UMass CS 585, 9/28/17 in-class notes

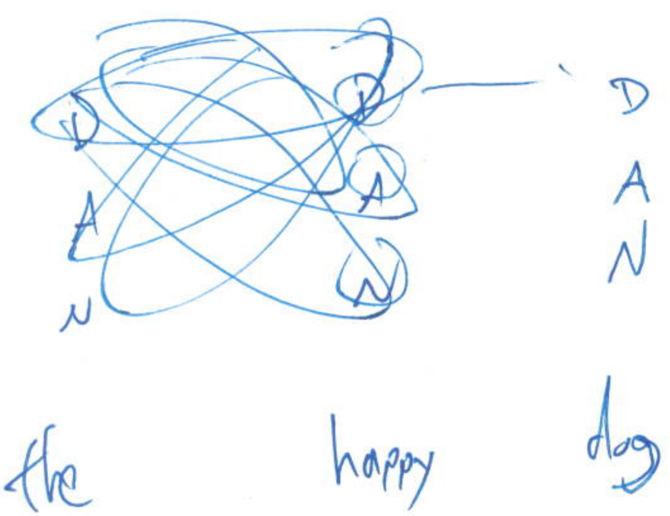
Forward Alg

Want: $P(w) = \sum_{\vec{y}} P(w, y)$

$\sum_{\vec{y}} \prod_t P(w_t | y_t)\, P(y_t | y_{t-1})$

N-th order HMM?

$P(y_t | y_{t-1}\, y_{t-2})$



1st order F. Algo.

Runtime?

K state vocab.

N tokens long

K^N: Naive Enum Alg.

$O(K^2 N)$: F. Algo

Want $P(\vec{w}) = \sum_{\vec{y}} P(\vec{w}, \vec{y})$

$$= \sum_{\vec{y}} \prod_t P(y_t \mid y_{t-1})\, P(w_t \mid y_t)$$

$y_1 = ?$ $\quad$ $y_2 = ?$ $\quad$ $y_3 = ?$

D $\quad$ D $\quad$ D

A $\quad$ A $\quad$ A

N $\quad$ N $\quad$ N

$w_1 =$ the $\quad$ $w_2 =$ happy $\quad$ $w_3 =$ dog

Forward Algo: left-to-right

Log-linear form for an HMM — for handout

$$A(y_1, y_2) = \log P_{trans}(y_2 \mid y_1)$$

$$A(y_2, y_3) = \log P_{trans}(y_3 \mid y_2)$$

$$B_1(y_1) = \log P_{emit}(w_1 \mid y_1)$$

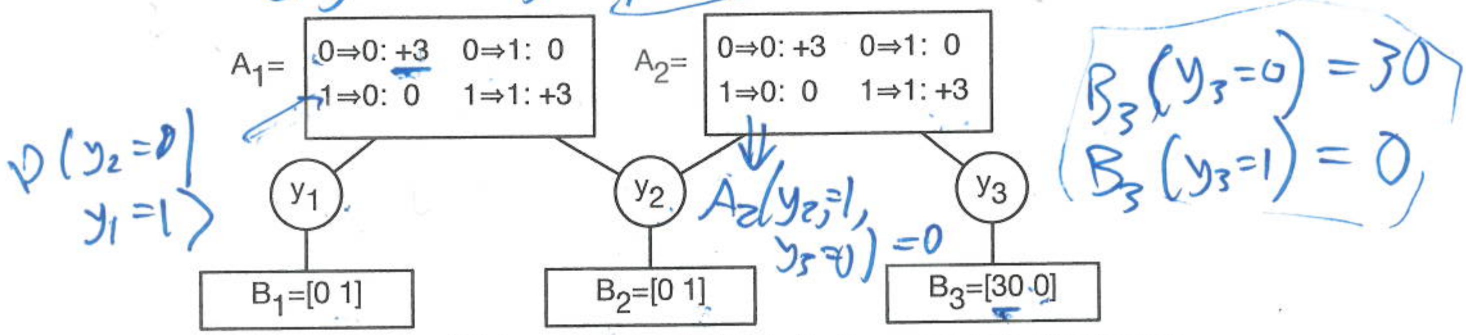
$$B_2(y_2) = \log P_{emit}(w_2 \mid y_2)$$

$$B_3(y_3) = \log P_{emit}(w_3 \mid y_3)$$

$y^* = (\_\_ , \_\_ , 0)$

CS585 Viterbi exercise

$\log P(0|0)$ $\quad A_2(y_2=0, y_3=0) = 3$



Sticky-favoring model over hidden state vocab {0,1}. Factor scores are "goodness points" are in log-scale additive form. (They're positive, though for an HMM they would all be negative.)

log P(y | w) = (constant) + G(y1, y2, y3)

G(y1, y2, y3) = A(y1,y2) + A(y2,y3) + B1(y1) + B2(y2) + B3(y3)

Additive Viterbi

For t=1..T,
  For k in {0,1},

$$V_t[k] := \max_j \left(V_{t-1}[j] + A(j,k) + B_t(k)\right)$$

$$B[k] := \arg\max_j (\ldots)$$

For t=1, set A_0(anything)=0 and V_0[anything]=0

Final backtrace step: take best-scoring from last V_T, follow the backpointers all the way back

Run Viterbi and fill out the trellis with arcs like in the textbook's HMM example.

Most probable sequence $y^* = (0 , 0 , 0)$ $\quad G(y^*) = 36$