Stage-wise Fine-tuning for Graph-to-Text Generation

Qingyun Wang[1*], **Semih Yavuz**[2], **Xi Victoria Lin**[3],
Heng Ji[1], **Nazneen Fatema Rajani**[2]

[1] University of Illinois at Urbana-Champaign [2] Salesforce Research [3] Facebook AI

`syavuz,nazneen.rajani@salesforce.com`
`victorialin@fb.com`
`{qingyun4,hengji}@illinois.edu`

A Hyperparameters and Statistics of the Model

	Origin[1]	+ Position
BART-base	139.42M	139.43M
distil-BART-xsum	305.51M	305.53M
BART-large	406.29M	406.31M
T5-base	222.88M	222,90M
T5-large	737.64M	737.65M

Table 1: # of Model Parameters

Our model is built based on the Huggingface framework (Wolf et al., 2020)[2]. Because the average lengths for source and target text in the training set are 31 and 22 words respectively, we set the maximum length for both source and target to 100 words. For T5 preprocessing, we prepend "translate RDF to English:" before the input. For BART-base, distil-BART-xsum, and T5-base, we use a batch size of 32 and train the model. We use a batch size of 16 for Bart-large, and 6 for T5-large. We use the Adam optimizer (Kingma and Ba, 2015) to optimize each model with learning rate of 3×10^{-5} with $\epsilon = 1 \times 10^{-8}$ for a maximum of 10 epochs. We run each experiment on one Nvidia Tesla V100 GPU with 16G DRAM. We first fine-tuned the PLMs on crawled Wikipedia pairs for 3 epochs. The Wikipedia Fine-tuning stage takes about 24 hours for T5-large and 10 hours for the rest of models. The final WebNLG fine-tuning stage takes less than 1 hour for all the models. We chose our best model based on multi-BLEU score[3]. For inference, we use beam search with beam size in the range {3,5}. Table 1 shows the number of the parameters for each pre-trained model.

[2]https://github.com/huggingface/transformers

[3]https://gitlab.com/webnlg/webnlg-baseline/-/blob/master/multi-bleu.perl

[3]# of parameters are slightly different because we add special tokens to the vocabulary

B   Sample Generation Results

Category	Output
Input	S| Aaron Turner P| associatedBand/associatedMusicalArtist O| Twilight (band)
	S| Aaron Turner P| associatedBand/associatedMusicalArtist O| Lotus Eaters (band)
	S| Aaron Turner P| genre O| Sludge metal
	S| Aaron Turner P| origin O| Massachusetts
	S| Aaron Turner P| activeYearsStartYear O| 1995
Reference	Aaron Turner was born in Massachusetts and started performing in 1995 . He formerly played with the band Twilight and is now in the Lotus Eaters . He is is closely associated with the Sludge metal genre.
T5-large	Aaron Turner, who is from Massachusetts , is a member of the Twilight band and the Lotus Eaters band. He is also associated with the Sludge metal genre and started playing in 1995 .
T5-large + Wiki	Aaron Turner was born in Massachusetts and started playing in 1995 . He is associated with the band Twilight and the band Lotus Eaters . His music style is Sludge metal .
T5-large + Position	Aaron Turner, who was born in Massachusetts , is associated with Twilight and Lotus Eaters . He is a member of Sludge metal band and has also been a musical artist. [1995]
T5-large + Wiki + Position	Aaron Turner was born in Massachusetts and started playing Sludge metal in 1995 . He is also associated with Twilight and Lotus Eaters .

Table 2: Human and System Generated Description We use the color box to frame each entity out with the same color as the corresponding entity. We highlight *fabricate facts*, [missed relations], and incorrect relations with different color.

Category	Output
Input	S| American Civil War P| commander O| Jefferson Davis
	S| Alfred Moore Scales P| country O| United States
	S| United States P| ethnicGroup O| Native Americans in the United States
	S| Alfred Moore Scales P| battles O| Siege of Petersburg
	S| Siege of Petersburg P| isPartOfMilitaryConflict O| American Civil War target
Reference	American Civil War , Siege of Petersburg involved U.S. born Alfred Moore Scales. The American Civil War was commanded by Jefferson Davis . An ethnic group of the United States includes Native Americans .
T5-large	Alfred Moore Scales is from the United States where Native Americans are one of the ethnic groups. He fought in the Siege of Petersburg which was part of the American Civil War commanded by Jefferson Davis .
T5-large + Wiki	**The American Civil War was fought in the** Siege of Petersburg **.** Jefferson Davis was the commander of the war. Alfred Moore Scales was born in the United States where Native Americans are an ethnic group. [S| Alfred Moore Scales P| battles O| Siege of Petersburg]
T5-large + Position	Alfred Moore Scales was born in the United States , where Native Americans are an ethnic group. He fought in the American Civil War, which was led by Jefferson Davis . The Siege of Petersburg is part of the American Civil War . [S| Alfred Moore Scales P| battles O| Siege of Petersburg]
T5-large + Wiki + Position	Alfred Moore Scales is from the United States where Native Americans are one of the ethnic groups. He fought in the Siege of Petersburg which is part of the American Civil War . Jefferson Davis was the commander of the American Civil War.

Table 3: Human and System Generated Description We use the color box to frame each entity out with the same color as the corresponding entity. We highlight *fabricate facts*, [missed relations], and **incorrect relations** with different color.

Category	Output
Input	S| 101 Helena P| discoverer O| James Craig Watson
	S| James Craig Watson P| almaMater O| University of Michigan
	S| 101 Helena P| discovered O| 1868-08-15
	S| James Craig Watson P| nationality O| Canada
	S| James Craig Watson P| deathPlace O| Madison, Wisconsin
Reference	James Craig Watson, who discovered 101 Helena on 15th August 1868 , is a Canadian national who attended the University of Michigan . He died in Madison, Wisconsin .
T5-large	James Craig Watson is a Canadian who graduated from the University of Michigan . He was the discoverer of 101 Helena which was discovered on 15 August 1868 . He died in Madison, Wisconsin .
T5-large + Wiki	James Craig Watson, a Canadian , graduated from the University of Michigan and discovered 101 Helena on August 15th, 1868 . He died in Madison, Wisconsin .
T5-large + Position	James Craig Watson, a Canadian , graduated from the University of Michigan and was the discoverer of 101 Helena on August 15, 1868 . He died in Madison, Wisconsin .
T5-large + Wiki + Position	James Craig Watson, a Canadian , graduated from the University of Michigan and discovered 101 Helena on August 15th, 1868 . He died in Madison, Wisconsin .

Table 4: Human and System Generated Description We use the color box to frame each entity out with the same color as the corresponding entity. We highlight *fabricate facts*, [missed relations], and **incorrect relations** with different color.

References

Diederik P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.