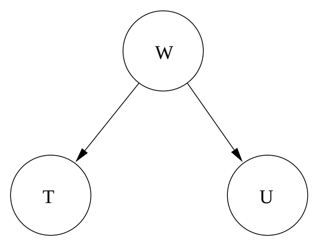
Problem: Bayesian Network (*taken from Spring'97 Final*)

Consider three random BOOLEAN variables W, T, and U, such that:

- W=TRUE indicates that the weather is bad
- T=TRUE indicates that there is traffic on the turnpike between Philadelphia and Pittsburgh
- U=TRUE indicates that USAir cancels all flights between Philadelphia and Pittsburgh

1. Give a Bayes network representing the following facts:
   - For a given state of weather, cancelled flights don't affect turnpike traffic.
   - Similarly, for a given state of weather, turnpike traffic does not affect USAir's flight cancellation decisions.



2. Use the following additional information gathered from observing a large set of data to compute ALL the conditional probability values for your Bayes network:
   - There is a 20% change of bad weather
   - If the weather is bad, there is an 80% chance that USAir will cancel all its flights from Philadelphia to Pittsburgh. In fact, even if the weather is good, there is a 40% chance those flights are cancelled.
   - Bad weather causes turnpike traffic 30% of the time, but in good weather, there is only a 10% chance of delays due to traffic.

$$
\begin{aligned}
P(W) &= 0.2 \\
P(T|W) &= 0.3 \\
P(T|\neg W) &= 0.1 \\
P(U|W) &= 0.8 \\
P(U|\neg W) &= 0.4
\end{aligned}
$$

$$
\begin{aligned}
P(\neg W) &= 0.8 \\
P(\neg T|W) &= 0.7 \\
P(\neg T|\neg W) &= 0.9 \\
P(\neg U|W) &= 0.2 \\
P(\neg U|\neg W) &= 0.6
\end{aligned}
$$

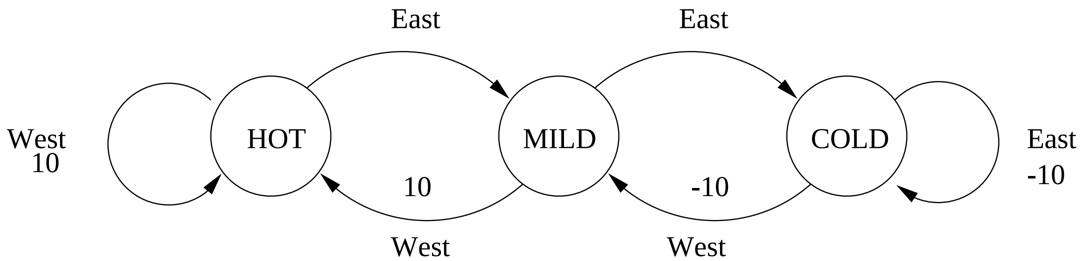
3. Compute $P(U|T)$. Show your work completely.

$$P(U|T) = \frac{P(U,T)}{P(T)}$$

$$\begin{aligned}
P(U,T) &= \sum_w P(U,T,W) \\
&= P(U,T,W) + P(U,T,\neg W) \\
&= P(W)P(T|W)P(U|W) + P(\neg W)P(T|\neg W)P(U|\neg W) \\
&= (0.2)(0.3)(0.8) + (0.8)(0.1)(0.4) \\
&= 0.08 \\
P(T) &= \sum_w P(T,W) \\
&= \sum_w P(W)P(T|W) \\
&= P(W)P(T|W) + P(\neg W)P(T|\neg W) \\
&= (0.2)(0.3) + (0.8)(0.1) \\
&= 0.14 \\
P(U|T) &= \frac{0.08}{0.14} \\
&= 0.57
\end{aligned}$$

Problem: Deterministic Reinforcement Learning

Consider the simple 3-state deterministic weather world sketched in the figure below. There are two actions for each state, namely moving West and East. The non-zero rewards are marked on the transition edges, hence edges without a number correspond to a zero reward.



1. A robot starts in the state MILD. It is actively learning its $\widehat{Q}$-table and moves in the world for 4 steps choosing actions: **West, East, East, West**. The initial values of the $\widehat{Q}$-table are zero and the discount factor is $\gamma = 0.5$. Fill in the $\widehat{Q}$-values for the sequence of states visited.

	Initial State: MILD		Action: West New State: HOT		Action: East New State: MILD		Action: East New State: COLD		Action: West New State: MILD	
	East	West	East	West	East	West	East	West	East	West
HOT	0	0	0	0	**5**	0	5	0	5	0
MILD	0	0	0	**10**	0	10	**0**	10	0	10
COLD	0	0	0	0	0	0	0	0	0	**-5**

2. How many possible **policies** are there in this 3-state deterministic world?

8

(There are 3 states, each with two possible actions. Therefore we have 2^3 possible policies.)

3. Why is the policy $\pi(s) =$ West, for all states, *better than* the policy $\pi(s) =$ East, for all states?

The policy of always moving West (π_w) is better than that of always moving East (π_e) because it gives you a higher cumulative value for all of the three states.

For the policy of always moving West:

$$V^{\pi_w}(HOT) = 10 + (0.5)V^{\pi_w}(HOT)$$

$$
\begin{aligned}
&= 10\sum_{i=0}^{\infty} 0.5^i \\
&= 10(\frac{1}{1-0.5}) \\
&= 20
\end{aligned}
$$

$$
\begin{aligned}
V^{\pi_w}(MILD) &= 10 + (0.5)V^{\pi_w}(HOT) \\
&= 20
\end{aligned}
$$

$$
\begin{aligned}
V^{\pi_w}(COLD) &= -10 + (0.5V^{\pi_w}(MILD) \\
&= -10 + (0.5)(20) \\
&= 0
\end{aligned}
$$

For the policy of always moving East:

$$
\begin{aligned}
V^{\pi_e}(COLD) &= -10 + (0.5)V^{\pi_e}(COLD) \\
&= -10\sum_{i=0}^{\infty} 0.5^i \\
&= -10(\frac{1}{1-0.5}) \\
&= -20
\end{aligned}
$$

$$
\begin{aligned}
V^{\pi_e}(MILD) &= 0 + (0.5)V^{\pi_e}(COLD) \\
&= 0 + (0.5)(-20) \\
&= -10
\end{aligned}
$$

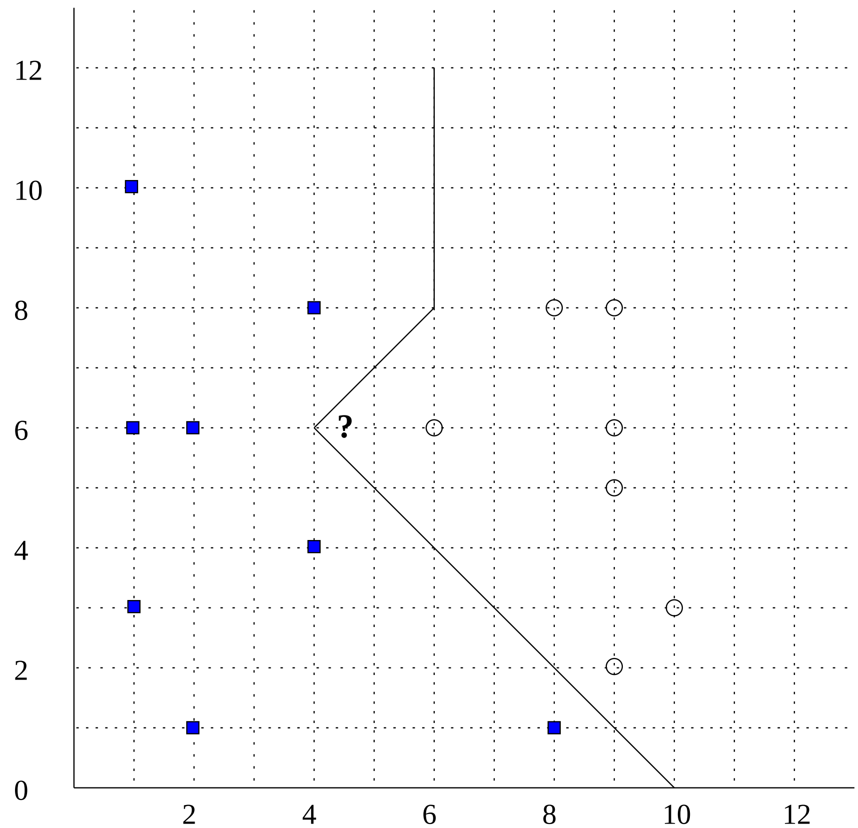
$$
\begin{aligned}
V^{\pi_e}(HOT) &= 0 + (0.5)V^{\pi_e}(MILD) \\
&= 0 + (0.5)(-10) \\
&= -5
\end{aligned}
$$

Problem: Nearest Neighbor Learning

A professor of an undergraduate AI course decided to ask some former students a few key questions so as to predict the performance of the current students on the upcoming AI final. The professor asked just two questions:

- How many hours did you sleep the night before the exam?
- How many hours did you study for the exam?

The training sample is shown below, with 15 individuals plotted according to their answers to the two questions, with a dark square (didn't get an A) and a light circle (got an A) indicating how they did on the final exam. There is also a question mark "?" on the plot, indicating a current student who the professor has queried just prior to the start of the exam.



1. What would the 1-nearest neighbor predict the current student's performance to be, assuming Euclidian distance is used as the similarity metric (your answer should be either "got an A" or "didn't get an A").

"Got an A"

2. On the plot above, carefully draw the precise boundary lines that the 1-nearest neighbor algorithm would indicate as separating the two spaces.