Summarization and simplification

Quiz, 5 questions

✔ **Congratulations! You passed!**

✔   1 / 1 point

1.

Which type of information is not used for evaluation in machine translation (e.g. in BLEU score), but is used for evaluation in simplification task (e.g. in SARI score)?

◯ System output

⦿ System input

Correct
Bingo!

◯ Human reference

✔   1 / 1 point

2.

Let us consider a simplification task and denote an input by I, a reference by R, and a system output by O. We have discussed several types of operations for simplification. How would you compute a precision score for the copying operation?

⦿ $\dfrac{|O \cap I \cap R|}{|O \cap I|}$

Correct
Exactly!

◯ $\dfrac{|O \cap I \cap R|}{|I \cap R|}$

◯ $\dfrac{|O \cap \bar{I} \cap R|}{|O \cap \bar{I}|}$

Summarization and simplification

3.
In the summarization video we talked about attention distribution and denoted its elements as p_i^j. How are they normalized?

- ⦿ They sum to 1 over all positions of an input sentence: $\sum_i p_i^j = 1$.

  ▲

  Correct
  True!

- ◯ They sum to 1 over all positions of an output sentence: $\sum_j p_i^j = 1$.

- ◯ They are logits, we need to apply softmax to get the normalization constraint.

✔ 2 / 2 points

4.
Imagine you have trained an encoder-decoder-attention model to generate a text summary. Let's say you have a vocabulary [big, black, bug, bear] and the vocabulary distribution at some decoding moment is [0.3, 0.4, 0.1, 0.2].

Now, let us consider how it changes if we add the pointer part from the paper "Get to the point! Summarization with pointer-generator network" to be able to copy some input words.

Consider an input sentence: "a big black bug bit a big black bear". And the attention distribution [0.1, 0.1, 0.1, 0.1, 0.2, 0.1, 0.1, 0.1, 0.1].

How will the final distribution look like, if the pointer network (copy distribution) is **weighted equally** with the generator network (vocabulary distribution)?

Enter the probability for "big".

| 0.25 |

▲

Correct Response
The generator part gives you 0.3, the copy part gives you 0.2. So the final weighted probability is 0.5 * 0.3 + 0.5 * 0.2 = 0.25.

✔ 1 / 1 point

5.

Check the correct statements about the summarization models discussed in the video.

☑ use coverage that helps to avoid repetitions of the input fragments.

Correct
Yes! We accumulate some scores about what have already been seen and try not to repeat ourselves.

☑ The pointer-generator network performs abstractive summarization.

Correct
Yes, it has an ability to generate words that dod not occur in the input due to the "generation" part of the network.

☐ The pointer-generator network with coverage trick outperforms all other baselines.

Un-selected is correct

☐ The copy mechanism encourages the model to generate new fragments that did not occur in the input.

Un-selected is correct

☐ The pointer-generator network performs extractive summarization.

Un-selected is correct