Homework 4 Part A: HMM, Viterbi

Shamya Karumbaiah [Collaborated with - Rafael Lizarralde]

October 21, 2016

1 HMM

1.1

$P(o_2 = \Delta | s_1 = B) = \frac{\sum_{s \in A,B} P(o_2=\Delta, s_2=s, s_1=B)}{P(s_1=B)}$
Simplifying this using output independence and markov assumption of HMM, we get

$$= T_{B \to A} E_{A \to \Delta} + T_{B \to B} T_{B \to \Delta}$$

$P(o_2 = \Delta | o_1 = \Box) = \frac{\sum_{s \in A,B} \sum_{s' \in A,B} P(s_1=s, o_1=\Box, s_2=s', o_2=\Delta)}{\sum_{s'' \in A,B} P(s_1=s'') P(o_1=\Box)}$
Simplifying this using output independence and markov assumption of HMM, we get

$$= \frac{T_{START \to A}(T_{A \to A} E_{A \to \Delta} + T_{A \to B} E_{B \to \Delta}) + T_{START \to B}(T_{B \to A} E_{A \to \Delta} + T_{B \to B} E_{B \to \Delta})}{T_{START \to A} E_{A \to \Box} + T_{START \to B} E_{B \to \Box}}$$

Hence, No, $P(o_2 = \Delta | s_1 = B) \neq P(o_2 = \Delta | o_1 = \Box)$

1.2

$P(s_2 = B | s_1 = A) = T_{A \to B}$
$P(s_2 = B | s_1 = A, o_1 = \Delta) = P(s_2 = B | s_1 = A) = T_{A \to B}$ by markov assumption
Hence, Yes, $P(s_2 = B | s_1 = A, o_1 = \Delta) = P(s_2 = B | s_1 = A)$

1.3

$P(o_2 = \Delta | s_1 = A) = T_{A \to A} E_{A \to \Delta} + T_{A \to B} T_{B \to \Delta}$ similar to 1.1

$P(o_2 = \Box | s_1 = A, s_3 = A) = \frac{\sum_{s \in A,B} P(s_1=A, s_2=s, s_3=A, o_2=\Box)}{\sum_{s' \in A,B} P(s_1=A, s_2=s', s_3=A)}$
Simplifying this using output independence and markov assumption of HMM, we get

$$= \frac{T_{START \to A} T_{A \to A} E_{A \to \Box} T_{A \to A} + T_{START \to A} T_{A \to B} E_{B \to \Box} T_{B \to A}}{T_{START \to A} T_{A \to A} T_{A \to A} + T_{START \to A} T_{A \to B} T_{B \to A}}$$

Hence, No, $P(o_2 = \Delta | s_1 = A) \neq P(o_2 = \Box | s_1 = A, s_3 = A)$

1.4

$P(o_1 = \Box) = \sum_{s' \in A,B} P(s_1 = s', o_1 = \Box)$
$= P(s_1 = A, o_1 = \Box) + P(s_1 = B, o_1 = \Box)$
$= T_{START \to A} T_{A \to \Box} + T_{START \to B} T_{B \to \Box}$
$= 0.5 * 0.5 + 0.5 * 0.7$
$= 0.6$

1.5

Goal - Find $P(s_1 = A | o_2 = \Delta, s_3 = END)$

$$P(s_1 = A | o_2 = \Delta, s_3 = END) = \frac{P(s_1 = A, o_2 = \Delta, s_3 = END)}{P(o_2 = \Delta, s_3 = END)}$$

$$= \frac{\sum_{s' \in A,B} P(s_1 = A, s_2 = s', o_2 = \Delta, s_3 = END)}{\sum_{s'_1 \in A,B} \sum_{s'_2 \in A,B} P(s_1 = s'_1, s_2 = s'_2, o_2 = \Delta, s_3 = END)}$$

Simplifying this using output independence and markov assumption of HMM, we get

$$= \frac{T_{START \to A} T_{A \to A} E_{A \to \Delta} T_{A \to END} + T_{START \to A} T_{A \to B} E_{B \to \Delta} T_{B \to END}}{\begin{array}{l} T_{START \to A} T_{A \to A} E_{A \to \Delta} T_{A \to END} + T_{START \to A} T_{A \to B} E_{B \to \Delta} T_{B \to END} \\ \quad + T_{START \to B} T_{B \to A} E_{A \to \Delta} T_{A \to END} + T_{START \to B} T_{B \to B} E_{B \to \Delta} T_{B \to END} \end{array}}$$

$$= 0.3542$$

2 Viterbi

2.1

In HMM,
$log\ P(\vec{y}, \vec{w}) = \sum_t log\ P(y_t | y_{t-1}) + log\ P(w_t | y_t)$
$A(y_{t-1}, y_t)$ gives the transition log probabilities equivalent to $log\ P(y_t | y_{t-1})$
Along with $B_t(y_t)$ giving the preference for the tag y_t at t i.e, $log\ P(y_t)$, we could also define a table to give the preference for the tag y_t at t given the observation w_t, $B'_t(w_t, y_t) = log\ P(y_t | w_t)$. If we keep this constant over t, this could just be $B'(w_t, y_t)$.
If we don't worry about normalizing, we could use $B_t(w_t, y_t)$ as is for emission log probability in HMM. Else we could get these values using condition probability in HMM, $P(w_t | y_t) = \frac{P(y_t | w_t) P(w_t)}{P(y_t)}$