

Appendix: Concept Pointer Network for Abstractive Summarization

Wang Wenbo¹, Gao Yang^{1,2*}, Huang Heyan^{1,2} and Zhou Yuxiang¹

¹ School of Computer Science and Technology, Beijing Institute of Technology

² Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications

gyang@bit.edu.cn

A Examples of Microsoft Concept Graph

Table 1: $p(c|x)$ is the probability of a concept c that is for word x . This table shows the concept distribution for word “Microsoft” and “apple”.

x	c	$p(c x)$
Microsoft	company	0.611
	vendor	0.089
	client	0.048
	firm	0.045
	large company	0.043
	organization	0.043
	corporation	0.036
	brand	0.034
	software company	0.028
	technology company	0.024
apple	fruit	0.415
	company	0.286
	food	0.076
	brand	0.05
	fresh fruit	0.049
	fruit tree	0.037
	crop	0.032
	corporation	0.018
	manufacturer	0.018
	firm	0.017

In this section, we give several examples to show how the Microsoft Concept Graph provide a distributional set of candidate concepts for a specific word in our model. Suppose we have words like “Microsoft” and “apple”, Table 1 shows concept distributions for these words. Formally, the concept graph specifies the probability that each instance x belongs to a concept c , $p(c|x)$. Given a word x_i , a set of k concept candidates, $C_i = \{c_i^1, c_i^2, \dots, c_i^k\}$, is linked to the word from the knowledge base, with distributional

probabilities over the concepts, i.e., $P(C|x_i) = \{p(c_i^1), p(c_i^2), \dots, p(c_i^k)\}$. These statistics are derived from the public API <https://concept.research.microsoft.com/Home/API>.

B Generated Summary Examples

This appendix provides example source articles from the test set, with side-by-side comparisons of the human generated (reference) summaries and the summaries produced by baseline models and our models. We provide two representative baselines which are **seq2seq+att** and **pointer-generator** models.

To demonstrate the performance of our proposed concept pointer network, we compare it with the baselines in Table 2. In Table 3, we further show examples of the concept pointer with reinforcement learning training strategy and distant supervision adaption of our models.

In Table 2, we highlight those better expressions generated by the proposed concept pointer model. As the NO.1 examples show, the summary generated by the seq2seq+att model lacks of “first writer” and the pointer-generator doesn’t have “writer” either. While, our model generates a complete meaning of the source article and the summary is most semantically close to what human generate (reference). Similar advantages are also shown in NO.2. Comparing with the baseline seq2seq+att and pointer-generator in the NO.3 examples, our model can generate a concept word “agreement” to the summary, which can make a deeper understanding of the source article. In the NO.4 examples, seq2seq+att has the problem of $\langle \text{unk} \rangle$ token while pointer-generator generates factual errors, such as “during Asian games”. The source article uses “upcoming Asian games”, and our model replaces the phrase with a new one “ahead of” while maintaining the same meaning.

* Corresponding author

In Table 3, we find that the RL training strategy and DS training strategy can assist the model to obtain new and longer, more fluent and read-

able summaries compared to the concept pointer model, which are highlighted in the table.

Table 2: Examples of the generated summaries by seq2seq+att, pointer-generator and concept pointer, respectively.

NO.1: <i>Source article</i>
Saramago became the first writer in Portuguese to win the Nobel prize for literature on Thursday, his personal delight was seconded by a burst of public elation in his homeland.
<i>Reference summary</i>
Saramago is the first writer in the Portuguese language to win Nobel.
<i>Generated summary</i>
seq2seq+att: Portuguese writer Saramago wins Nobel for literature. pointer-generator: Saramago first wins Nobel prize for literature in Portugal. Concept pointer: Saramago becomes first writer in Portuguese to win Nobel literature prize.
NO.2: <i>Source article</i>
Cambodia’s two-party opposition asked the Asian development bank Monday to stop providing loans to the incumbent government , which it calls illegal.
<i>Reference summary</i>
opposition asks end to loans to “illegal” Cambodian government.
<i>Generated summary</i>
seq2seq+att: Cambodian opposition asks ADB to stop giving government loans pointer-generator: Cambodian opposition asks Asian development bank to stop providing loans to embattled government. Concept pointer: Cambodian opposition asks ADB to stop giving loans to incumbent government
NO.3: <i>Source article</i>
King Norodom Sihanouk on Tuesday praised agreements by Cambodia’s top two political parties , previously bitter rivals, to form a coalition government led by strongman Hun Sen.
<i>Reference summary</i>
King praises Cambodian coalition government of top two political parties.
<i>Generated summary</i>
seq2seq+att: Norodom Sihanouk praises Cambodian political parties. pointer-generator: King Sihanouk praises Cambodia’s coalition government to form government. Concept pointer: King Norodom Sihanouk praises agreements to form coalition government in Cambodia.
NO.4: <i>Source article</i>
Thai police have detained more than 300 beggars , most from neighboring countries, in a campaign to make Bangkok ’s streets safer for spectators and athletes arriving for the upcoming Asian games , a senior police officer said Friday.
<i>Reference summary</i>
Thai police clear streets of beggars in preparation for Asian Games.
<i>Generated summary</i>
seq2seq+att: More than [UNK] beggars detained in bangkok. pointer-generator: Thai police detain more than 300 beggars during Asian games. Concept pointer: Thai police detain more than 300 beggars ahead of Asian games.

Table 3: Comparisons of the concept pointer, concept pointer+RL and concept pointer+DS, respectively.

NO.1: <i>Source article</i>
Horses belonging to Iran 's equestrian team will not be allowed to compete in next month 's Asian games because they failed to meet the requirements of the games ' veterinary commission , the Thai organizers announced Thursday.
<i>Reference summary</i>
Horses of Iran's equestrian team flunk Asian Games test.
<i>Generated summary</i>
pointer-generator: Iran won't be allowed to compete in next month's Asian games. Concept pointer: Horses belonging to Iran will not be allowed to compete in Asian games. Concept pointer+RL: Iran's equestrian horses not allowed to compete in Asian games. Concept pointer+DS: Horses belonging to Iran's equestrian team will not be allowed to compete in Asian games.
NO.2: <i>Source article</i>
President Fernando henrique cardoso's efforts to repair the largest economy in Latin America may depend on the outcome of this weekend's gubernatorial elections.
<i>Reference summary</i>
Cardoso's economic efforts may depend on upcoming gubernatorial elections.
<i>Generated summary</i>
pointer-generator: Brazil's silva may depend on outcome of gubernatorial elections. Concept pointer: Brazilian president seeks to depend on outcome of presidential elections. Concept pointer+RL: Brazil's president's efforts to latin america may depend on local elections . Concept pointer+DS: Cardoso's efforts to repair Latin American economy may hinge on outcome of elections.