE NEWS!**

(7) [Tweet AUG.26th 2019] **Fake news is INSANE!** They are very disturbed individuals led by diabolical companies.

(5) [Post AUG.27th 2019] Trump denies that he suggested nuking hurricanes. But the government once studied the idea. Regardless of Trump's feelings on the subject, the idea of hurling bombs in the path of incoming storms was **once given serious consideration**.

LC Only	LC NetC	LC RC	LC Both
✗	✗	✓	✓
✗	✓	✗	✓

Roadmap

- Motivation
- Challenges and Problem Definition
- Proposed Framework
- Experiments
- Summary

Summary

- A new problem in behavior modeling – identifying intention of referential behaviors. A novel intention schema based on a theory in behavioral and social sciences.
- A new mechanism called Interactive Hierarchical Attention for learning from interactive heterogeneous contexts.
- Experiments on two real-world datasets to demonstrate the effectiveness of our proposed framework.

Web Conference 2020, Taipei, Taiwan

Thank you!


Wenhao Yu


Mengxia Yu


Tong Zhao


Dr. Meng Jiang

If you have any question, please contact wyu1@nd.edu

UNIVERSITY OF NOTRE DAME
College of Engineering

Any Questions?