

Programming Project 3

April 15, 2019

Due May 1, 2019

On Patriots' Day (a holiday observed by several states, including Massachusetts, on the third Monday in April), the Tortoise and the Hare have their annual footrace in a small town in eastern Massachusetts. On the morning of the race, either the short or the long course is chosen at random. Another major factor is the weather, which can be cold and wet, very hot, or very nice. The short course favors the Hare. The Hare does not like cold and wet weather, frequently leaving the course and taking shelter. The Tortoise does not like the heat, although on the long course there are some muddy puddles that it can use to cool off.

The Bayesian network representing this situation consists of the following variables:

- *Course*, which can have values *short* or *long*.
- *Weather*, which can have values *coldWet*, *hot*, or *nice*.
- *HarePerf*, which can have the values *slow*, *medium*, *fast*, representing the performance of the Hare.
- *TortoisePerf*, which does the same for the Tortoise.
- *HareWins*, a Boolean variable which is true when the Hare wins.

Generally, a "faster" performance will result in a win by that creature, but not always. The Hare, for instance, is apt to get overconfident and stop for selfies and autographs.

Implement this network with the probabilities indicated below. You will need to implement two algorithms for sampling:

- PRIOR-SAMPLE, which is run without any evidence to generate a sample for the general case. Run many times, it should give a probability distribution for the Hare's winning.
- Your choice of REJECTION SAMPLING or LIKELIHOOD WEIGHTING, which are run in the presence of evidence.

Using these two algorithms, generate answers to the following three queries:

1. In general, how likely is the Hare to win?
2. Given that it is *coldWet*, how likely is the Hare to win?
3. Given that the Tortoise won on the *short* course, what is the probability distribution for the *Weather*?

Rather than interpreting queries at the keyboard, these may be hard-coded with a menu selection to run one of them.

Here are the prior and posterior probabilities for the network:

<i>Course</i>	$P(\text{Course})$
<i>short</i>	0.5
<i>long</i>	0.5

<i>Weather</i>	$P(\text{Weather})$
<i>coldWet</i>	0.3
<i>hot</i>	0.2
<i>nice</i>	0.5

<i>Course</i>	<i>Weather</i>	$P(\text{HarePerf})$
<i>short</i>	<i>coldWet</i>	<0.5, 0.3, 0.2>
<i>short</i>	<i>hot</i>	<0.1, 0.2, 0.7>
<i>short</i>	<i>nice</i>	<0.0, 0.2, 0.8>
<i>long</i>	<i>coldWet</i>	<0.7, 0.2, 0.1>
<i>long</i>	<i>hot</i>	<0.2, 0.4, 0.4>
<i>long</i>	<i>nice</i>	<0.1, 0.3, 0.6>

<i>Course</i>	<i>Weather</i>	$P(\text{TortoisePerf})$
<i>short</i>	<i>coldWet</i>	<0.2, 0.3, 0.5>
<i>short</i>	<i>hot</i>	<0.4, 0.5, 0.1>
<i>short</i>	<i>nice</i>	<0.3, 0.5, 0.2>
<i>long</i>	<i>coldWet</i>	<0.2, 0.4, 0.4>
<i>long</i>	<i>hot</i>	<0.2, 0.5, 0.3>
<i>long</i>	<i>nice</i>	<0.4, 0.4, 0.2>

<i>HarePerf</i>	<i>TortoisePerf</i>	$P(\text{HareWins})$
<i>slow</i>	<i>slow</i>	0.5
<i>slow</i>	<i>medium</i>	0.1
<i>slow</i>	<i>fast</i>	0.0
<i>medium</i>	<i>slow</i>	0.8
<i>medium</i>	<i>medium</i>	0.5
<i>medium</i>	<i>fast</i>	0.2
<i>fast</i>	<i>slow</i>	0.9
<i>fast</i>	<i>medium</i>	0.7
<i>fast</i>	<i>fast</i>	0.5

Grading Rubric:

Design/Clarity/Style:	25%
Overall function with just PRIOR-SAMPLE:	60%
Posterior Probability Sampling:	15%

To hand in:

- Source Code
- A run that addresses each of the three queries above
- A README file with any instructions for compilation or input