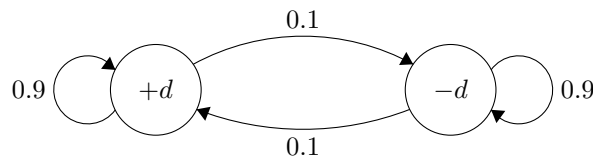
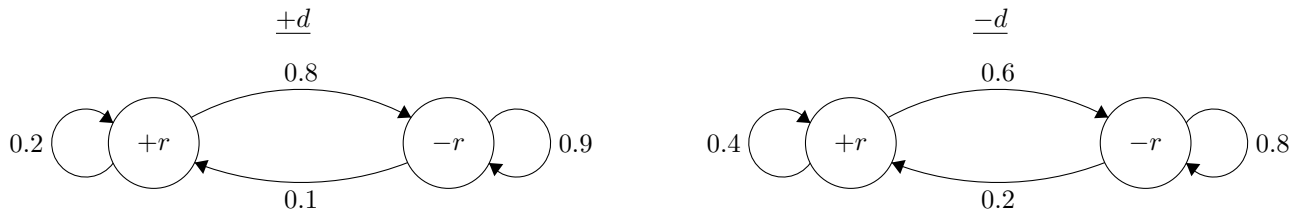


### Q1. I Heard You Like Markov Chains

In California, whether it rains or not from each day to the next forms a Markov chain (note: this is a terrible model for real weather). However, sometimes California is in a drought and sometimes it is not. Whether California is in a drought from each day to the next itself forms a Markov chain, and the state of this Markov chain affects the transition probabilities in the rain-or-shine Markov chain. This is the state diagram for droughts:



These are the state diagrams for rain given that California is and is not in a drought, respectively:



(a) Draw a dynamic Bayes net which encodes this behavior. Use variables  $D_{t-1}$ ,  $D_t$ ,  $D_{t+1}$ ,  $R_{t-1}$ ,  $R_t$ , and  $R_{t+1}$ . Assume that on a given day, it is determined whether or not there is a drought before it is determined whether or not it rains that day.

(b) Draw the CPT for  $D_t$  in the above DBN. Fill in the actual numerical probabilities.

(c) Draw the CPT for  $R_t$  in the above DBN. Fill in the actual numerical probabilities.

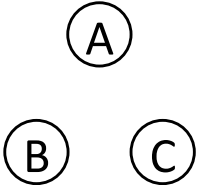
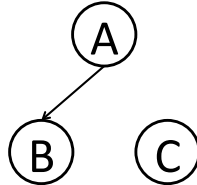
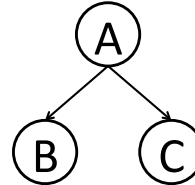
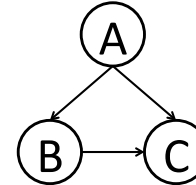
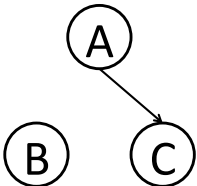
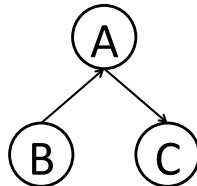
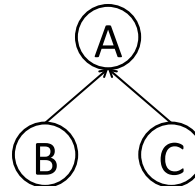
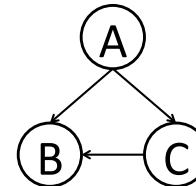
Suppose we are observing the weather on a day-to-day basis, but we cannot directly observe whether California is in a drought or not. We want to predict whether or not it will rain on day  $t + 1$  given observations of whether or not it rained on days 1 through  $t$ .

- (d) First, we need to determine whether California will be in a drought on day  $t + 1$ . Derive a formula for  $P(D_{t+1}|r_{1:t})$  in terms of the given probabilities (the transition probabilities on the above state diagrams) and  $P(D_t|r_{1:t})$  (that is, you can assume we've already computed the probability there is a drought today given the weather over time).

- (e) Now derive a formula for  $P(R_{t+1}|r_{1:t})$  in terms of  $P(D_{t+1}|r_{1:t})$  and the given probabilities.

## Q2. Learning a Bayes' Net Structure

- (a) You want to learn a Bayes' net over the random variables  $A, B, C$ . You decide you want to learn not only the Bayes' net parameters, but also the structure from the data. You are willing to consider the 8 structures shown below. First you use your training data to perform maximum likelihood estimation of the parameters of each of the Bayes' nets. Then for each of the learned Bayes' nets, you evaluate the likelihood of the training data ( $l^{\text{train}}$ ), and the likelihood of your cross-validation data ( $l^{\text{cross}}$ ). Both likelihoods are shown below each structure.

				
$l^{\text{train}}$	0.0001	0.0005	0.0015	0.0100
$l^{\text{cross}}$	0.0001	0.0004	0.0011	0.0009
	(a)	(b)	(c)	(d)
				
$l^{\text{train}}$	0.0008	0.0015	0.0020	0.0100
$l^{\text{cross}}$	0.0006	0.0011	0.0010	0.0009
	(e)	(f)	(g)	(h)

- (i) Which Bayes' net structure will (on expectation) perform best on test-data? (If there is a tie, list all Bayes' nets that are tied for the top spot.) Justify your answer.
- (ii) Two pairs of the learned Bayes' nets have identical likelihoods. Explain why this is the case.
- (iii) For every two structures  $S_1$  and  $S_2$ , where  $S_2$  can be obtained from  $S_1$  by adding one or more edges,  $l^{\text{train}}$  is higher for  $S_2$  than for  $S_1$ . Explain why this is the case.