

**Natural Language Processing
Final Exam**

June 23rd, 2022

1. [1 points] Introduce Herdan’s law and discuss its relevance for natural language processing.
2. [5 points] With reference to the feedforward neural language model (FNLM), answer the following questions.
 - (a) How does FNLM approximate the probability $P(w_t | w_{1:t-1})$?
 - (b) Provide and explain the formal equations defining FNLM.
 - (c) Discuss the advantages of feedforward NLM with respect to the N -gram language model.
3. [8 points] The following transition and emission matrices specify a hidden Markov model for POS tagging in terms of negative log probabilities, also called scores. As an example, the transition from PP to PL has score 12, and the transition from PL to PP has score 2.

	PL	PN	PP	VB	⟨EOS⟩
⟨BOS⟩	11	2	3	4	19
PL	17	3	2	5	7
PN	5	4	3	1	8
PP	12	4	6	7	9
VB	3	2	3	3	7

	hen	vilar	ut
PL	17	17	4
PN	3	19	19
PP	19	19	3
VB	19	8	19

- (a) Draw a finite automaton representation of the above model.
- (b) Consider the sentence ‘hen vilar ut’. In the following table each entry represents the best score of reaching the associate word/POS pair starting from ⟨BOS⟩. Use the Viterbi algorithm to fill in the table and to calculate the most probable POS tag sequence. **Important:** since we are using negative log probabilities, the Viterbi algorithm needs to find the lowest score path.

	⟨BOS⟩	hen	vilar	ut	⟨EOS⟩
⟨BOS⟩	0	-	-	-	-
PL	-				-
PN	-				-
PP	-				-
VB	-				-
⟨EOS⟩	-	-	-	-	

(see next page)

4. **[2 points]** In the context of phrase structures, introduce the notions of argument and modifier and provide some examples.
5. **[7 points]** Consider the English sentence ‘I have found the picture of Alice.’ along with the projective dependency tree consisting of the following unlabeled dependency relations

head	found	found	⟨ROOT⟩	picture	found	Alice	picture	found
dependent	I	have	found	the	picture	of	Alice	.

Answer the following questions.

- (a) Draw a graphical representation of the dependency tree above, with arcs directed from the head to the dependent.
- (b) Apply to the above tree the arc-standard oracle presented in class, to construct a sequence of training instances.
6. **[2 points]** In the context of machine translation, introduce the BLEU evaluation metric and explain why this measure is criticised nowadays.
7. **[8 points]** Consider the end-to-end application of text-based question answering, and answer the following questions.
- (a) Introduce the machine reading task and the notions of query, passage and span. Define the span probability $P(p_i, \dots, p_j | q, p)$ and explain how this probability is approximated by means of start and end probabilities.
- (b) Introduce the Stanford attentive reader for machine reading, and discuss the main representation and equations used by this neural model.