11-711: Algorithms for NLP

Recitation #5

Greg Hanneman

October 31, 2007

Bigram Markov POS Tagging

Suppose we are doing part-of-speech tagging based on a bigram Markov model, as discussed in class. Show the tag sequence that the Viterbi algorithm would assign to the following sentence:

`the large can can hold the water .`

The following partial tables of probabilities may be useful:

$P(w_i \mid t_i)$: Look in the first column for w_i, then read across to t_i. Assume that the end-of-sentence token "." has tag **SENT** with probability 1.

	SENT	**DET**	**ADJ**	**N**	**V**	**AUX**
the	0	5.89×10^{-1}	9.71×10^{-5}	4.42×10^{-4}	8.76×10^{-5}	0
large	0	0	4.20×10^{-3}	0	0	0
can	0	0	0	2.78×10^{-5}	0	9.40×10^{-2}
hold	0	0	0	1.11×10^{-4}	5.06×10^{-3}	0
water	0	0	0	6.76×10^{-3}	2.80×10^{-5}	0

$P(t_i \mid t_{i-1})$: Look in the first column for t_i, then read across to t_{i-1}. Assume that every sentence begins with an implicit **SENT** tag for a null first word.

	SENT	**DET**	**ADJ**	**N**	**V**	**AUX**
SENT	1.31×10^{-4}	1.59×10^{-3}	2.51×10^{-2}	9.60×10^{-2}	3.36×10^{-2}	1.98×10^{-3}
DET	2.03×10^{-1}	1.67×10^{-3}	4.21×10^{-3}	7.56×10^{-3}	1.60×10^{-1}	4.12×10^{-3}
ADJ	4.25×10^{-2}	2.30×10^{-1}	7.39×10^{-2}	1.11×10^{-2}	7.62×10^{-2}	6.86×10^{-4}
N	2.59×10^{-1}	6.67×10^{-1}	7.03×10^{-1}	2.61×10^{-1}	1.35×10^{-1}	2.29×10^{-3}
V	2.18×10^{-2}	2.81×10^{-2}	9.50×10^{-3}	1.33×10^{-1}	1.29×10^{-1}	7.98×10^{-1}
AUX	6.00×10^{-4}	2.07×10^{-3}	4.36×10^{-4}	1.76×10^{-2}	7.29×10^{-4}	1.52×10^{-4}