

1. Assume a given tagger performs full disambiguation, i.e. it assigns a single tag to each token. What is the formula for the *accuracy* of the tagger?

$$\text{accuracy} = \frac{\# \text{ correctly tagged tokens}}{\text{total number of tokens}}$$

2. What is a *base tagger*?

A tagger which assigns the most frequent tag to each token.

3. What is a *data-driven tagger*?

A tagger which learns its tagging model from data, i.e. from a tagged corpus. A data-driven tagger is (usually) language independent.

4. Statistical taggers are often based on a Hidden Markov Model (HMM). In the context of PoS tagging, what is **hidden** and what is **observed** (known)?

The PoS tags are hidden, the words are observed.

5. The formula below is used in statistical taggers based on a HMM. Which part stands for the *transition/contextual* probabilities and which part stands for the *emission/lexical probabilities*?

$$P(t_1)P(t_2|t_1) \prod_{i=3}^n P(t_i|t_{i-2}, t_{i-1}) \prod_{i=1}^n P(w_i|t_i)$$

Transition/contextual: $P(t_1)P(t_2|t_1) \prod_{i=3}^n P(t_i|t_{i-2}, t_{i-1})$

Emission/lexical: $\prod_{i=1}^n P(w_i|t_i)$