

# Employing Text Matching Network to Recognise Nuclearity in Chinese Discourse



**Sheng Xu**, Peifeng Li, Guodong Zhou, and Qiaoming Zhu  
Soochow University

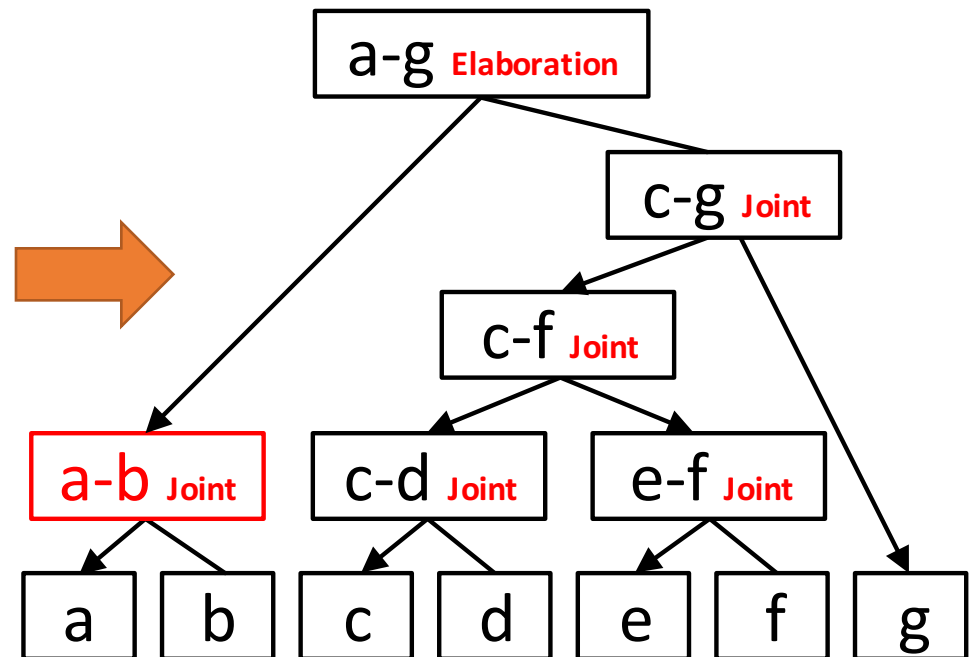
COLING 2018

# Discourse Parsing

- **Task:** Analysing the discourse structure of document
- **Theory:** Rhetorical Structure Theory (RST) [Mann and Thompson, 1988]

## Elementary Discourse Units (EDUs)

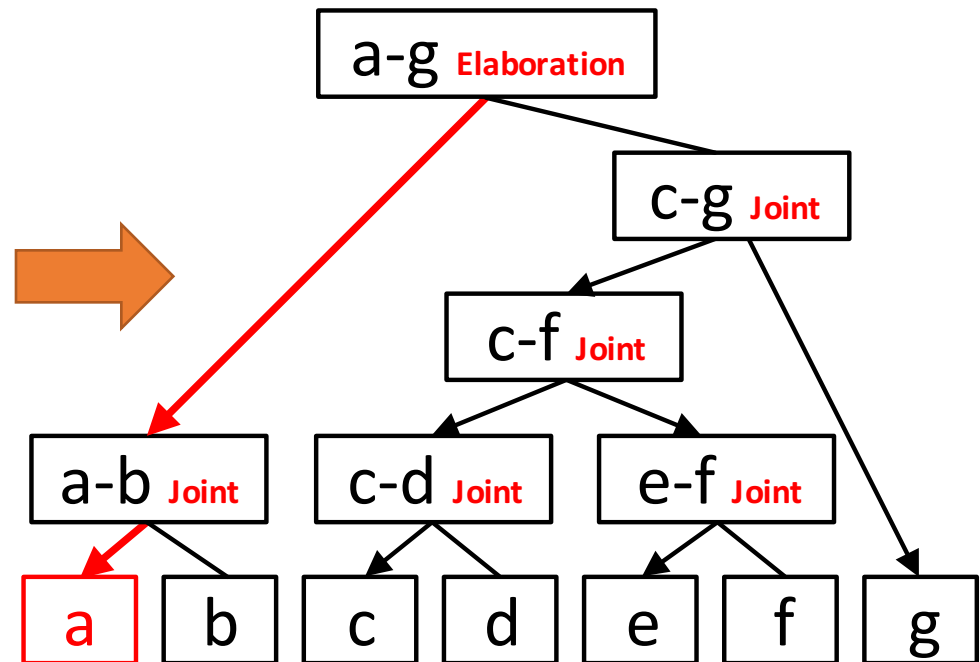
[The import and export trade of China's mechanical and electronic products continues to increase]<sup>a</sup> [and its proportion of the total imports and exports also continues to rise]<sup>b</sup> [Among them, the exports amounted to 5.79 billion dollars]<sup>c</sup> [accounting for 32.5 percent of the total exports]<sup>d</sup> [and the imports of 8.52 billion dollars]<sup>e</sup> [accounting for 46.4 percent of the total imports]<sup>f</sup> [all of them were higher than those in the same period last year]<sup>g</sup>.



# Nuclearity Recognition

- **Task:** Identifying the nuclearity between discourse units
- **Mononuclear:** Nucleus-Satellite / Satellite-Nucleus
- **Multinuclear:** Nucleus-Nucleus

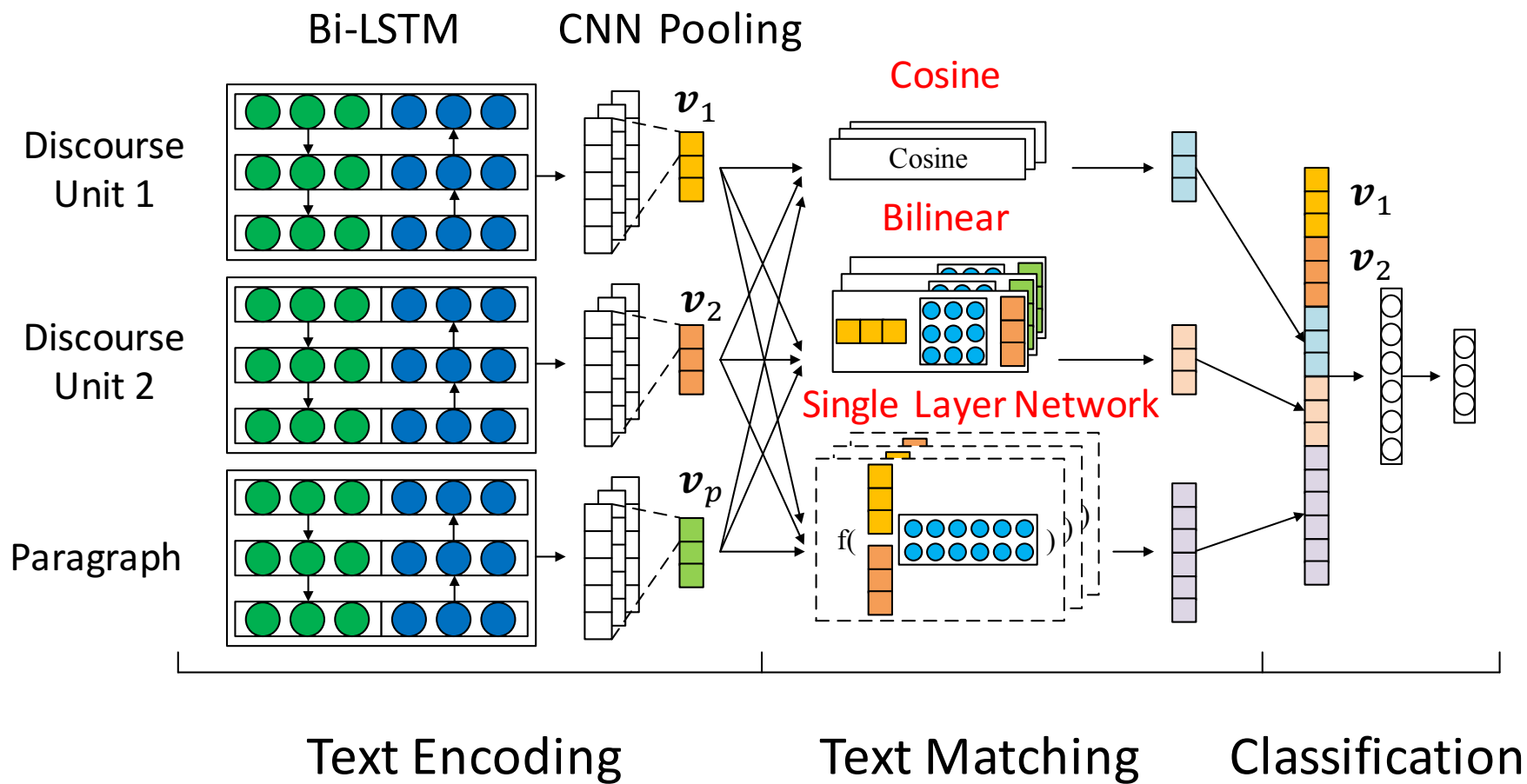
[The import and export trade of China's mechanical and electronic products continues to increase]<sup>a</sup> [and its proportion of the total imports and exports also continues to rise]<sup>b</sup> [Among them, the exports amounted to 5.79 billion dollars]<sup>c</sup> [accounting for 32.5 percent of the total exports]<sup>d</sup> [and the imports of 8.52 billion dollars]<sup>e</sup> [accounting for 46.4 percent of the total imports]<sup>f</sup> [all of them were higher than those in the same period last year]<sup>g</sup>.



# Hypotheses

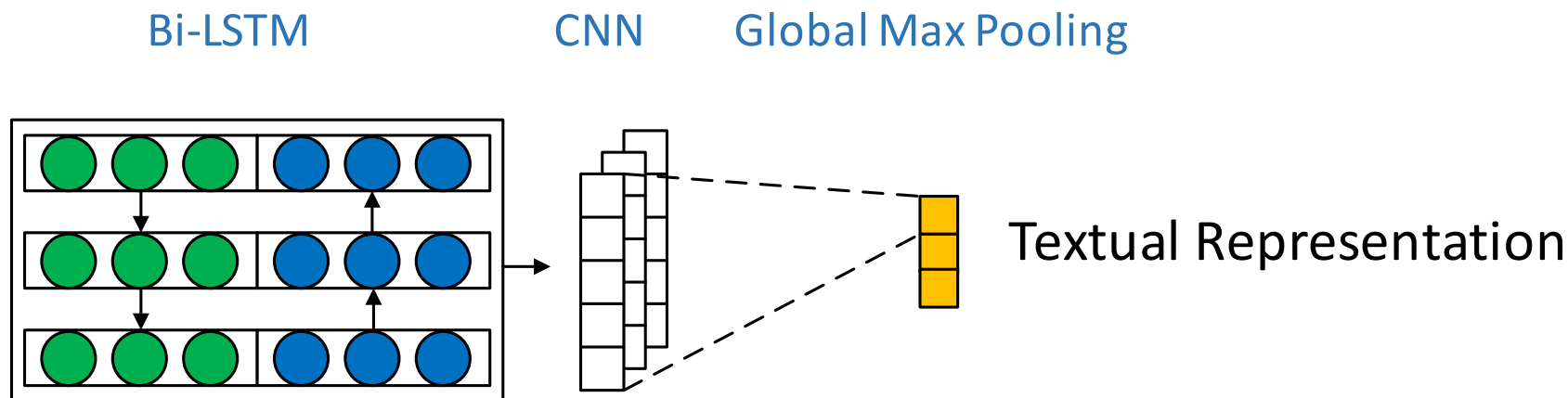
- There are strong correlations between the nuclearity and the semantic similarity or interactions of two discourse units.
  - Discourse units with similar semantics → **Multinuclear**
  - Discourse units with semantic interactions → **Mononuclear**
- The nuclearity of two discourse units is relevant to the topic of the paragraph or document.
  - The nucleus unit is usually semantically closer to the topic of the paragraph.

# Text Matching Network



# Text Encoding

- We combine Bi-LSTM and CNN to jointly learn proper representation of the discourse units, which can capture both **global dependency information** and **local n-gram information**.



# Text Matching

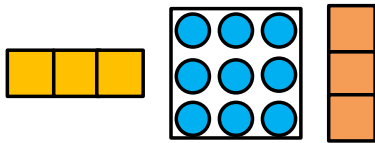
- **Cosine**

$$\cos(\mathbf{v}_1, \mathbf{v}_2) = \frac{\mathbf{v}_1^\top \mathbf{v}_2}{\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|}$$



similarities

- **Bilinear** [Sutskever et al., 2009; Jenatton et al., 2012]



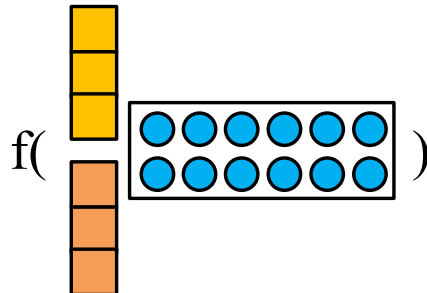
$$s(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^\top \mathbf{W} \mathbf{v}_2$$



linear interactions

$$s(\mathbf{h}_{x_i}, \mathbf{h}_{y_j}) = \mathbf{h}_{x_i}^\top \mathbf{W} \mathbf{h}_{y_j} \rightarrow \text{enormous matching matrix}$$

- **Single Layer Network (SLN)** [Collobert and Weston, 2008]



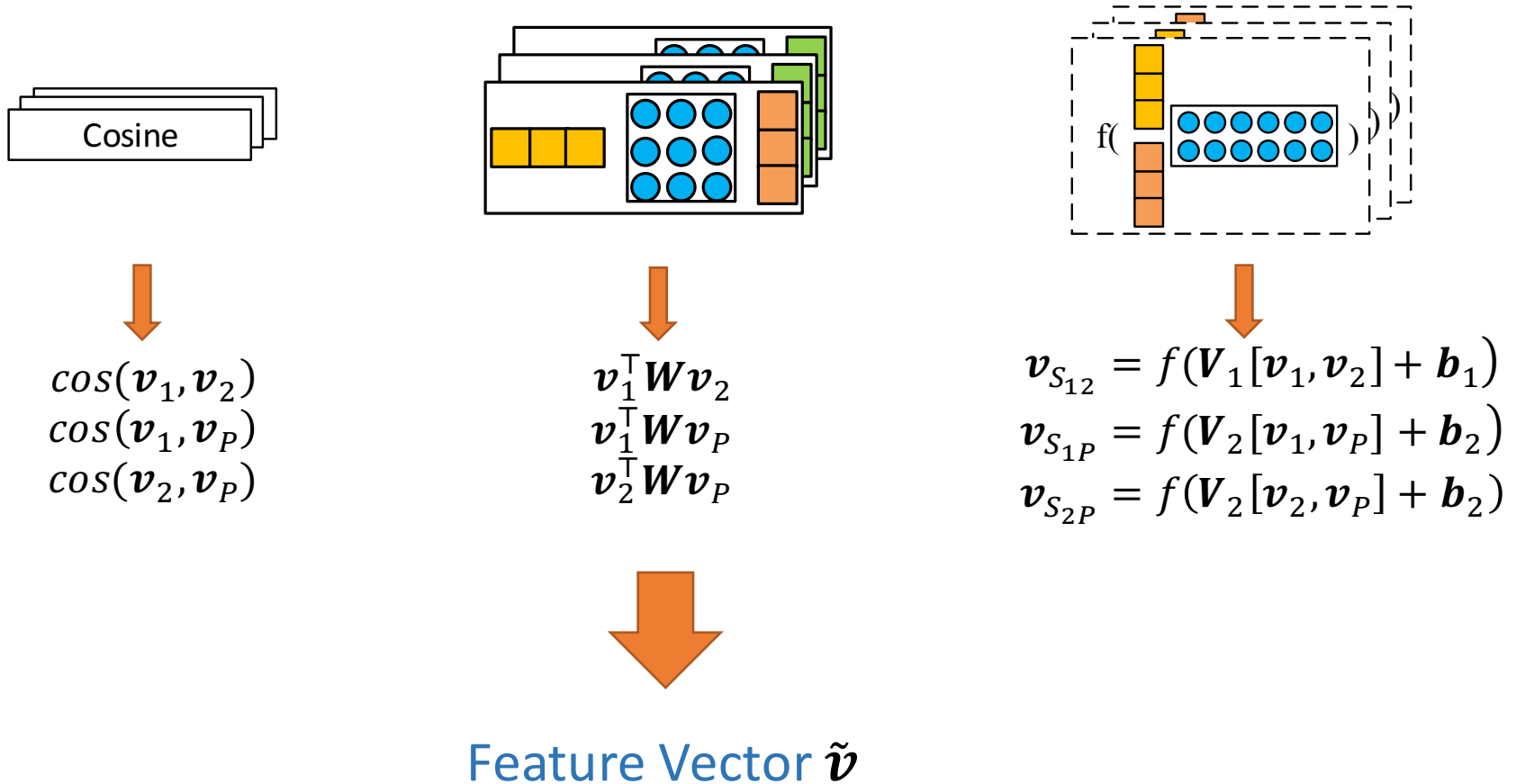
$$s(\mathbf{v}_1, \mathbf{v}_2) = f(\mathbf{V}[\mathbf{v}_1, \mathbf{v}_2] + \mathbf{b})$$



Nonlinear interactions

# Text Matching

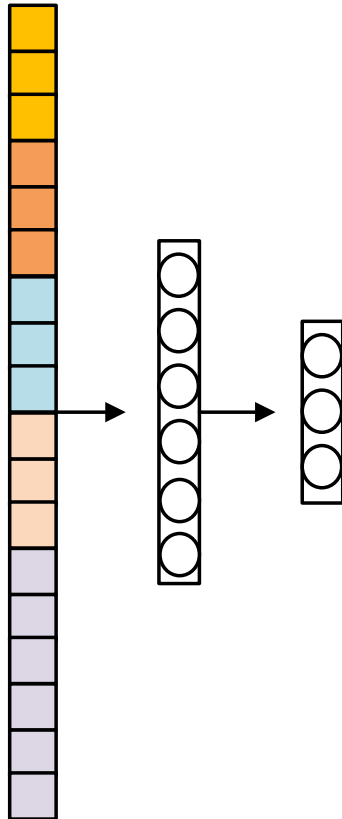
- Among discourse units and between the unit and the paragraph.





# Nuclearity Classification

- Two-layer feedforward neural network
- The input vector is first sent to a nonlinear transformation and then fed into a standard softmax layer.



$$\mathbf{t} = \text{Relu}(\mathbf{W}_t \tilde{\mathbf{v}} + \mathbf{b}_t)$$
$$\tilde{\mathbf{y}} = \text{softmax}(\mathbf{W}_s \mathbf{t} + \mathbf{b}_s)$$

# Experimental Results

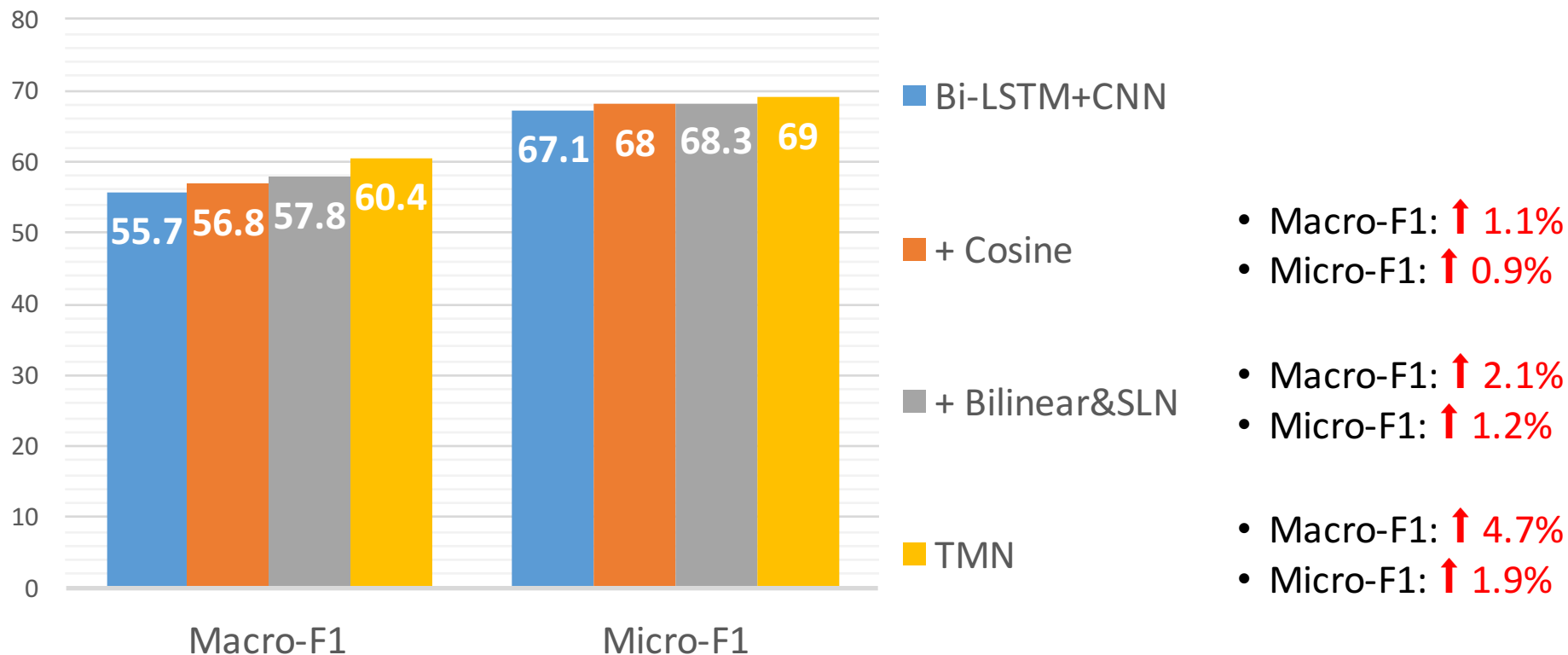
- We evaluate our model on CDTB, report the F1 on each nuclearity type and also give the micro-average and macro-average F1.

Model		NS	SN	NN	Macro-F1	Micro-F1
Neural Network Models	Kong and Zhou, 2017	20.5	21.8	75.0	42.3	60.5
	Bi-LSTM with Attention	49.7	35.7	77.2	54.4	65.2
	Bi-LSTM+CNN	52.1	35.5	78.0	55.7	67.1
	Li et al., 2016	53.6	40.2	77.5	57.2	66.3
	<b>Ours</b>	<b>54.8</b>	<b>43.6</b>	<b>79.6</b>	<b>60.4</b>	<b>69.0</b>

# Analysis: Compare with Bi-LSTM+CNN

- Justify our first hypothesis:

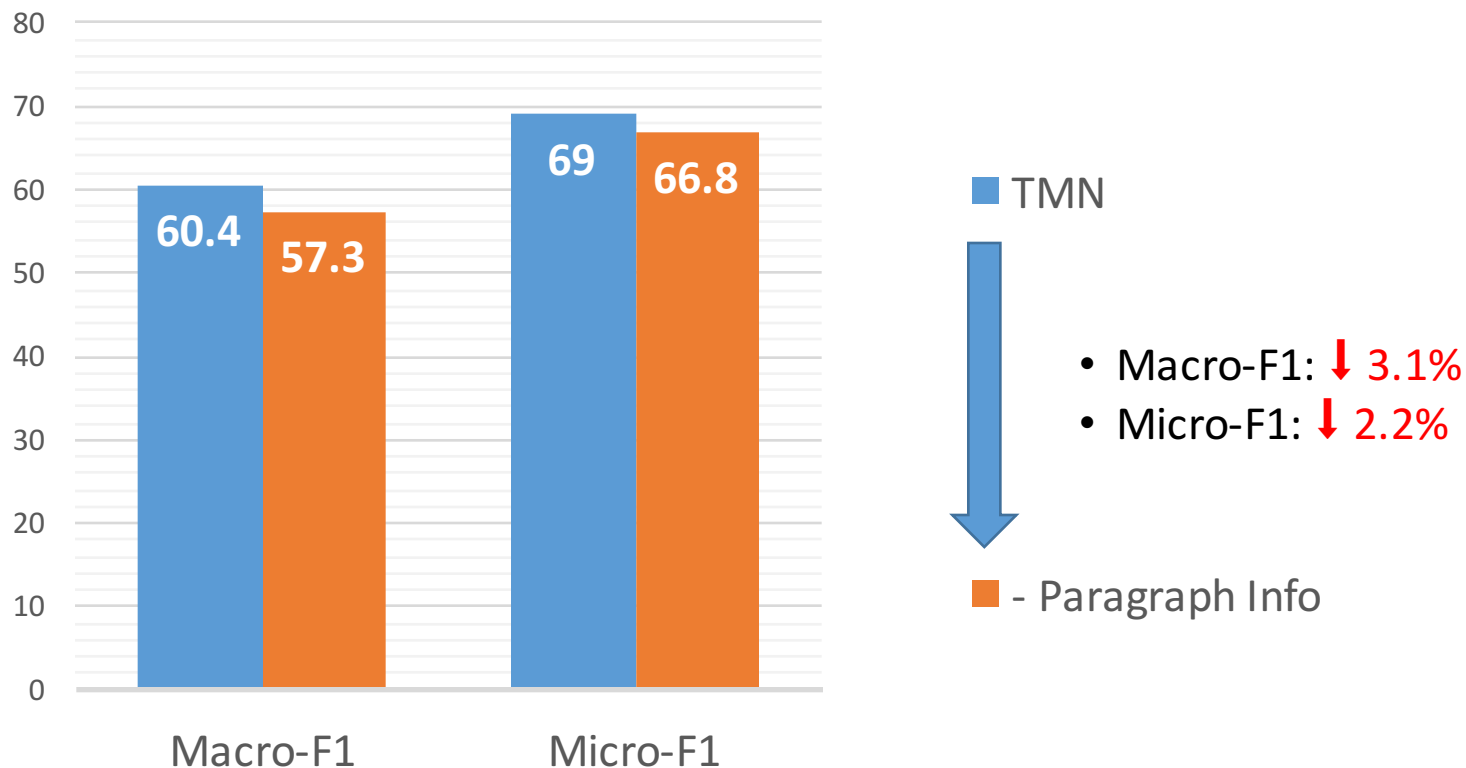
“There are strong correlations between the nuclearity and the semantic similarity or the interactions of two different discourse units”



# Analysis: remove the input of paragraph

- Justify our second hypothesis:

“Nuclearity recognition is relevant to the topic of the paragraph”



# Conclusion

- **Summary:**

- A Text Encoding module to capture both the global dependency information and the local n-gram information via Bi-LSTM and CNN.
- A Text Matching module to capture various similarities and interactions by Cosine, Bilinear and Single Layer Network.
- Experimental results on the CDTB corpus show that our TMN model is effective.

# Thank You!

Contact: [sxu@stu.suda.edu.cn](mailto:sxu@stu.suda.edu.cn)