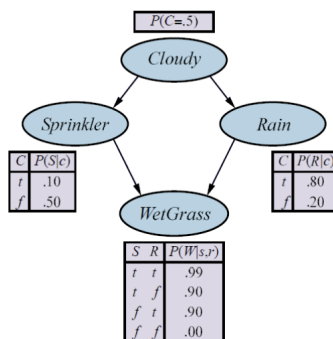1. **(20 pts)** (Exercise 14.21) Consider the query $\mathbf{P}(Rain|Sprinkler = true, WetGrass = true)$ in the figure below and how Gibbs sampling via uniform choice of nonevidence variables can answer it.

(a) **(2 pts)** How many states does the Markov chain for this query have?

4

(b) **(10 pts)** Calculate the *transition matrix* Q containing the kernel $k(\mathbf{y} \to \mathbf{y}')$ for every directed edge from, say, a state $\mathbf{y}$ to another state $\mathbf{y}'$ or itself (in the case $\mathbf{y} = \mathbf{y}'$), in the Markov chain. Label the rows and columns of the matrix by the states such that an entry in the matrix corresponds to the probability of a transition from the state given by the row label to the state given by the column label.

	(c,r)	$(c,\neg r)$	$(\neg c,r)$	$(\neg c,\neg r)$
(c,r)	0.6295	0.0925	0.278	0
$(c,\neg r)$	0.4075	0.1165	0	0.476
$(\neg c,r)$	0.222	0	0.386	0.392
$(\neg c,\neg r)$	0	0.024	0.108	0.868

(c) **(2 pts)** What does Q^2, the square of the transition matrix, represent?

It represents the probability of starting with the state in the row and ending in the column state after 2 steps.

(d) **(3 pts)** What about Q^n as $n \to \infty$?

As $n \to \infty$, Q^n reaches a stationary distribution which represents the probability of being in a state.

(e) **(3 pts)** Explain how to do probabilistic inference in a Bayesian network, assuming that Q^n is available. Is this a practical way to do inference?

You can multiply Q^n however many times you need to reach different numbers of steps. And you can do it efficiently by multiplying the newly calculated matrices together to get higher power matrices in less calculations. However, matrix multiplication requires a lot of computation, so I don't think it is very practical.

2. **(10 pts)** (Exercise 14.22) This exercise explores the stationary distribution for Gibbs sampling methods.

(a) **(5 pts)** The convex composition $[\alpha, k_1; 1 - \alpha, k_2]$ of two transition kernels k_1 and k_2 is a transition probability distribution that first chooses one of k_1 and k_2 with probabilities α and $1 - \alpha$, respectively, and then applies whichever is chosen. Prove that if k_1 and k_2 are in detailed balance with a stationary distribution π, then their convex composition is also in detailed balance with π.

$\pi(\mathbf{x})(\alpha k_1(\mathbf{x} \to \mathbf{x}') + (1 - \alpha)k_2(\mathbf{x} \to \mathbf{x}')) = \pi(\mathbf{x}')(\alpha k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)k_2(\mathbf{x}' \to \mathbf{x}))$
$\alpha\pi(\mathbf{x})k_1(\mathbf{x} \to \mathbf{x}') + (1 - \alpha)\pi(\mathbf{x})k_2(\mathbf{x} \to \mathbf{x}') = \pi(\mathbf{x}')(\alpha k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)k_2(\mathbf{x}' \to \mathbf{x}))$
We know that k_1 and k_2 are in detailed balance with π, so apply that rule...
$\alpha\pi(\mathbf{x}')k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)\pi(\mathbf{x}')k_2(\mathbf{x}' \to \mathbf{x}) = \pi(\mathbf{x}')(\alpha k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)k_2(\mathbf{x}' \to \mathbf{x}))$
$\pi(\mathbf{x}')(\alpha k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)k_2(\mathbf{x}' \to \mathbf{x})) = \pi(\mathbf{x}')(\alpha k_1(\mathbf{x}' \to \mathbf{x}) + (1 - \alpha)k_2(\mathbf{x}' \to \mathbf{x}))$

(b) **(5 pts)** Prove that if each of k_1 and k_2 has π as its stationary distribution, then the sequential composition $k = k_2 \circ k_1$ defined by

$$(k_2 \circ k_1)(\mathbf{x} \to \mathbf{x}') = \sum_{\mathbf{x}''} k_1(\mathbf{x} \to \mathbf{x}'')k_2(\mathbf{x}'' \to \mathbf{x}')$$

also has π as its stationary distribution.

$$\sum_{\mathbf{x}} \pi(\mathbf{x})(k_2 \circ k_1)(\mathbf{x} \to \mathbf{x}') = \sum_{\mathbf{x}} \pi(\mathbf{x}) \sum_{\mathbf{x}''} k_1(\mathbf{x} \to \mathbf{x}'') k_2(\mathbf{x}'' \to \mathbf{x}')$$

$$\pi(\mathbf{x}') = \sum_{\mathbf{x}''} k_2(\mathbf{x}'' \to \mathbf{x}') \sum_{\mathbf{x}} \pi(\mathbf{x}) k_1(\mathbf{x} \to \mathbf{x}'')$$

$$\pi(\mathbf{x}') = \sum_{\mathbf{x}''} k_2(\mathbf{x}'' \to \mathbf{x}') \pi(\mathbf{x}'')$$

$$\pi(\mathbf{x}') = \sum_{\mathbf{x}''} \pi(\mathbf{x}'') k_2(\mathbf{x}'' \to \mathbf{x}')$$

$$\pi(\mathbf{x}') = \pi(\mathbf{x}')$$