# Invasive Species and Probability: Percolation of the Emerald Ash Borer

Nate Shank, Laurie Heyer, Gwen Spencer

December 23, 2012

- **Module Summary:** This module explores the probabilistic spread of the Emerald Ash Borer (an invasive species) through a series of stylized landscapes.
- **Informal Description:** This module explains the ideas of percolation and their application on several different levels. Students learn through simulation or proof about percolation and percolation thresholds and how they can be applied to many areas including invasive species control.
- **Target Audience:** The module materials include a number of components (handout, simulation/lab, lecture notes on mathematical proofs) in the interest of customization (both across levels of math courses and across disciplines). The introduction, description of percolation and subsequent proofs would be an appropriate lesson for an advanced undergrad probability class. The introduction, math background, description of percolation, and simulation component would be an appropriate lesson for an earlier course, or for a general course on quantitative methods in ecology.
- **Prerequisites:** The module can be taught at two different levels; beginner and advanced.
  - Beginner: High school through early college non-mathematics majors. Basic understanding of probability is helpful, but not necessary.
  - Advanced: Intro to proofs or beyond. Understanding of limits, probability, counting, and proofs is sufficient. Little to no background in graph theory is needed.
- **Mathematical Fields:** Probability, Simulation, Discrete Math, Proof writing, Graph Theory.
- **Application Areas:** Invasive species threaten local ecology as well as renewable economic resources systems (in this case, the timber industry). Local species may be overconsumed or out-competed for food sources: extinctions result in gaps that destabilize greater ecological processes.

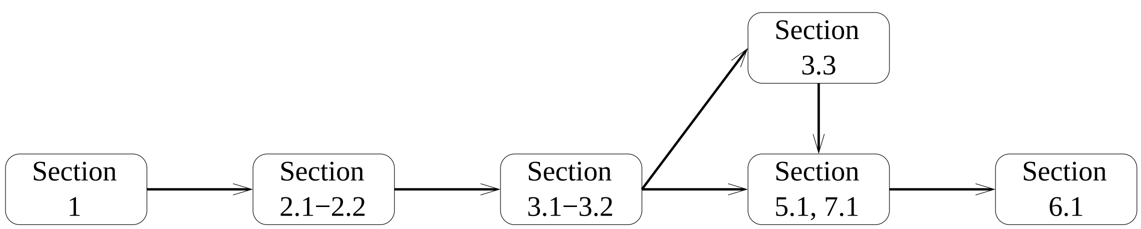
- **Goals and Objectives:** Ideas accessed:
  probabilistic independence
  Simulation as a tool to reach an initial hypothesis
  idea of threshhold boundary between two regimes
  variation due to connectivity
  interplay of connectivity and probability
  compelling pointers to unsolved problems

- **Technology/Software Needs:** If available, Excel or internet access allows for demonstrations, otherwise

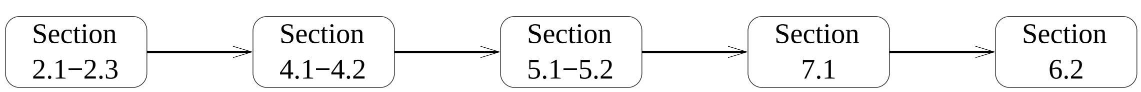
**Time:** This self contained module is designed to be completed within 50-90 minutes.

Suggested content based on level:

*High school, beginning undergraduate, non-math major:*




*Advanced Undergraduate Mathematics - Proof Based*

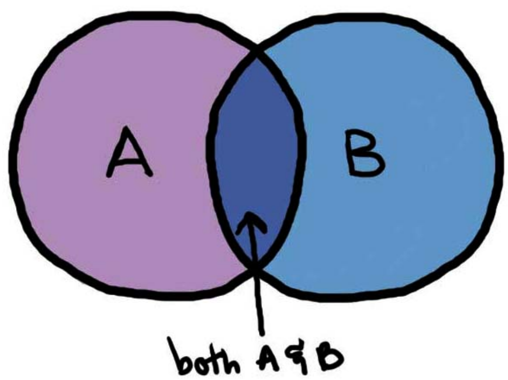


# Contents

# 1 Math Background

This module uses the basic probability rules given below. You can skip this section if you already know basic probability.

- *Independence*: if event A has probability $p_A$ and event B has probability $p_B$ and events A and B are independent (that is, whether A happens does not change the probability B happens), then the probability of A and B is $p_A p_B$.
- For two independent events A and B, the probability that either A happens or B happens (or both happen) is $p_A + p_B - p_A p_B$.



# 2 Introduction: Invasive Species and Percolation

## 2.1 Invasive Species

The term *invasive species* is used to describe animal or plant species that have colonized regions outside where they are naturally found. The introduction of non-native species can cause imbalance in the environment: if the new species finds a good food source and has no natural predators, its numbers can explode. There are many examples where this explosion of population causes major problems including driving local species to extinction either through over-consumption or through direct competition for resources.

This unit focuses on the **Emerald Ash Borer**.

(Image from Tennessee Government page
`http://www.tn.gov/agriculture/regulatory/eab.shtml`)

- A beetle introduced to North America in the 1990s, has spread to 15 states so far. The ash borer probably hitch-hiked to Michigan on wood products imported from Asia.
- **Spread of the Ash Borer:** The ash borer is able to fly short distances from infected trees to new ash trees. The ash borer infects all species of ash trees in North America, killing trees approximately three years after initial infection.
- So far this invasive species is estimated to have killed 50-100 million ash trees. It is considered one of the most destructive non-native insects in the United States.
- The timber industry produces approximately 25 billion dollars of ash saw timber per year, and ash trees are planted extensively in urban neighborhoods throughout the Southeast. The ash borer threatens an estimated 7.5 billion ash trees in North America.

**Problem:** We will use ideas from probability to explore how the Emerald Ash Borer will spread through a (stylized) landscape.

**Modeling:** Foresters have determined that the probability that a tree infected with ash borers will spread the infection to nearby trees depends on the distance between the two trees. For now, assume that an infected tree will have one chance to infect the trees near it before it dies, and that the ash borers remaining at a dead tree will die from lack of food.

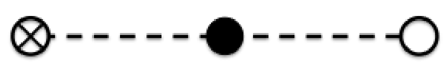
A mathematical model for the spread of an invasive species like the emerald ash borer is *percolation*. The word percolate comes from the Latin word *percolare*, meaning to be filtered through. Older adults and outdoor enthusiasts may be familiar with the word percolate because they have made coffee using this process. Others may have used the word percolate to describe the way an idea gradually spreads through a social network. Scientists use the word percolate to describe the movement of a fluid through a porous medium, such as water through soil, shale, or sandstone.

Percolation theory can be used to study the movement of anything that spreads from one discrete location to another, ignoring all the space in between. For example, the emerald ash borer moves from one tree to another. Many other applications are active areas of research in applied mathematics and engineering. Your teacher can give you some examples, or you can research them for yourself.

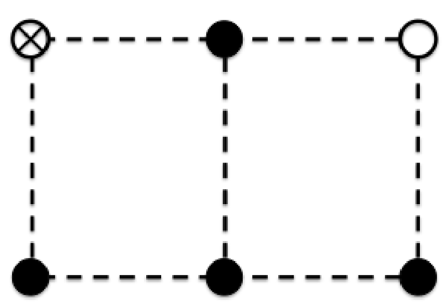
**Stop and Think:** I live on one end of a street with 3 ash trees on it (the last tree is in my yard). The trees are spaced 20 feet apart. Suppose that the probability of ash borers infecting a tree within 25 feet of an infected tree is 60% but that if the distance is more than 25 feet then the probability of transmission is 0.

I notice that the tree at the other end of the street is infected. The following figure illustrates this situation, where the X indicates the infected tree.

How could my tree become infected? What is the probability of this event?



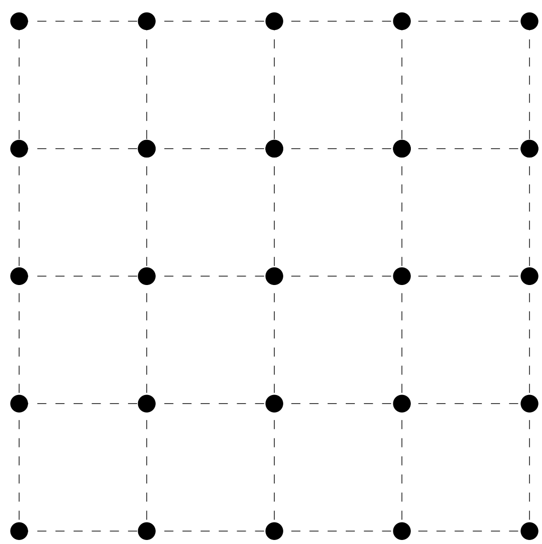
Suppose now that the next street also has 3 ash trees, as shown in the following figure. Assuming the same separation distances and probability of transmission, what are the different ways in which my tree could become infected? Is my tree more likely or less likely to be infected than in the previous setup? How might you go about finding the probability of this event?



## 2.2 Percolation

The previous examples are known as *bond percolation* because it is the bonds between the vertices which transmit the disease from one vertex to another. We say a graph *percolates* if there is a connected path from a vertex in the top row to a vertex in the bottom row (this is top-down percolation). This can represent a diseases spreading through a neighborhood of trees.

From the example in the previous section, you have discovered how tedious it can be to find the probability that a particular tree could be infected. It can be very difficult to keep track of the many possible paths for the borers to follow from the infected tree. Imagine how difficult this calculation would be if you were working with the trees in the neighborhood shown below!



**Stop and Discuss:** What do we see happening in this graph?
Is it what you would expect?
What do you think happens when the number of lattice points gets larger?

See section 3 for teacher notes on the simulations.

## 2.3 Percolation Threshold

Consider a $n \times n$ lattice. For each value of $0 \leq p \leq 1$ and any positive integer $n$, we have a probability of percolation on an $n \times n$ lattice with a transmission probability of $p$. Technically speaking we an define

$$f_n(p) = \mathbb{P}(n \times n \text{ lattice percolates given the transmission probability is } p).$$

**Stop and Think:** What do you think the value of $f_n(p)$ would be if $p$ is close to 0? Close to 1? What happens if $n$ gets large?

The *percolation threshold* is a value $0 \leq p_c \leq 1$ (if one exists) such that

$$\lim_{n \to \infty} f_n(p) = \begin{cases} 0 \text{ if } p < p_c \\ 1 \text{ if } p > p_c. \end{cases}$$

In non-technical terms, the percolation threshold is the critical point of a *very large* (infinite) graph; if the transmission probability is below the threshold, then the graph will not percolate, if it is above the threshold, then the graph will percolate.

Often times there is no particular critical value, but there is a percolation function $\lim_{n\to\infty} f_n(p) = f(p)$.

**Stop and Think:** How would we find the percolation threshold for a lattice (or other variations!)?

See section **??** for more history and information about bond percolation and thresholds.

# 3 Simulation In-class Activity

## 3.1 Simulation Background and Notes

In an activity with your class, you will learn an alternative way to study the transmission of an invasive species through a grid graph like the one above. You will use a spinner or other random outcome to decide whether or not each line segment will allow borers to spread, and then look for paths in the graph. This method is called a *simulation*, because you are "acting out" the random process of spreading from tree to tree. A simulation requires many repetitions of the process to produce a reliable estimate of the answer.

Complete section 3.2 as a handout for all the students. Then graph the results of the simulation as a function of the transmission probability $p$. (Note: The probability of the graph percolating is small for $p < .5$ and large for $p > .5$. So you will most likely see a shift in the simulation near $p = .5$.)

Bond percolation is easily simulated by randomly assigning each bond, or edge, in a square lattice or other graph to be open with probability p. By simulating percolation in a finite graph, students can quickly collect data that allows them to conjecture the value of the asymptotic phase transition probability. Handouts are provided for both bond and, as an extension, site percolation.

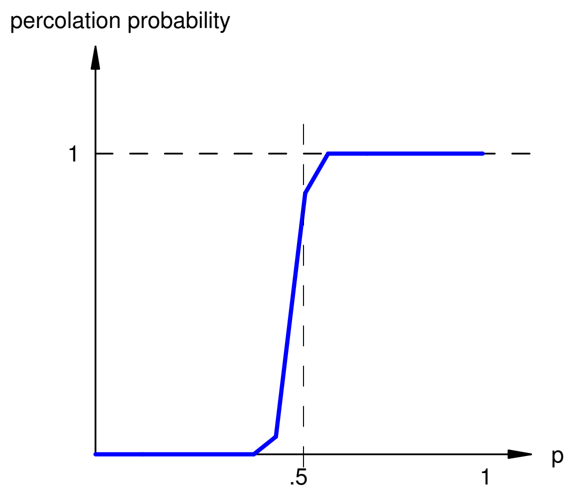
Collectively, the class should explore a range of probabilities between 0 and 1, being sure to include the value 0.5. A rule of thumb is that each probability should be simulated at least 5 times. You might rule out very small probabilities and very large probabilities after discussing threshold behavior with the class.

In a small class, you may wish to have every student simulate percolation for each probability value. In larger classes, you can accomplish the task more quickly by dividing into groups, and having each group do the simulation with a different probability. Electronic methods of simulation enable the class to explore larger lattices, a wider range of probabilities, and perform more replications of each. Alternatively, you could hand out strips of paper with pre-computed simulations, done with either of the methods described below, or project these strips of pre- computed simulations on the board for the whole class to evaluate together.

Choose one of the following methods to perform the simulation. For each probability value, track the proportion of simulations for which percolation occurs (there is an open path from any node at the top of the lattice to any node at the bottom). Put the results in a table like the following (with idealized results) on the board:

| P(open) | Percolation Proportion |
|---|---|
| 0.1 | 0 |
| 0.2 | 0 |
| 0.3 | 0 |
| 0.4 | 0.1 |
| 0.5 | 0.6 |
| 0.6 | 0.9 |
| 0.7 | 1.0 |
| 0.8 | 1.0 |
| 0.9 | 1.0 |

The module asks students to graph the results, which should look something like the graph below:




**Simulation Method 1: Pencil and Paper (Handout)** Students use a physical (game-board style) spinner with a fraction of the circle colored black corresponding to their assigned probability. (Note, however, the significant time investment required to build and flick physical spinners. A less time-consuming option that still lets students actively participate in the simulation is to use a random number generator. Use the rand() command in Excel, or use an online random number generator. For example, at `http://www.random.org/decimal-fractions/`, where they can generate all their random numbers ) to determine whether each bond is open or closed, and fill in with a solid line each open bond. Closed bonds are left as is. Once each bond is determined to be open or not, the students must determine by eye whether or not there is percolation.

**Simulation Method 2: Excel Workbook** Open the Excel file BondPerc.xlsx. Each time the probability of a bond being open is changed, the entire lattice is redrawn with the simulated open and closed bonds. Students must determine by eye whether or not there is an open path, consisting of adjacent colored edges, from any node at the top of the lattice

to any node at the bottom of the lattice.
With a smaller class, you might want to use BondPercExplore.xlsx, which lets each student with a computer generate 3 results for each of 4 different probabilities. Note that these probabilities can be modified, as in BondPerc.xlsx.

**Discussing the graph:** First, look at the extremes:

If $p$ is very close to 0: hardly any edges, not likely to percolate.

If $p$ is close to 1: lots of edges, very likely to percolate.

**What is happening between these two extremes?**

- When $p$ is less than $\frac{1}{2}$, the graph does not percolate
- When $p$ is greater than $\frac{1}{2}$ the graph does percolate.

This simulation is illustrating the concept of **percolation threshhold**, a common flavor of results in the theory of random graphs.

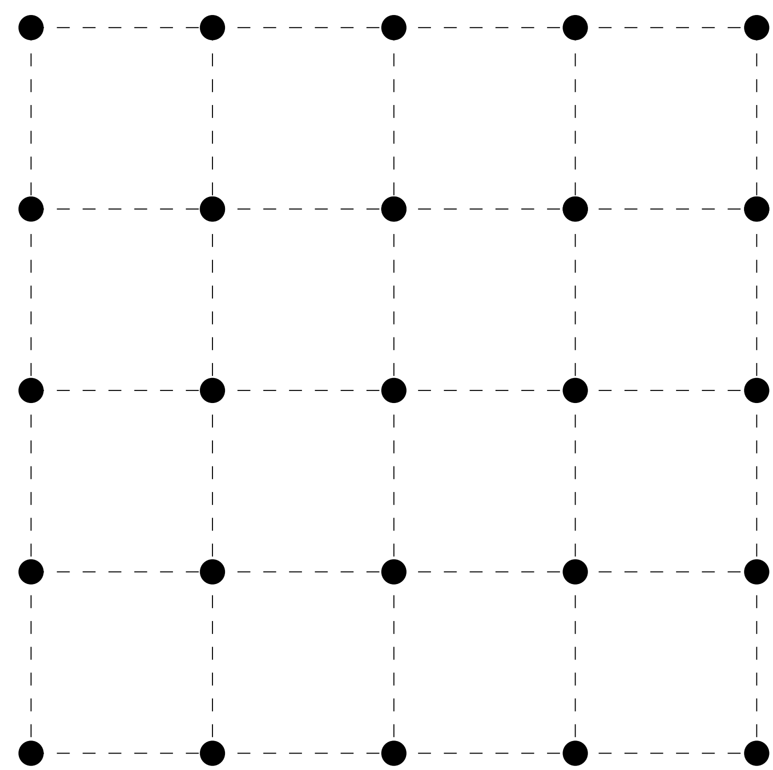
If a particular probability is above $p_c$ then one property is exhibited, if the probability is below $p_c$, some other property is exhibited. We say $p_c$ is the *threshhold* value of the property. In our case, the property is "percolation happens."

Q: What happens when the number of lattice points gets larger?
A: graph becomes increasingly sharp.

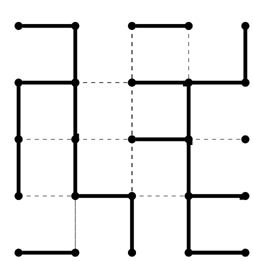
See section **??** - **??** for more information and background.

## 3.2 Bond Percolation Simulation Example

Use the spinner for each of the 40 edges on the graph. If you spin and land on black, color the edge. You must go in order (top to bottom, left to right). This represents an open edge. An example is done for you.



Example:



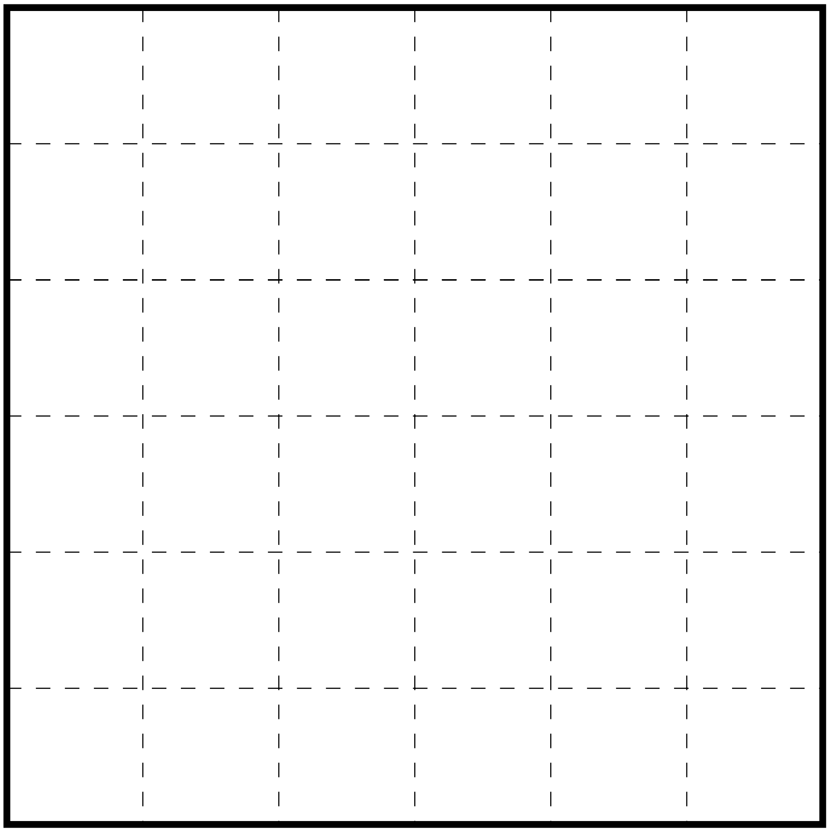
Can you find a path from the top to the bottom through the colored (open) edges? If you can, then your graph percolated.

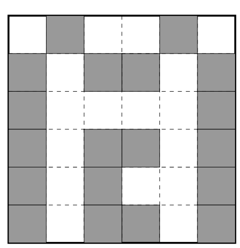
Did your random graph percolate? (Can you find a path from the top to the bottom through the colored (open) edges?)

What do you think is the *bond percolation threshold*?

### 3.3 Site Percolation Simulation Example

Use the spinner for each of the 36 blocks on the graph. If you spin and land on black, fill in the block. You must go in order (top to bottom, left to right). An example is done for you.





Can you find a path from the top to the bottom through the black boxes (open sites)? You may not go diagonal. If you can, then your graph percolated.

What do you think is the *site percolation threshold*?

# 4 Proofs

## 4.1 Discussion and Background

Question: How do we find the percolation threshold?

Finding the percolation threshold on a square lattice is quite difficult. However we can understand the percolation threshold for an infinite complete binary tree which has some similarities to the square lattice structure.

See background in section **??**.

In order for students to understand the proof, they will need a little background in:

- Probability: independent events, addition law, basic counting techniques
- Calculus: limits
- Proof techniques: induction (optional for homework assignment)
- Graph theory: See below definitions

We will now consider a complete binary tree. First we need a few definitions.

- **Tree**: A tree is an graph in which any two vertices are connected by a unique simple path.
- **Rooted Tree:** A rooted tree is a tree with a designated root vertex. All vertices have 'parent' vertices except the root vertex.
- **Full Binary Tree:** A full binary tree is a tree in which each node has exactly two children, except for *leaf* nodes which have no children.
- **Complete Binary Tree:** A complete binary tree (or a perfect binary tree) is a full binary tree such that at level $n$, there are $2^n$ vertices, where $n = 0$ is the root vertex.
- **Infinite Complete Binary Tree:** A complete binary tree with countable infinite number of levels.

## 4.2 Proof for infinite complete binary tree

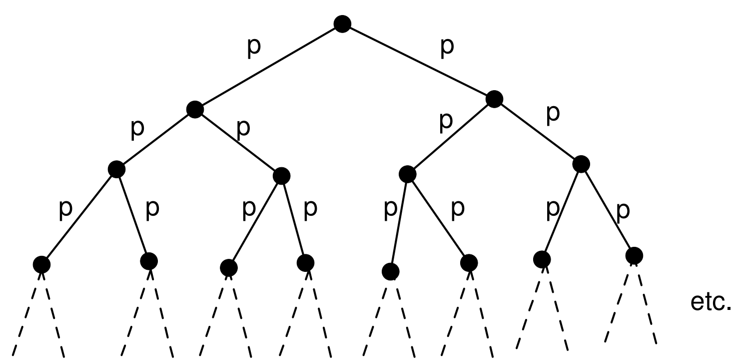
**Theorem 1.** *Let $T_n$ be a complete binary tree with $n$ levels (not including the root vertex). Let $f_n(p)$ denote the probability there is a path from the root vertex, 0, to level $n$ if the probability of an edge is $p$. Then*

$$\lim_{n\to\infty} f_n(p) = f(p) = \begin{cases} 0 \ \textit{if}\ p \leq \frac{1}{2} \\ \frac{2p-1}{p^2} \ \textit{if}\ p > \frac{1}{2} \end{cases}$$

*Proof.* Consider the following definition:

Let 0 be the root vertex of our complete binary tree. Let $1_1$ and $1_2$ be the first sub-row of the tree etc..

Let $f_n(p)$ denote the probability there is a path from the root vertex of $T_n$ to level $n$ if the probability of an edge is $p$.




We are interested in finding $\lim_{n\to\infty} f_n(p) = f(p)$.

Big picture of the proof:

Step 1: Show $f_{n+1}(p) = 2pf_n(p) - (pf_n(p))^2$. Thus $f(p) = 2pf(p) - (pf(p))^2$. Thus $f(p) = 0$ or $f(p) = \frac{2p-1}{p^2}$.

Step 2: Show if $p \leq \frac{1}{2}$, then $f(p) = 0$.

Step 3: Show if $p \geq \frac{1}{2}$, then $f(p) = \frac{2p-1}{p^2}$.

**Step 1:** Let $A_1$ be the event there is a path from 0 to level $n+1$ which goes through vertex $1_1$. Let $A_2$ be the event there is a path from 0 to level $n+1$ which goes through vertex $1_2$.

Notice that

$$f_{n+1}(p) = \mathbb{P}(A_1 \cup A_2)$$

By using the addition law for events we see that

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

Now we will compute each of the probabilities on the right hand side.

If the event $A_1$ is going to occur, we have to have the edge from 0 to $1_1$ and a path from $1_1$ to level $n+1$. The probability there is a path from $1_1$ to level $n+1$ is $\mathbb{P}(A_1) = pf_n(p)$. Similarly, if the event $A_2$ is going to occur, we have to have the edge from 0 to $1_2$ and a path from $1_2$ to level $n+1$. The probability there is a path from $1_2$ to level $n+1$ is $\mathbb{P}(A_2) = pf_n(p)$.

Also notice that the events $A_1$ and the events $A_2$ are independent because they involve disjoint pieces of the binary tree. Thus

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) = (pf_n(p))^2 .$$

Combining the above statements, we see:

$$\begin{aligned} f_{n+1}(p) &= \mathbb{P}(A_1 \cup A_2) \\ &= \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \\ &= pf_n(p) + pf_n(p) - (pf_n(p))^2 \\ &= pf_n(p)\,(2 - pf_n(p)) \end{aligned}$$

Taking limits as $n$ approaches infinity on both sides gives the quadratic equation:

$$f(p) = 2pf(p) - (pf(p))^2 , \tag{1}$$

which is a quadratic in $f(p)$ and thus has two solutions: $f(p) = 0$ or $f(p) = \frac{2p-1}{p^2}$.

This completes step 1.

**Step 2:**

If $p \leq \frac{1}{2}$ we see that $\frac{2p-1}{p^2} \leq 0$, therefore the only plausible solution to (1) is $f(p) = 0$.

(Note: This step can also be easily proved by overestimating the number of paths from 0 to level $n$ and assuming the paths are independent. There are $2^n$ paths from 0 to level $n$, and the probability of one of these paths being open is $p^n$. Therefore

$$f_n(p) \leq 2^n p^n,$$

which, if $p < \frac{1}{2}$, goes to zero as $n$ goes to infinity. This alternate approach can be applied to a homework problem. )

**Step 3:**

We will need to make two simple observation: (The first observation can be easily proved by induction - Homework!)

- In order to percolate to the $(n+1)^{st}$ level, you need to percolate to the $n^{th}$ level. Thus $f_{n+1}(p) \leq f_n(p)$.
- Since there is at least one path from the root to level $n$, we see $f_n(p) \geq p^n > 0$.

Therefore, we know from our recursion $f_{n+1}(p) = pf_n(p)\,(2 - pf_n(p))$ that

$$p\,(2 - pf_n(p)) \leq 1.$$

Solving this inequality for $f_n(p)$, we see

$$\frac{2p-1}{p^2} \leq f_n(p).$$

Taking limits as $n$ goes to infinity on both sides shows

$$\frac{2p-1}{p^2} \leq f(p).$$

Notice if $p > \frac{1}{2}$, then

$$0 < \frac{2p-1}{p^2} \leq f(p).$$

Therefore since $f(p) = 0$ or $f(p) = \frac{2p-1}{p^2}$ by (1), we know $f(p) = \frac{2p-1}{p^2}$ for all $p > \frac{1}{2}$.

$\square$

# 5 Other Applications and Resources

Other Applications: You may use the following as discussion, handouts, or an assignment.

## 5.1 Other Applications

- Spread of disease in trees:
  Consider a farmer who would like to plant the trees in an orchard in order to minimize the spread of disease between trees, yet maximize the yield from the orchard. The increased distance between trees represents a smaller probability of the spread of a disease. What is the optimal distance and lattice structure desired?

- Forest Fires:
  Often forest rangers would like to be able to predict how far and how quickly a fire would be able to spread. Giving a time component to a simple percolation model allows forest rangers to make these predictions based on wind speed and density of the forest.

- Oil Field:
  Often gas or oil is found insides porous rocks. The pores in rock form a network in which the oil or gas flows. Percolation models are used to predict how much oil or gas can be found in rocks of different porosity.

- Electrical network grid:
  Electricity is passed from one component to another through connections (edges) which could represent power lines. For example, there is a nice structure for the power grid in a neighborhood of houses, or they may represent larger power lines connecting neighborhoods in a particular city. In order for power to pass from one city to the next, the power must percolate through the connections. What configuration should we use for the network in order to make it reliable yet cost effective?

- Communication network and social media:
  Recently information has been able to spread through the use of social media and communication networks. Often the spread of text messages, tweets, etc can be analyzed through the use of a percolation network on graphs which model social network structures. The model percolates as information is passed from one person to another.

- Epidemiology:
  The transmission of disease through a particular species is of great importance. If a particular individual or group becomes infected, how long will the disease propagate? Transmitability and virulent are two key factors in this model. The big question: Will there be an epidemic?

- Child immunization:
  Children that get immunizations for particular diseases act as a buffer for the spread of disease. If the percentage of children who receive a particular vaccine drops below a particular point, then the probability of an outbreak in children who are not immune increases dramatically.

- Gelatinous substance:
  As a substance forms, bonds are made between neighboring chemicals. This formation allows the liquid to become a gelatinous substance. For example: the process of boiling an egg. These small clusters eventually bond together to form larger and larger molecules.

- Structural integrity of material:
  Most materials have imperfections in them. These imperfections or impurities often make a substance weaker. Under a large amount of stress a crack often forms between these impurities. To percolate, the substance would have a crack from one end to the other, thus breaking the substance. How 'pure' must we make our material in order to have a particular strength?

- Groundwater flow:
  As water flows through the soil, it percolates through the soil layers by moving through cracks and capillaries. This flow of water can be studies by a percolation model for different types of soils.

- Rumors:
  Given a social network structure, if a rumor is started with an individual how far is will it percolate? The social network structure would be highly dependent on the number of friends each person has and the strength of the friendship.

- Others: Lightning, Brine Ice Formation, underground lava flow.

## 5.2 Other Resources

References:

For more information on percolation we suggest the following texts. This list is not meant to be an extensive list, but serve as a beginners reference list.

1. 'Introduction to Percolation Theory' - Dietrich Stauffer and Ammon Aharony
2. 'Percolation' - Bela Bollobas and Oliver Riordan
3. 'Percolation' - Geoffrey Grimmett
4. 'Applications of Percolation Theory' - M Sahimi

In addition, you can find several apps online which demonstrate different percolation models.

# 6 Homework

Homework 1 is designed for high school, non-mathematics majors, or beginner level college courses.

Homework 2 is designed for a proof based course with a little in probability.

## 6.1 Homework 1

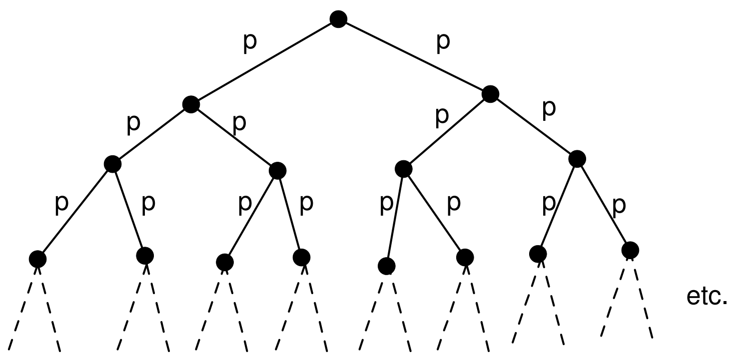
1. Suppose that a new invasive species, the *Ruby Oak Chomper*, is mistakenly introduced in Texas. The *Ruby Oak Chomper* spreads by attaching itself to car hubcaps. Suppose that every state in the USA introduces border checks to inspect the hubcaps of all vehicles crossing the border and this procedure succeeds in stopping transmission from an infected state to a neighboring uninfected state with probability 70%. Will Washington's Oaks become infected by the *Ruby Oak Chomper*?

   Use the following map of the USA to model this situation as a percolation problem in a graph.

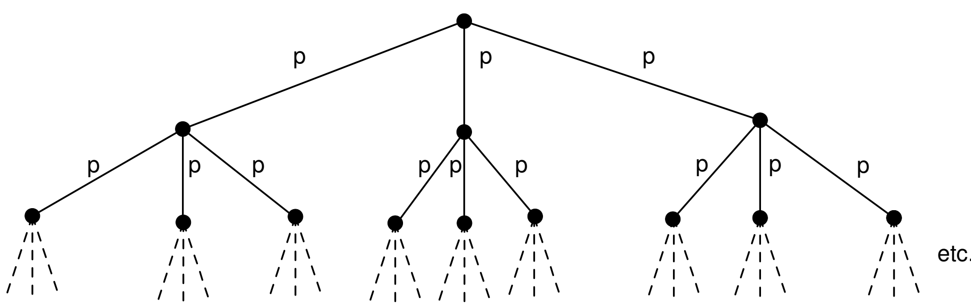
- What is the right graph (what are the nodes, edges)?
- What are the transmission probabilities?
- How many states are in the shortest path the *Ruby Oak Chomper* can take from Texas to Washington?
- What is the probability that the *Ruby Oak Chomper* infestation will travel this path?

2. In this problem you'll compare the percolation threshholds for two special types of graphs.

   The first kind of graph is a *binary tree* (the word *binary* means 2). Starting from a single node at the top, each node has two child nodes in the level below it, each of these child nodes also has two child nodes of its own in the level below it, etc:

   The second kind of graph is a *ternary tree* (the word *ternary* means 3). Starting from a single node at the top, each node has three child nodes in the level below it, each of these child nodes also has three child nodes of its own in the level below it, etc:

   Suppose that you have a binary tree and a ternary tree which each have the same number of levels. Each tree becomes infected at its single top node, and the probability of transmission along each edge is a value $p$. We'll say that a tree *percolates* when a node in its lowest level gets infected.

   - Think about slowly increasing $p$ from 0 to 1: in which tree will the infection percolate first? Why?
   - Which tree has the lower percolation threshhold?

## 6.2 Homework 2

1. Consider the percolation function defined by $f_{n+1}(p) = 2pf_n(p) - (pf_n(p))^2$. If $f_0(p) = 1$, prove
$$f_{n+1}(p) \leq f_n(p).$$
2. If we did not require the binary tree to be complete, how might this change the value of the percolation threshold? Support your answer.
3. Assume we consider a infinite complete ternary tree (each vertex has degree 4 except the root vertex. Prove that if $p < \frac{1}{3}$, then the probability of the graph percolating is zero. (i.e. $f(p) = 0$ for $p < 1/3$).
4. Find a generalization of the previous problem for any infinite complete $m$-ary tree.
5. Based on the previous problem, why does finding the exact value of $f(p)$ for a complete $m$-ary tree for all values of $p$ become increasingly difficult as $m$ increases?

# 7 Teacher's Notes

## 7.1 Generalizations and Brief History

**Bond Percolation:**

The mathematical concepts of percolation were first introduced by Broadbent and Hammersley (1957: "Percolation processes I. Crystals and mazes.") Given a $n \times n$ lattice, what is the probability a path forms from the top of the lattice to the bottom of the lattice if each edge is present with probability $p$ independent of the other edges? As is often the case, it is easier to compute the probability of percolation (forming an infinite cluster) assuming an infinite lattice structure, i.e. $n$ tends to infinity. Based on this model, for any given value of $p$, Kolmogorov's zero-one law tells us that the lattice either percolates with probability 1 or 0. Therefore there is some critical value $p_c$ so that if $p < p_c$, then the model will percolate with probability 0, and if $p > p_c$, then the model will percolate with probability 1. In a very celebrated result, Harry Kesten (1982 - Percolation theory for mathematicians) proved that the critical value for the square lattice $\mathbb{Z}^2$ was $p_c = 1/2$.

**Site Percolation:**

A similar questions can be asked for site percolation; each site is open with probability $p$ and closed with probability $1 - p$. Is there a path from the top of the lattice to the bottom through open sites? For the square lattice $\mathbb{Z}^2$, bounds have been able to show that $p_c \approx .59$, although an exact answer is still unknown! (This is a great way to show students that mathematics is alive and people are still working on very interesting and useful problems.)

**Other Lattice Structures:**

Other generalizations include different lattice structures. Many of these lattice structures have exact results while others do not. In most simple cases, if the number of neighbors increases, this implies there are a larger number of possible percolation paths, and therefore the value of $p_c$ tends to decrease. However this may not be the case on lattice structures which have some vertices have very few neighbors. The wikipedia page on percolation threshold has a very comprehensive list of 2D lattice structures and bond and site percolation thresholds and bounds. (`http://en.wikipedia.org/wiki/Percolation_threshold#Thresholds_on_other_2d_lattices`)

**Higher Dimensional Variants:**

Percolation models have also been studied in higher dimensional variants. For example, consider the 3D square lattice where each vertex on the interior has 6 neighbors. Although exact values and bounds are more difficult to prove, there have been some results for simple structures. In almost all cases, the higher dimensional percolation thresholds $p_c$ decrease significantly because of the increased number of possible paths because of the lattice structure.

## 7.2 Homework 1 solutions

1. **The Ruby Oak Chomper:**

   The graph has a node for each state and an edge between each two state nodes which are adjacent. (Drawing the graph on top of the map is fine).

   Each edge has transmission probability 0.7 (or 70%).

   The shortest path I can find has four states (in addition to Texas and Washington). For example, a four state path would be: Texas to Arizona, Arizona to New Mexico, New Mexico to California, California to Oregon, Oregon to Washington. (Students may just draw a path on the map).

   Each edge in the path transmits with probability 0.7, so the probability that the path transmits is $(0.7)^5$.

2. **Binary vs. Ternary:**
   As $p$ increases, the ternary tree will percolate first because it has more potential paths to transmit along than the binary tree does.

   Thus, the ternary tree has a lower percolation threshold.

### 7.3 Homework 2 Solutions

1. Consider the percolation function defined by $f_{n+1}(p) = 2pf_n(p) - (pf_n(p))^2$ for $0 < p < 1$. If $f_0(p) = 1$, prove
$$f_{n+1}(p) \leq f_n(p).$$

   **Solution:** Base Case: Since $f_0(p) = 1$, then we can see that $f_1(p) = 2p - p^2$. Notice that $0 \leq (1-p)^2$. If we expand $(1-p)^2$ we see $f_1(p) = 2p - p^2 \leq 1 = f_0(p)$.

   Induction. Assume $f_{n+1}(p) \leq f_n(p)$. Show $f_{n+2}(p) \leq f_{n+1}(p)$.

   Simple algebra shows $f_{k+1}(p) = 1 - (1 - pf_k(p))^2$. Since $f_{n+1}(p) \leq f_n(p)$, we can multiply both sides by $-p$ and add 1 to get
$$1 - pf_{n+1}(p) \geq 1 - pf_n(p).$$

   Since both of these quantities are negative, we can square both sides and preserve the inequality:
$$(1 - pf_{n+1}(p))^2 \geq (1 - pf_n(p))^2 .$$
   Multiplying by (-1) and adding one gives
$$1 - (1 - pf_{n+1}(p))^2 \leq 1 - (1 - pf_n(p))^2 .$$

   However the left hand side is exactly $f_{n+2}(p)$ and the right hand side is $f_{n+1}(p)$, thus completing the proof.

2. If we did not require the binary tree to be complete, how might this change the value of the percolation threshold? Support your answer.

   **Solution:** If you decrease the number degree of a particular vertex, you are decreasing the total number of possible paths from the root vertex to the bottom. Therefore it would be harder to percolate. Thus more edges would have to be open to have a path from the top down. Thus we would expect $p_c$ to increase since there are fewer possible percolation paths.

3. Assume we consider a infinite complete ternary tree (each vertex has degree 4 except the root vertex. Prove that if $p < \frac{1}{3}$, then the probability of the graph percolating is zero. (i.e. $f(p) = 0$ for $p < 1/3$).

   **Solution:** Let $f_n(p)$ denotes the probability of reaching the $n^{th}$ level in a complete ternary tree where each edges is independently open with probability $p$. In a complete ternary tree, there are $3^n$ possible paths from the root vertex to the $n^{th}$ level. For any one particular path, the probability that path is open is $p^n$. Therefore
$$f_n(p) \leq 3^n p^n.$$

   So taking limits on both sides as $n$ goes to infinity shows if $p < \frac{1}{3}$, then $f(p) = 0$.

4. Find a generalization of the previous problem for any infinite complete $m$-ary tree.

   **Solution:** Similarly you can show $f_n(p) \leq m^n p^n$. So if $p \leq \frac{1}{m}$, we can see $f(p) = 0$.

5. Based on the previous problem, why does finding the exact value of $f(p)$ for a complete $m$-ary tree for all values of $p$ become increasingly difficult as $m$ increases?

   **Solution:** The recursive structure makes it increasingly difficult to solve. You will have to imply the inclusions exclusion principle inorder to effectively solve for larger values of $p$.