

Few-shot Named Entity Recognition with Self-describing Networks

Jiawei Chen^{1,3,*}, Qing Liu^{1,3,*}, Hongyu Lin^{1,†}, Xianpei Han^{1,2,4,†}, Le Sun^{1,2}

¹Chinese Information Processing Laboratory ²State Key Laboratory of Computer Science

Institute of Software, Chinese Academy of Sciences

³University of Chinese Academy of Sciences

⁴Beijing Academy of Artificial Intelligence

Few-shot Named Entity Recognition

■ Few-shot named entity recognition (FS-NER)

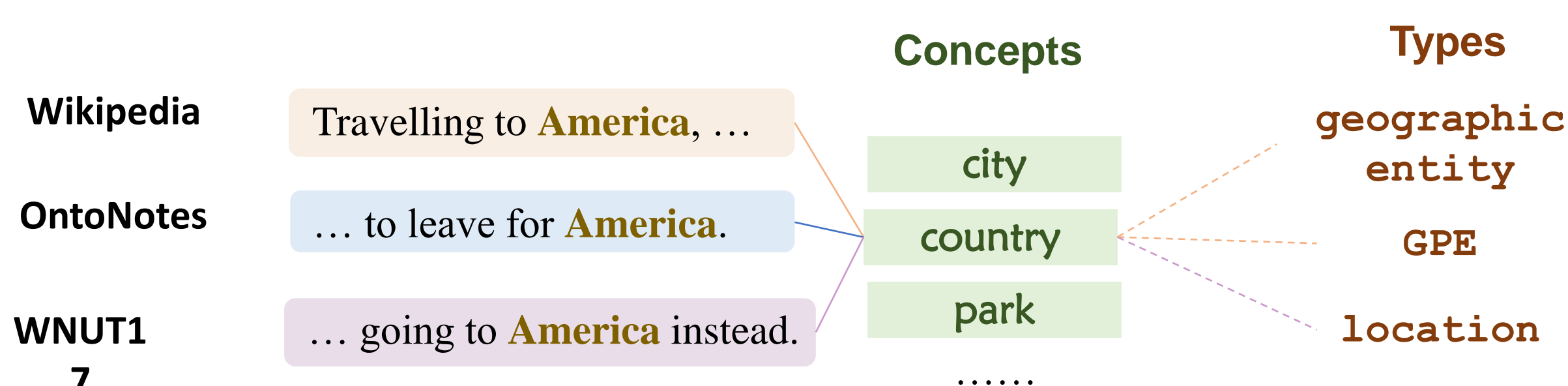
- Identify entity mentions corresponding to new entity types with only a few illustrative examples.

■ Challenge

- Limited information challenge.
- Knowledge mismatch challenge

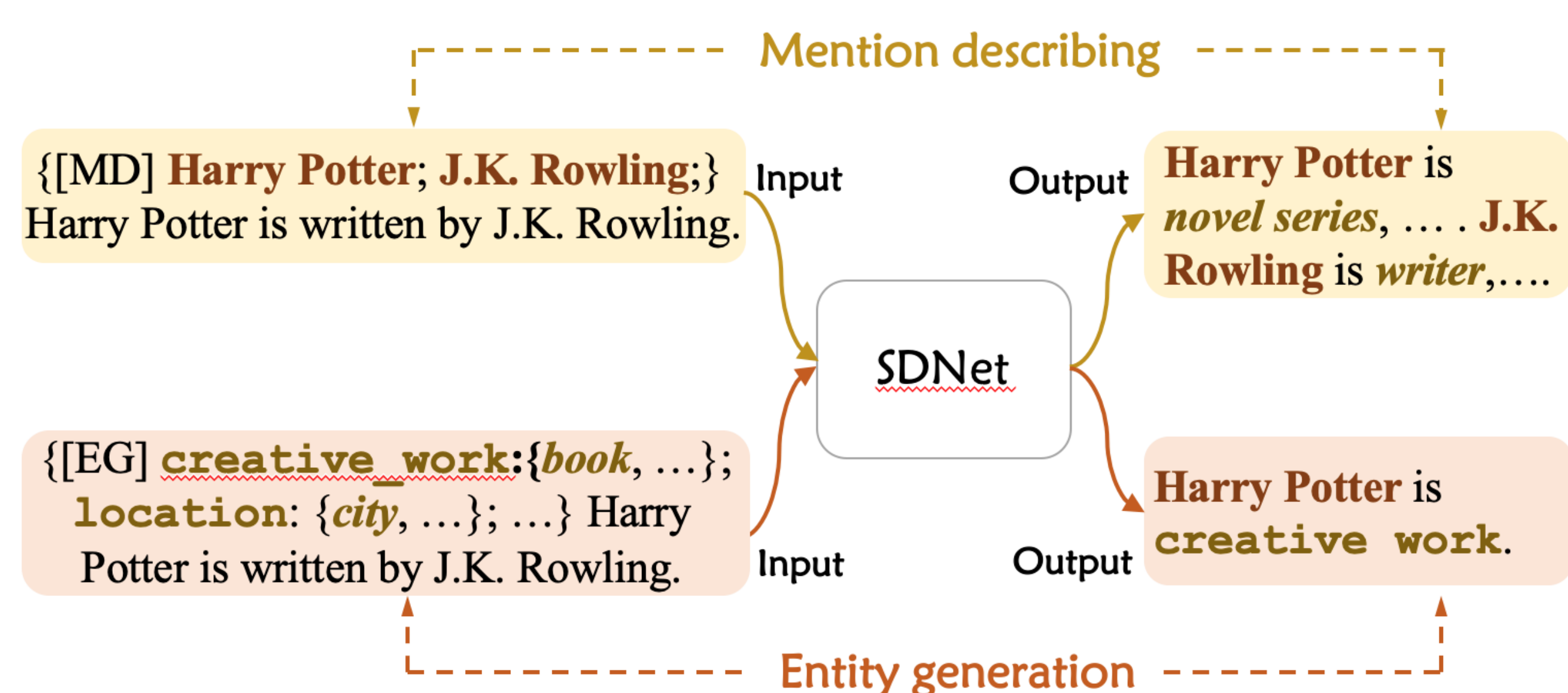
Self-describing Networks for FS-NER

- Knowledge mismatch challenge can be resolved by **uniformly describing different entity types using the same concept set.**



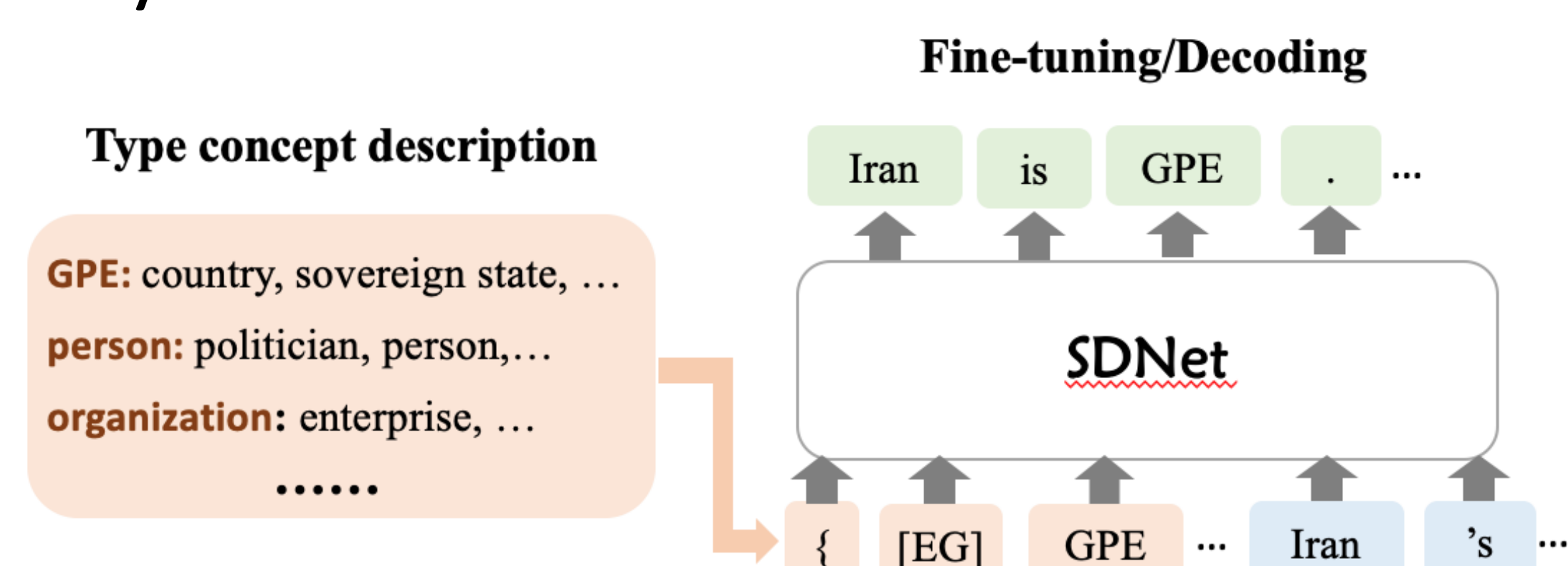
■ Self-describing Networks (SDNet):

- A Seq2Seq generation network
- Two tasks of SDNet:**
 - Entity generation
 - Adaptively generate entity mentions
 - Mention describing
 - Generate concept descriptions of mentions



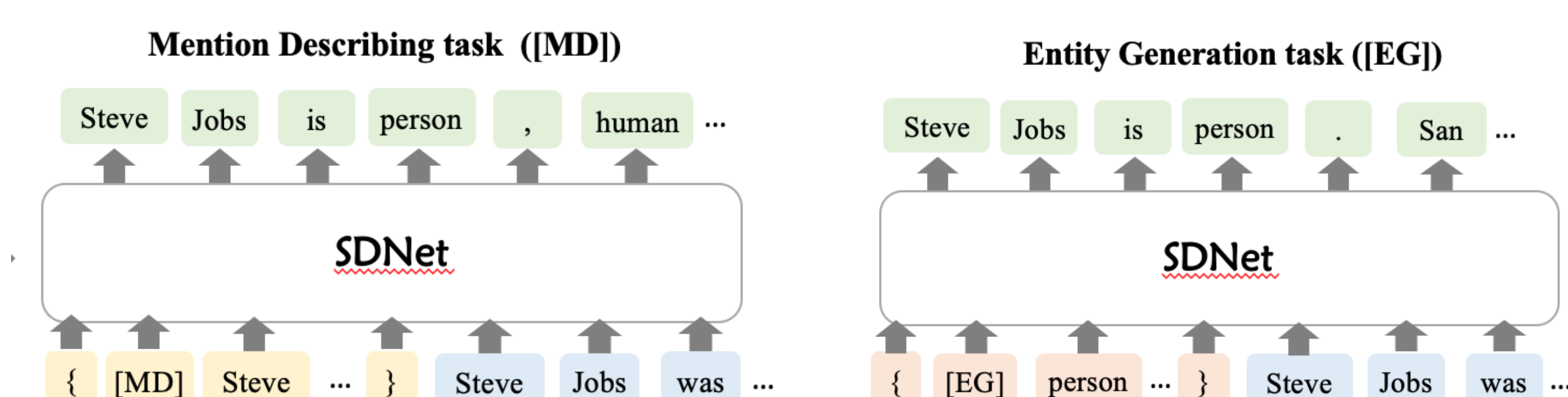
- After mention describing, SDNet summarizes the generated concepts to describe the precise semantics of specific novel types.

- The type descriptions are used to generate entity mentions.

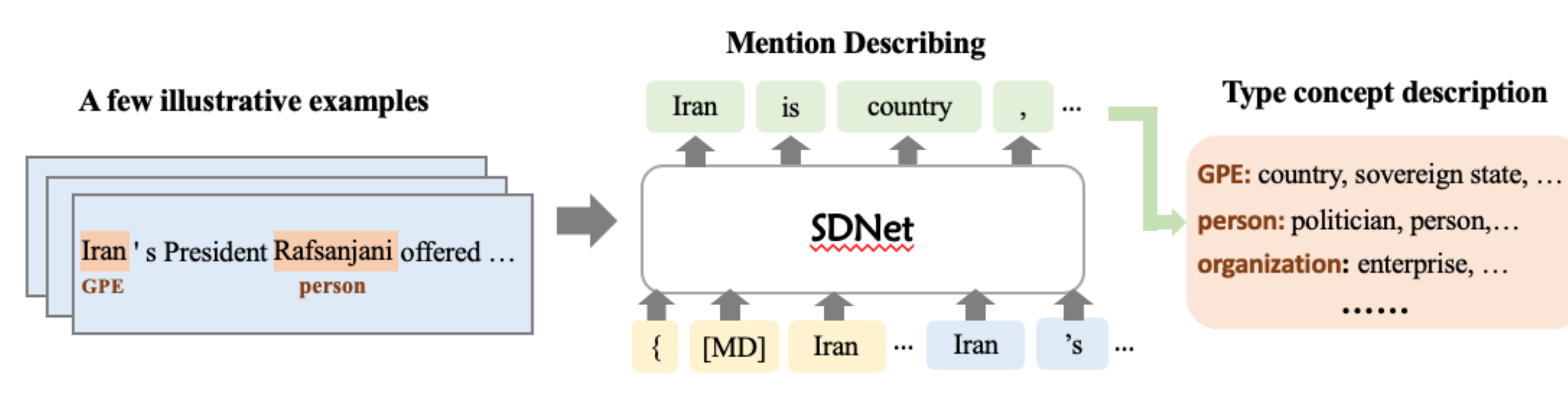


SDNet pre-training and fine-tuning

- SDNet is pre-trained by Wikipedia and Wikidata. The two tasks are pre-trained jointly.



- Before fine-tuning or decoding, SDNet will automatically build type descriptions using illustrative instances.



Experimental Results

		CoNLL	WNUT	Res	Movie1	Movie2	Re3d	I2B2	Onto	AVE
Baselines	RoBERTa (Huang et al., 2020)	53.5	25.7	48.7	51.3	/	/	36.0	57.7	/
	RoBERTa-DS (Huang et al., 2020)*	61.4	34.2	49.1	53.1	/	/	38.5	68.8	/
	Proto (Huang et al., 2020)	58.4	29.5	44.1	38.0	/	/	32.0	53.3	/
	Proto-DS (Huang et al., 2020)*	60.9	35.9	48.4	43.8	/	/	36.6	57.0	/
	spanNER (Wang et al., 2021)	71.1	25.8	49.1	/	65.4	/	/	67.3	/
	spanNER-DS (Wang et al., 2021)*	75.6	38.5	51.2	/	67.8	/	/	71.6	/
Baselines [in-house]	Bert-base	58.6	23.2	47.6	52.4	66.3	57.0	47.6	61.1	51.7
	T5-base	60.0	36.6	59.4	57.9	69.9	57.1	39.9	62.0	55.3
	T5-base-prompt	55.4	34.2	58.4	58.7	67.1	60.7	61.8	59.8	57.0
	T5-base-DS	68.2	34.9	59.7	58.4	70.8	56.0	34.1	58.8	55.1
Ours	SDNet	71.4	44.1	60.7	61.3	72.6	65.4	64.3	71.0	63.8

- By universally modeling and pre-training NER knowledge in a generation architecture, the self-describing network can effectively handle few-shot NER.
- Due to the limited information problem, transferring external knowledge to fewshot NER models are critical
- Due to the knowledge mismatch, it is challenging to transfer external knowledge effectively to novel downstream types..

Conclusions

- We propose SDNet, a Seq2Seq generation model for few-shot named entity recognition.
- SDNet can automatically map novel entity types to concepts, and adaptively recognize entities on-demand..
- SDNet is pre-trained by easily available and large-scale Wikipedia and Wikidata data.