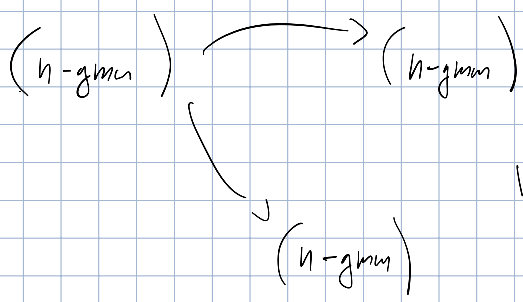
Zadání úkolu:

- 1. část: ez, pouze hraní si s brill taggerem, embrace toho taggeru
- 2. část:
  HMM tagger, build and train
  - one state is the n-gram

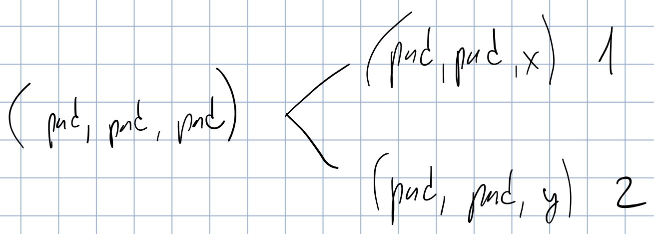


I need to count the transitions between n-grams, make probabilities from them

later run viterbi.

But how to smooth? Well for each (z|x,y), I can also check (z|x) and (z) and give weights to each model

→ actually I will pose „question" to all n-gram models and combine their outputs.



<PAD> <PAD> # Dobrý den pane Oliver.
<PAD> <PAD> # Adj

$P(\# \mid \langle PAD \rangle, \langle PAD \rangle) = 1$

$P(Adj \mid \langle PAD \rangle, \#) = 0{,}25$

$P(Dobrý \mid \# \; Adj)$

Chci: $P(\text{nový stav} \mid \text{slovo, předchozí stav})$

$P(\text{nový stav} \mid \text{předchozí stav})$ → to umím vypozorovat jako transition probability

→ já vždy hledám takový nový stav N: $\underset{N}{\text{argmax}} \; P(\text{slovo} \mid N) \cdot P(N \mid M)$

$P(\text{slovo} \mid N)$ — emission prob
- fallback: uniform

$P(N \mid M)$ — transition prob
- fallback: ?uniform?

get_transitions(alpha[7])

alpha[7]

$\beta(s', 7) = \beta(s, 8) \cdot p(s \mid s') \cdot p(obs[8] \mid s)$