Welcome to the Statistical Methods of Language Technologyb SoSe21 course

Dr. Seid Muhie Yimam

- Email: yimam@informatik.uni-hamburg.de (mailto:yimam@informatik.uni-hamburg.de)
- Office: Informatikum, F-415

Dr. Özge Alaçam

- Email: alacam@informatik.uni-hamburg.de (mailto:alacam@informatik.uni-hamburg.de)
- Office: Informatikum, F-435

Topic of this week;

In this first practice class, we are going to focus on three main topics, which will be useful to complete 1st assignment;

- HMM models
- Viterbi Algorithm

Deadline: 12/17 May

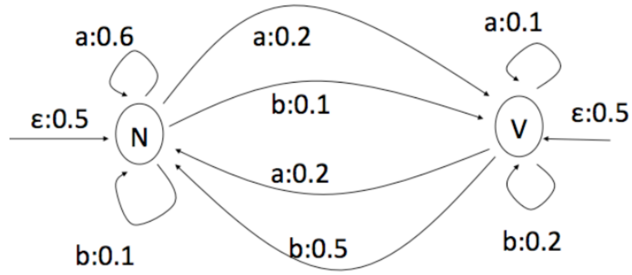
HMM

check this link for more hands-on practice on coding backward and forward HMM models and Viterbi algorithm from scratch!
https://notebook.community/danijel3/ASRDemos/notebooks/HMM_FST
(https://notebook.community/danijel3/ASRDemos/notebooks/HMM_FST)

In class Exercises

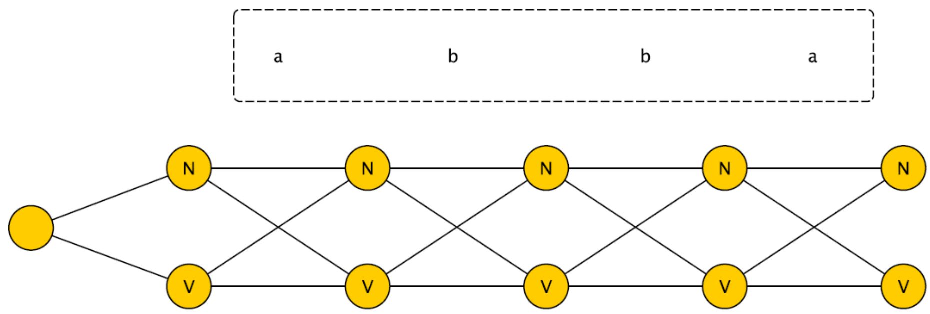
Problem 4.1 HMM: Probability of observing a sequence

Consider the HMM given



a) For the sequence **abba** , compute the probability $P(abba)$ of the observation using the **forward** procedure

HINT: fill this HMM diagram

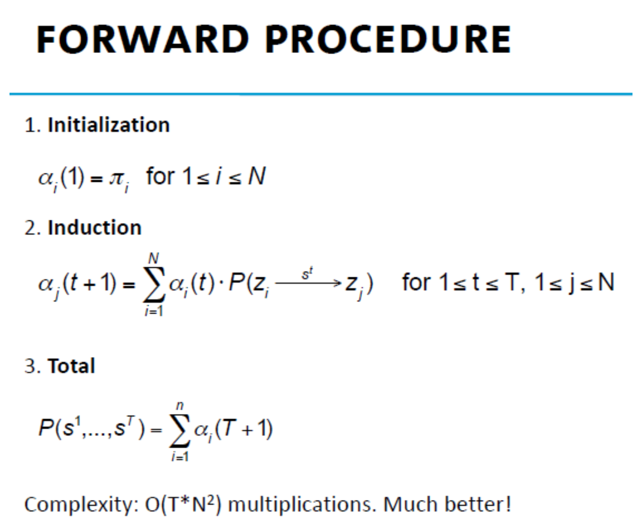


hmm1_empty

t	α**N**	α**V**
1	0.5	0.5
2		
3		
4		
5		

$P(abba) = ?$ \

Hint (Lecture notes, slide 20):



b) For the sequence **abba** , compute the probability $P(abba)$ of the observation using the

backward procedure

t	βN	βV
1	1.0	1.0
2		
3		
4		
5		

$P(abba) = ?$ \

Hint (Lecture notes, slide 21):

BACKWARD PROCEDURE

1. **Initialization**

$$\beta_i(T+1) = 1 \quad \text{for } 1 \le i \le N$$

2. **Induction**

$$\beta_j(t) = \sum_{i=1}^{N} \beta_i(t+1) \cdot P(z_j \xrightarrow{s^t} z_i) \quad \text{for } 1 \le t \le T,\ 1 \le j \le N$$

3. **Total**

$$P(s^1, \ldots, s^T) = \sum_{i=1}^{n} \pi_i \beta_i(1)$$

Problem 4.2 HMM: Best state sequence

For the HMM given in 4.1, what is the sequence of states that marks the path with the highest probability for the input sequence abba, and what is the probability?

HINT: Viterbi algorithm
Lecture Notes, slide 24:

VITERBI ALGORITHM

1. **Initialization**

$$\delta_i(1) = \pi_i \quad \text{for } 1 \le i \le N$$

2. **Induction**

$$\delta_j(t+1) = \max_{i=1..N} \delta_i(t) * P(z_i \xrightarrow{s_t} z_j) \quad \text{for } 1 \le j \le N$$

store backtrace: per state *j*, memorize the previous state for $\delta_j(t+1)$

$$\psi_j(t+1) = \operatorname*{argmax}_{i=1..N} \delta_i(t) * P(z_i \xrightarrow{s_t} z_j) \quad \text{for } 1 \le j \le N$$

3. **Termination**:

last state in *MAXPATH*: $z_{\max}^{T+1} = \operatorname*{argmax}_{i=1..N} \delta_i(T+1)$

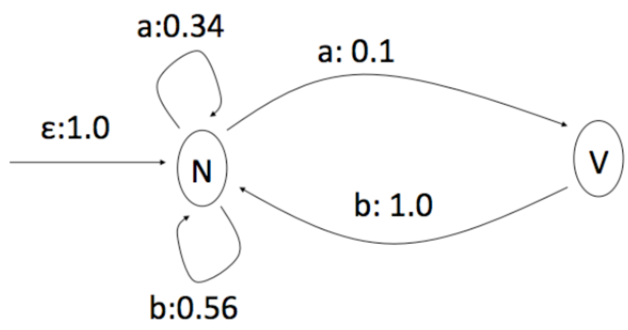
read sequence according to: $z_{\max}^{t} = \psi_{z_{\max}^{t+1}}(t+1)$

Ties: resolve randomly or store n-best-list.

HMM: Training

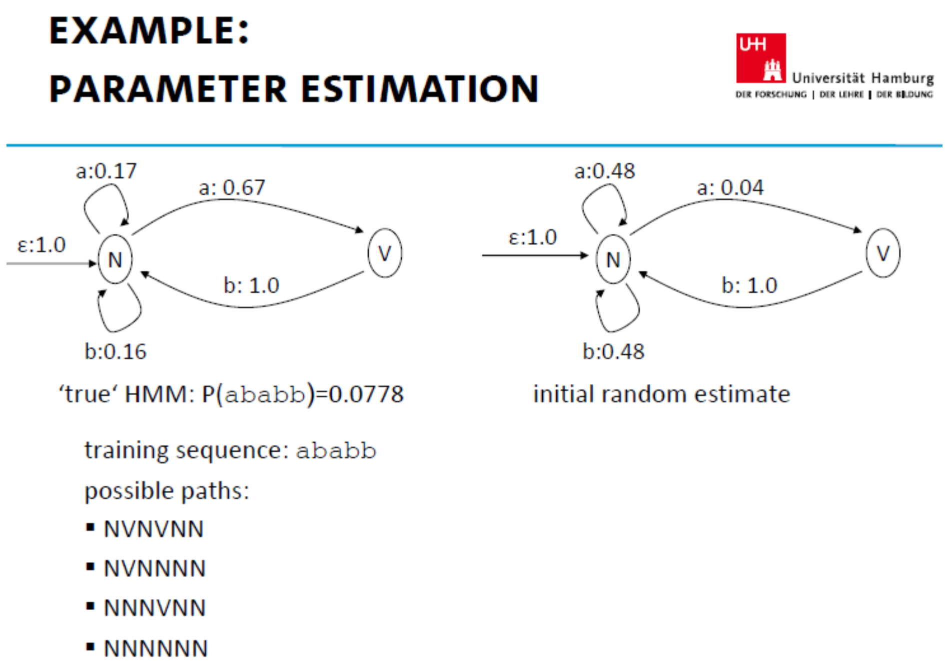
In the lecture, we trained an HMM using sequence ababb. We arrived at the following:

Iteration 2 (HMM2):



Training a Markov Chain

Lecture Notes, slide 33-34:



TWO ITERATIONS

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

training sequence: `ababb`

1.

path	P(path)	P(V|a,N)	P(N|b,V)	P(N|a,N)	P(N|b,N)	sum P(.|.,N)
NVNVNN NVNNNN NNNVNN NNNNNN	0.00077 0.00442 0.00442 0.02548	0.00154 0.00442 0.00442 0.0	0.00154 0.00442 0.00442 0.0	0.0 0.00442 0.00442 0.05096	0.00077 0.00884 0.00884 0.07644	
sum over paths	0.03509	0.01038	0.01038	0.05970	0.09489	0.165
new P		0.06	1.0	0.36	0.58	

2.

path	P(path)	P(V|a,N)	P(N|b,V)	P(N|a,N)	P(N|b,N)	sum P(.|.,N)
NVNVNN NVNNNN NNNVNN NNNNNN	0.00209 0.00727 0.00727 0.02529	0.00418 0.00727 0.00727 0.0	0.00418 0.00727 0.00727 0.0	0.0 0.00727 0.00727 0.05058	0.00209 0.01454 0.01454 0.07587	
sum over paths	0.04192	0.01872	0.01872	0.06512	0.10704	0.191
new P		0.10	1.0	0.34	0.56	

probability of sequence for iteration 3: 0.0472

a) Train the third iteration (HMM3).

b) Compute the entropy of the HMM in Iteration 2.

Good luck with your assignment :-)