

MT, SS 21: Problem Set 2

Sebastian Pado

April 27, 2021

This problem set is due on May 4th (10am), please submit as a PDF file through ILIAS. It will be discussed in the tutorial on May 6th. Please include the names of all team members. Feel also free to bring any other questions that you might have. Ideally, submit them to the Wiki beforehand.

Problem 1) Word Alignments

Suggest word alignment **functions** (English \rightarrow German) for the following sentence pairs. Discuss the cases where word translation is not one-to-one.

1. My₁ brother₂ is₃ tall₄ and₅ he₆ is₇ blond₈. – Mein₁ Bruder₂ ist₃ groß₄ und₅ hat₆ blondes₇ Haar₈.
 2. Balloons₁ scare₂ me₃. – Ich₁ fürchte₂ mich₃ vor₄ Ballons₅. (“I fear myself of balloons”)
 3. beef₁ packaging₂ regulation₃ – Rindfleischverpackungsverordnung₁
-

Problem 2) Word Alignments

Consider the following alignment functions.

- Gold alignment: $\{ 1 \rightarrow 3, 2 \rightarrow 4, 3 \rightarrow 4, 4 \rightarrow 5, 5 \rightarrow 1 \}$.
- Predicted alignment 1: $\{ 1 \rightarrow 3, 1 \rightarrow 4 \}$.
- Predicted alignment 2: $\{ 1 \rightarrow 3, 2 \rightarrow 4, 3 \rightarrow 4, 4 \rightarrow 5 \}$.

Compute precision, recall, and balanced F-Score for both predictions.

Problem 3) Combinatorics of Alignment Functions

a) Assume that you have two sets X, Y with cardinalities $|X|$ and $|Y|$. How many mappings are there that define a total function from X to Y ?

b) In this light, what is the interpretation of the denominator $(1 + l_e)^{l_f}$ in the definition of $P(f, a|e)$ in IBM Model 1? Where does the +1 come from?

c) Why does this quantity occur in the denominator?

d) In a bilingual sentence pair (s_1, s_2) , are there the same number of alignment functions from s_1 to s_2 than from s_2 to s_1 ? (Ignore NULL tokens.) If yes, why? If no, why not, and are there conditions under which this is true?
