Reach out and torch someone

Introduction to Computational Reasoning

Random Behavior in Models

Susan Ragan – Maryland Virtual High School of Science and Mathematics
Scott Sinex – Prince George's Community College

This module is an introduction to computational reasoning. We are going to examine a pre-built computer model to run some experiments and collect data. We will analyze the data by graphing it, look at variation in the data, and curve fit in a pre-built Excel spreadsheet. No prior knowledge of the computer model or Excel is needed.

Reach out and torch someone!

Many computer models rely on the concepts of probability and random numbers to represent events in the real world. In this module, we'll use coin flipping and the spread of fire as examples of probabilistic behavior.

From Physical to Virtual Manipulatives

Consider flipping a coin. What are the chances or probability of getting a heads on a flip?

Now flip a coin 10 times and complete the table below. Compare your results to another group.

For 10 flips of a coin	Number of heads	Number of tails
In theory:		
In experiment:		

How do they compare?

Computer models use formulas to generate what are known as pseudo-random number generators. Usually, a random number generator will generate a number between 0 and 1. To simulate the flipping of a coin, we would determine a rule based on the assumption that half of our random numbers will be less than 0.5 and half will be more than 0.5. However, if we generate only 10 random numbers, we won't necessarily get 5 numbers that are less than 0.5 and 5 numbers that are more than 0.5.

To see an example of this, bring up the interactive Excel spreadsheet for flipping pennies (flipping_pennies_CAST.xls).

Go to the "10 flips" tab on the spreadsheet and press the F9 key to generate a new simulation of flipping 10 pennies. Note that the results are shown in three different ways. The words 'heads' and 'tails' are used, the numbers 1 and 0 are used where 1 means 'heads', and a bar graph shows the distribution of heads and tails for each simulation. The number of heads is recorded at the top and the percent error is calculated. The %error is based on the fact that the expected value for heads is 5. The formula for the %error calculation can be seen by putting the mouse over the red triangle above the '=' sign.

By pressing the F9 key on the computer keyboard, flip your ten pennies for ten trials and record the number of heads you get on each trial.

Number of heads out of ten flips of a penny

Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Trial 6	Trial 7	Trial 8	Trial 9	Trial 10

How do your trials compare to each other?

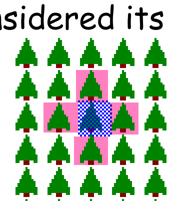
Ten flips of 10 pennies each gives you 100 flips of a penny. How many heads did you get? How does that total number of heads compare to others in the class?

Go to the "100 flips" tab and try a few trials flipping a hundred pennies at a time. What do you notice about the error?

If you have time, you can try the "1000 flips" tab and "10,000 flips" tab and watch the error.

<u>Computer Model and Data Collection</u>

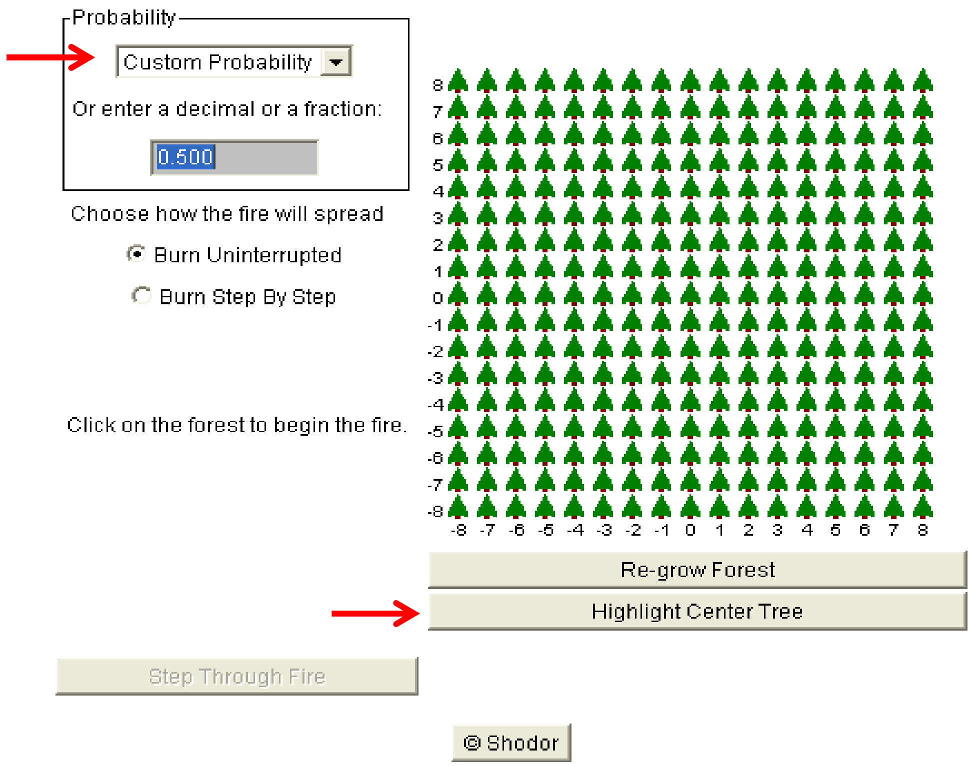
Now let's examine a computer model of burning a forest that uses probability to determine whether a fire started by a lightning strike to a single tree will spread to neighboring trees. In this simulation, we assume that all the trees are the same and there is no wind. At a certain probability, the fire can spread from one tree to its nearest neighbors (the tree highlighted in blue is on fire and the four trees around it highlighted in pink are considered its nearest neighbors).



Open Fire!! (a Java applet from Shodor's Interactivate)

http://www.shodor.org/interactivate/activities/fire1/index.html

Here is the opening screen shot of Fire!! Go to the dropdown menu and select Custom Probability. That will allow us to set the probability. To set the fire, we will also click on the center tree (you may want to hit the "Highlight Center Tree" button to help).



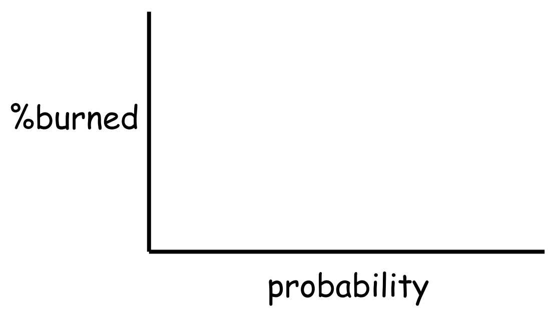
First we will look at some simple cases. If there is zero probability that a tree will be ignited by the burning tree next to it, then how many trees will burn after we start the first tree burning?

Let's try it. Set the probability to 0 and ignite the center tree. Repeat this several times. What did you observe?

What should happen if the probability is set to 100%? Try it and record your observations.

Now we want to investigate a range of probabilities to see what percentage of the forest will burn. For each probability we will run ten trials and then average the %burned.

How do you think the %burned will vary as the probability changes? Sketch a prediction on the graph to the right. You already have the values to fill in for 0% and 100% probabilities.



Will 50% of the forest burn if the probability of the fire spreading to a neighboring tree is set to 0.5?

Will we all get the same result when we each set the probability to 0.5 and run the model?

Let's collect the %burned from each person and use the average as the next point on our graph.

<u>Collecting and Analyzing Data</u>

You will be assigned a probability between 0.05 and 0.95. Be sure to change your probability to the value you were assigned, then being sure to always ignite the center tree, burn the forest 10 times and average your results.

Probability = _______ Average of 10 trials = _______

Percent of trees burned

Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Trial 6	Trial 7	Trial 8	Trial 9	Trial 10

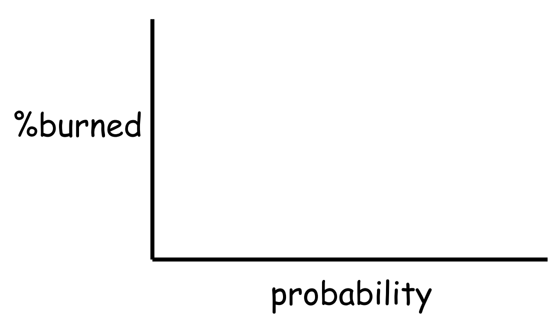
For your ten trials, did the burn pattern look the same for each trial? Describe.

Now share the data with the class and create a table of the class data.

Percent of trees burned given different probabilities

P = 0.05	P = 0.10	P = 0.15	P = 0.20	P = 0.25
P = 0.30	P = 0.35	P = 0.40	P = 0.45	P = 0.50
P = 0.55	P = 0.60	P = 0.65	P = 0.70	P = 0.75
P = 0.80	P = 0.85	P = 0.90	P = 0.95	P = 1.00

Bring up the interactive Excel spreadsheet for Fire Analysis (fire_analysis_CAST.xls) and go to the "data analysis" tab. Here you will enter the mean percent burned for the various probabilities. The graph will automatically plot as you enter the data. Sketch the graph.



Did the graph of the actual data match your prediction? Explain.

Some Other Considerations

Is the position of the lightning strike a factor? We hit the center tree in all the runs. Investigate another position for a P = 0.50. Compare to other groups.

Is the model's grid of the forest realistic? Explain.

Are there other factors that influence the spread of a forest fire? Explain.

Curve-fitting and Statistical Analysis – Developing a Mathematical Model

The %burned vs. probability graph is known by several names, including s-shaped curve, bounded growth, logistic curve, and sigmoid curve.

Now let's see if we can develop a mathematical model for the data collected. The s-shaped curve or sigmoidal curve is mathematically called a logistic function.

$$y = \frac{c}{1 + ae^{-bx}}$$

Go to the "curve fit" tab (fire_analysis_CAST.xls) and change the values of a, b, and c to see if you can manually fit the green curve to the data points (your mean %burned values from the data analysis tab). Error is defined as the measured value minus the predicted value. The better the fit of the green curve to the data, the lower the SSE (sum of the squared errors) will get. The best fit will minimize the SSE. You would like a random pattern in the error as well.

This is not an easy manual fit to do. So here are some hints. The value of c is the maximum where the logistic curve levels out. The value of b controls the rate of growth of the curve and will be in the range of 10-20 for this model. Start the value of a at 5000 and increase it. Play and see how good a fit you can get. Compare your values to other groups.

Record your equation using the names of the variables studied (not x and y) and the numerical values of a, b, and c. Record your SSE.

Equation by Manual Fit	SSE

Getting the best fit by minimizing the SSE can be accomplished by running the Solver in Excel. Since the Solver has been pre-set, all you need to do is run it (See "Solver info" tab for instructions). Your goal is to obtain a curve that is close to the data points.

Record your equation using the names of the variables studied (not x and y) and the numerical values of a, b, and c. Record your SSE.

Equation by Solver Fit	SSE

Does the error as a function of probability graph look random? There is no pattern to the distribution of the data points.

How close did your manual fit come to the best fit?

Acknowledgments

The authors graciously wish to thank Bob Panoff of Shodor for the idea behind this activity.