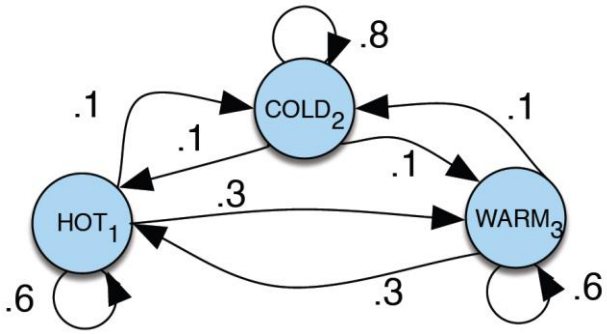
1. Explain how BIO tagging helps with the segmentation problem with NER

2. Consider the following Markov chain:



Take the initial probabilities

$$\pi = [.1\ (\text{cold}),\ .7(hot),\ .2(\text{warm})]$$

Compute the probabilities for seuqneces

A. hot hot hot hot
B. cold hot cold hot

3. Explain how POS tagging can be formulated as an HMM problem.

4. Explain how NER tagging can be formulated as an HMM problem.

5. Consider the sentence "Janet will back" for POS tagging. Consider only the following POS tags: NNP (proper noun, sing), MD (modal verb), VB (verb base):

   Generate POS tags for the above sentence solving the corresponding HMM formulation. Consider the following transition probabilities A: $P(t_i|t_{i-1})$ – precomputed from a corpus:

	NNP	MD	VB
<s>	0.2767	0.0006	0.0031
NNP	0.3777	0.011	0.0009
MD	0.0008	0.0002	0.7968
VB	0.0322	0.0005	0.005

Consider the following Observation likelihoods/ emission probabilities: B: $b_i(o_t)$ or $p(w_i|t_i)$

	Janet	will	back
NNP	0.000032	0	0
MD	0	0.3084	0
VB	0	3E-05	0.0007

6. Given the corpus of training data (the notation word/TAG means word tagged with a specific part-of-speech tag).

- eat/**VB** breakfast/**NN** at/**IN** morning/**NN** time/**NN**
- take/**VB** time/**NN** with/**IN** arrow/**NN** projects/**NN**
- horse/**NN** riders/**NN** like/**VB** the/**DT** airport/**NN**
- paper/**NN** flies/**VB** on/**IN** hydrogen/**NN** gas/**NN**
- bees/**NN** sting/**VB** like/**IN** some/**DT** flies/**NN**
- beans/**NN** soil/**VB** an/**DT** iron/**NN** grill/**NN**
- flies/**NN** smell/**VB** an/**DT** arrow/**NN** drink/**NN**
- people/**NN** like/**VB** an/**DT** army/**NN** arrow/**NN**
- dinner/**NN** time/**NN** flies/**VB** all/**DT** day/**NN**
- horse/**NN** flies/**NN** time/**VB** morning/**NN** rays/**NN**

Consider the test sentence “Time flies like an arrow”. Based on the above corpus draw transition probability and emission probability for the words in the above sentence and relevant tags. Estimate the POS tags based on these probabilities through HMM.