**CS6320, Spring 2018**

**Dr. Mithun Balakrishna**

**Homework 3**

**Due Friday, March 2nd, 2018 11:59pm**

1. **Submission Instructions:** 
   * + Submit your solutions via eLearning.
     + Please submit a single zip file with the following files:

o For programming questions:

* + - * + Source code file(s) in C/C++, Java, or Python. For using any other programming language, please get prior approval from the TA.
        + A ReadMe file with instructions on how to compile/run the code.

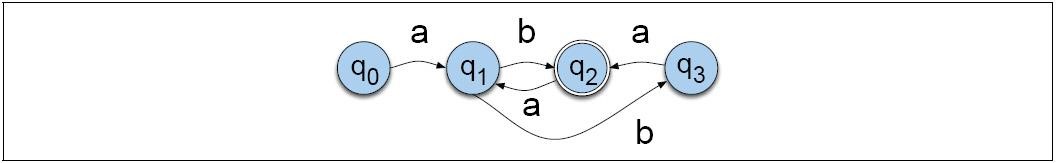
o For all other questions, a PDF/Doc/PS/Image file with the solutions.

 Late Submission Penalty:

o up to 2 hours late — 10% deduction o 2 - 4 hours late — 20% deduction o 4 - 12 hours late — 35% deduction o 12 - 24 hours late — 50% deduction o 24 - 48 hours late — 75% deduction o more than 48 hours late — 100% deduction (zero credit)

1. **Problems:** 
   1. **NFSA to Regular Expression (20 points)**

Write a regular expression for the language accepted by the NFSA:



**ANS: ^(ab)+$|^(ab)+a$**

* 1. **POS Tagging Errors (10 points)**

Find one tagging error in each of the following sentences that are tagged with the Penn Treebank POS tagset (Figure 1):

* + 1. I\_PRP need\_VBP a\_DT flight\_NN from\_IN Atlanta\_NN

**ANS:** Atlanta-NNP

* + 1. Does\_VBZ this\_DT flight\_NN serve\_VB dinner\_NNS

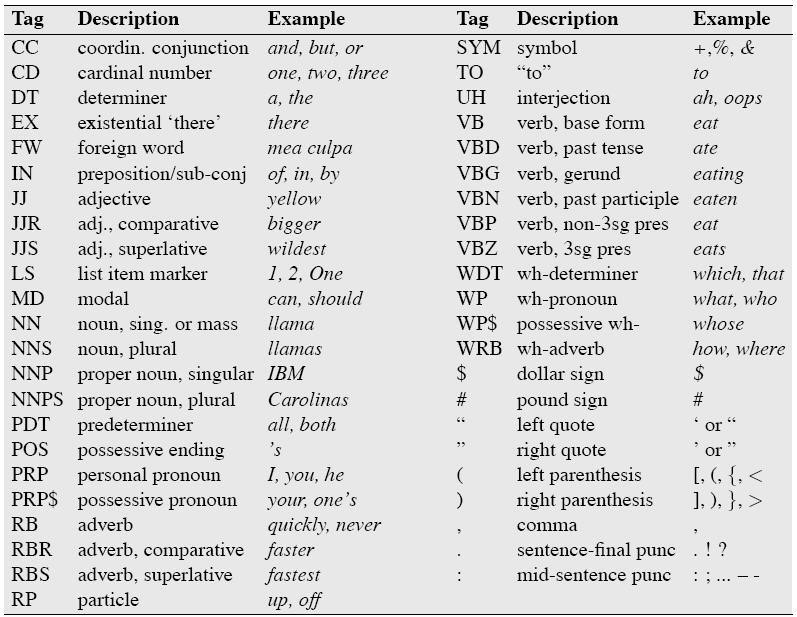
**ANS:** dinner-NN

* + 1. I\_PRP have\_VB a\_DT friend\_NN living\_VBG in\_IN Denver\_NNP

**ANS:** have-VBP

* + 1. Can\_VBP you\_PRP list\_VB the\_DT nonstop\_JJ afternoon\_NN flights\_NNS

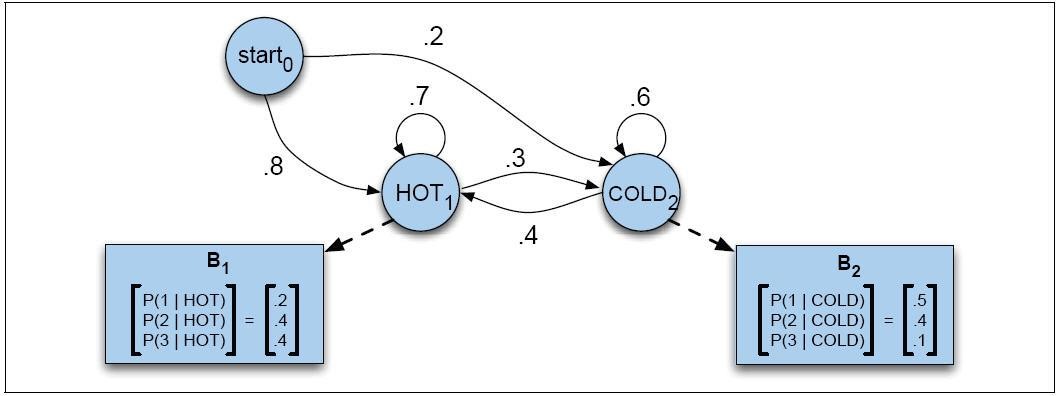
**ANS:** Can-MD



**Figure 1. Penn Treebank POS tagset**

**3. HMM Decoding: Viterbi Algorithm (70 points):**

Programmatically implement the Viterbi algorithm and run it with the HMM in Figure 2 to compute the most likely weather sequence and probability for a given observation sequence. Example observation sequences: *331, 122313, 331123312*, etc.



**Figure 2. HMM**