Master Degree in Computer Engineering

Natural Language Processing
Final Exam

July 20th, 2022

1. **[1 points]** Introduce the linguistic notion of compositionality and provide some examples.

2. **[5 points]** With reference to word embeddings, answer the following questions.

   (a) Introduce the basic idea underlying the skip-gram algorithm, and define the two probabilities $P(+ \mid w, u)$ and $P(- \mid w, u)$.

   (b) Describe the objective function used by the skip-gram algorithm.

3. **[6 points]** Some text T has been tokenized based on white spaces. The resulting dictionary and word frequencies are reported in the following table

word	clear	clearer	large	largest
freq	15	4	21	6

   Apply the byte pair encoding algorithm to derive subword tokens for T, using the character '_' to mark the end of each word. Report and comment each of the first eight iterations (merge operations) in your run of the algorithm, including the frequency updates.

4. **[2 points]** Explain why the sentence 'Alice sees the man with the telescope' is syntactically ambiguous, by drawing two different phrase structure representations using the following syntactic categories: S, NP, VP, PP, V, N, P, Det (article).

(see next page)

5. **[6 points]** Considering contextualized word embeddings, answer the following questions.

   (a) What is the basic difference between static word embeddings and contextualized word embeddings?

   (b) Introduce the basic architecture of the model known as ELMo (embeddings from language model).

6. **[5 points]** In the context of transition-based parsing, answer the following questions.

   (a) Define the notion of spurious ambiguity, as we have introduced it in class.

   (b) State two different sequences of transitions that make an arc-standard parser produce the projective dependency tree consisting of the following unlabeled dependency relations

head	w_3	w_1	$\langle \text{ROOT} \rangle$	w_5	w_3
dependent	w_1	w_2	w_3	w_4	w_5

7. **[6 points]** In the context of neural machine translation, answer the following questions.

   (a) How do we model the probability $P(y \mid x)$, with $x = x_1 \cdots x_n$ a sentence in the source language and $y = y_1 \cdots y_m$ a sentence in the target language?

   (b) Introduce the encoder-decoder with RNN greedy inference algorithm we have presented in the lectures, and report the model equations.

8. **[2 points]** Outline the six components of a typical dialogue-state system. In the natural language understanding component, explain the three main tasks: domain classification, intent extraction, and slot filling.