RANDOM WALKS AND PERCOLATION: AN ANALYSIS OF CURRENT RESEARCH ON MODELING NATURAL PROCESSES

AARON ZWIEBACH

Abstract. In this paper we will analyze research that has been recently done in the field of discrete mathematics, specifically relating to modeling natural processes. We examine, summarize, and explain the proofs related to the following natural processes: Percolation and Random Walks. We present the main results of two papers and provide the proofs to demonstrate how combinatorics, expected value, and generating functions, among other tools, can be used to model such natural processes using discrete mathematics.

1. INTRODUCTION

Both papers presented in this revew were written by: Hochstadt, Padilla, and Zwiebach, students of Mathematics at the Massachusetts Institute of Technology. The research papers are of interest because they tackle the issues of modeling natural processes in nature via mathematics. The two main methods of analysis are combinatorial proofs and empirical analyses. The two random processes discussed in the papers are: Random Walks and Percolation. We provide attachments containing the original paper with the complete set of results, but for the purposes of this survey paper, we focus on the main results of each paper- those that deal with finding proof based models of the processes. More specifically the structure of the paper will be as such:

(1) Introduction
   - Description of Random Walks in $\mathbb{Z}^2$
   - Description of Percolation.
(2) Proofs
   - Main result from Analysis of Random Walks in $\mathbb{Z}^2$
   - Main Result from Percolation
(3) Discussion/Conclusion

1.1. **Random Walks.** In this particular verison of Random Walks, we examine a "walker" taking random steps in $\mathbb{Z}^2$, the two dimensional lattice. Our walker takes a step at each discrete time interval t and must go in in one of the four cardinal directions (left, right, up, or down.) More precisely, if our walker finds themselves at location (x, y) at time t, then at time $t + 1$, they will be at either $(x, y \pm 1)$ or $(x \pm 1, y)$. Since this is a random process we stipulate that the walker goes in any of the four cardinal directions with probability $\frac{1}{4}$. Since the walker cannot stay in the same spot, they take a step at each time interval with probability 1.

In the paper that we are reviewing, we explore the dynamic between our walker and a series of boundary points. We define a boundary point as a point, (x', y'), that once reached ends the random walk. Imagine our walker finds themsemselves initialized at the center of a square boundary and begins the random walk, with what probability would they reach the boundary at any given time step, t? What is the walker's expected time until "absorbtion"? What is the likelihood that our walker exits at one particular points as opposed to the other? How could we use computer simulations to answer the previous question? How do the previous results vary when we vary the shape of the boundary?

These are all natural questions that arise when examining the question of the walker in a bounded area. In **Analysis of Random Walks on $\mathbb{Z}^2$**, Hochstadt, Padilla, and Zwiebach examine all of these questions. Given our desire to examine methods of modeling natural processes using discrete math, we will present the proof for one of the main results of the paper. Specifically we will look at the behavior of a walker intialized in a $5x5$ square boundary, and explore the question of the walker's probability of absorption at any given time step and their expected time until absorption. We cite the following theorems:

Theorem 1. Given a 5x5 array and an initial location $(0, 0)$, the probability that our walker will be absorbed by the boundary at any time step $n \geq 2$ is given by:
$$P[Absorbtion] = 2^{-1-k}$$
Where k is given by a function $F(n)$ that maps the set of even numbers and the set of odd numbers to the set of natural numbers:
$$k = F(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{n-1}{2} & \text{if } n \text{ is odd} \end{cases}$$

Theorem 2. The expected time until absorbtion for a walker starting at the center of a 5x5 lattice is $\frac{9}{2}$ time steps.

The reader may now be asking themselves: why did we choose to present a proof on such a particular shape as a $5x5$ square boundary? As we shall see in the proof, the authors not only provide us with the two theorems listed above, they also provide us with a general yet powerful framework for how to answer this question for any shape or dimension of boundary. For that reason we may look at their method for this particular instance and extrapolate this method to different instances with different dimensions.

1.2. **Percolation.** The second natural process that we will examine in this summary paper is Percolation. In nature percolation is the process by which a liquid moves and is filtered through a porous material. In their paper, Hochstadt, Padilla, and Zwiebach, bring the example of rain falling onto a patch of soil. As the rainwater lands on the surface, it seeps through the naturally formed paths in the soil until it is stopped. This process is called percolation and it is the way that the roots of plants are watered!

So, how can we model percolation using discrete mathematics? We do this by using graphs, specifically trees. We examine the paths accessible from a root vertex. We say that a node is accessible if there is a path comprised of "open edges" leading to it. At this point we should define our terms more precisely. We imagine that we have a root vertex, v, of a tree. We also imagine that this tree has degree n, which means that there are n edges emanating from it leading to n nodes. Futhermore, each one of those n nodes has n edges emanating from it leading to n distinct nodes (in a tree there are no cycles.) If we assume this tree is infinite then this process goes on forever.

For the purposes of percolation we assume that we have this infinite tree, and we start at an arbitrarily chosen root vertex. We "open" each adjacent edge with probability p and "close" it with probability $1 - p$. If we close an edge, the we stop the process along that path and consider all of its descendents unreachable. If we open an edge then we repeat the process for all of its descendents. The process can either stop when we close all candidate edges or it can go on ad infinitum.

We call a tree in which we open edges with probability p, a percolation graph with edge probability **p**. We measure the "cluster" size of our graph by counting all the reachable nodes via open edges. **Figure 1** gives an example of this process:
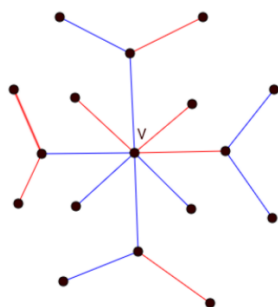

FIGURE 1. In this graph, blue edges are open and red edges are closed. The vertex v is a member of a cluster of size 8

In the paper we explore the behaviors on a specific kind of tree, trees of valence d. Valence trees are those in which every edge is adjacent to exactly d other edges. **Figure 2** gives an example of a tree of valence 3:
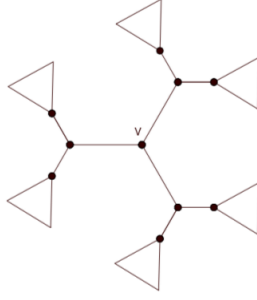
FIGURE 2. The tree of of valence 3

We present the main result from: **Percolation**. Namely, we show how to derive a closed form function for the size of a cluster formed in a percolation tree given a valence d and an open edge probability p.

Theorem 3:*In the graph of valence d with open edge probability p, the expected size of a cluster containing an arbitrary vertex v is:*

$$E[|\text{cluster size}|] = \begin{cases} \frac{1+p}{1-dp+p} & \text{if } p \in [0, \frac{1}{d-1}) \\ \infty & \text{if } p \in [\frac{1}{d-1}, 1] \\ 0 & \text{otherwise} \end{cases}$$

1.3. **Summary of Introduction.** In this section we have presented two problems that resemble natural processes. We have shown how they are relatable to discrete mathematics and by extension claimed that as such we can use combinatorics and probability to generate interesting results. In the next section we present two proofs from recent papers that we feel represent well this method of analysis and serve well as an example for those looking to model their own natural processes through discrete mathematics.

2. WALKER IN A 5x5 BOUNDARY

In this section we cite one of the main results from **Analysis of Random Walks in $\mathbb{Z}^2$** by Hochstadt, Padilla, and Zwiebach. The result shown is quoted exactly as it appears in their paper. We present the result here as it appears in their paper because of the clarity with which they present their concept would be difficult to replicate. In this proof they derive through induction the probability that the walker is absorbed at any time step t in a $5x5$ boundary. Additionally, they prove that the expected time until absorption is $\frac{9}{2}$ times steps. As we stated in the introduction, the proofs are of particular interest because they are based on probabilistic and combinatorial arguments, rather than on simulations and analysis of the data.

Proof as presented in Section 4 of the referred paper:

We have a random walker moving in a two-dimensional plane. At any given time step, the walker can move in any of the four cardinal directions (up,

down, left, and right) each with probability $\frac{1}{4}$.

In order to calculate an explicit formula for the expected time until the walker is absorbed, we first need to derive a formula for the probability that the walker is absorbed by the boundary at any time t. We plot for a few time steps t the probability that our walker will occupy a given location in the lattice. For our purposes we will consider a 5x5 array (Table 1.) We initialize our walker in the center of the plane, $(0,0)$ and consider the square boundary with corners: $(-2,-2), ...(2,-2), ...(2,2), ...(-2,2)$.

TABLE 1. Position of the walker at $t = 0$

		X		

At first the walker is positioned at the center of the array and the first step it takes can either be up, down, left or right, each with probability $\frac{1}{4}$ (Shown in Table 2.)

TABLE 2. Possible positions of walker at $t = 1$

		$\frac{1}{4}$		
	$\frac{1}{4}$		$\frac{1}{4}$	
		$\frac{1}{4}$		

Given the initial location of the walker and the size of the array, it is impossible for the boundary to absorb the walker at $t = 1$. At time $t = 2$, though, the walker can could be absorbed at four distinct points on the boundary, arriving at each with probability $\frac{1}{16}$ (Seen in Table 3.)

TABLE 3. Possible positions of walker at $t = 2$

		$\frac{1}{16}$		
	$\frac{1}{8}$		$\frac{1}{8}$	
$\frac{1}{16}$		$\frac{1}{4}$		$\frac{1}{16}$
	$\frac{1}{8}$		$\frac{1}{8}$	
		$\frac{1}{16}$		

At this point we encounter our first non-zero probability of the walker being absorbed by the boundary. We calculate the probability by summing the individual probabilities of our walker occupying some point along the

boundary. In this case the probability that the walker is absorbed at time $t = 2$ is $\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$.

TABLE 4. Possible positions of the walker at $t = 3$

	$\frac{1}{32}$		$\frac{1}{32}$	
$\frac{1}{32}$		$\frac{1}{8}$		$\frac{1}{32}$
	$\frac{1}{8}$		$\frac{1}{8}$	
$\frac{1}{32}$		$\frac{1}{8}$		$\frac{1}{32}$
	$\frac{1}{32}$		$\frac{1}{32}$	

We expressly plot out the possible positions that the walker can occupy at $t = 3$ in Table 4. We invite the reader to create their own plots for the next few time steps $t = 4, 5, 6, 7...$, calculating at each point the probability of absorption. This will help the reader see the pattern that forms, and understand the basis for **Theorem 1**.

Theorem 3. Given a 5x5 array and an initial location $(0, 0)$, the probability that our walker will be absorbed by the boundary at any time step $n \geq 2$ is given by:
$$P[Absorbtion] = 2^{-1-k}$$
Where k is given by a function $F(n)$ that maps the set of even numbers and the set of odd numbers to the set of natural numbers:
$$k = F(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{n-1}{2} & \text{if } n \text{ is odd} \end{cases}$$

Proof. This will be a proof by induction.

Inductive Hypothesis: For $n \geq 2$ the board can have two configurations, one for even number steps and one for odd number steps. The configuration for an even time step is given by **Table 5.**

TABLE 5. Possible positions of walker at even time step $t = n_{Even}$

		2^{-3-k}		
	2^{-2-k}		2^{-2-k}	
2^{-3-k}		2^{-1-k}		2^{-3-k}
	2^{-2-k}		2^{-2-k}	
		2^{-3-k}		

The configuration for an odd time step is given by **Table 6.**

Base Cases: We have two base cases. The first case is the configuration of the board at time $t = 0$ (the configuration is given by **Table 1.**) The probability that the walker is absorbed by the boundary is 0, while the the probability that the walker remains on the lattice is 1.

The second base case is the configuration of the board at time $t = 1$. The configuration of the board at $t = 1$ is given by **Table 2.** The probability

TABLE 6. Possible positions of walker at odd time step $t = n_{Odd}$

	2^{-4-k}		2^{-4-k}	
2^{-4-k}		2^{-2-k}		2^{-4-k}
	2^{-2-k}		2^{-2-k}	
2^{-4-k}		2^{-2-k}		2^{-4-k}
	2^{-4-k}		2^{-4-k}	

that the walker is absorbed by the boundary is 0 and the probability that the walker remains on the lattice is 1.

Inductive Step: We begin by considering an odd time step $t = n + 1$, we note that $k = F(n + 1) = F(n)$. By our inductive hypothesis board configuration at time n is given by **Table 5.** The general configuration of the board at time $t = n + 1$ is given by **Table 7.**:

TABLE 7. Possible positions of walker at odd time step $t = n + 1$

	A		B	
H		I		C
	L		J	
G		K		D
	F		E	

The probability that the particle occupies position A, is given by $P(A)$. We note that:

$$P(A) = P(H) = P(G) = P(F) = P(E) = P(D) = P(C) = P(B) = \frac{1}{4} \cdot 2^{-2-k} = 2^{-4-k}$$

Additionally:

$$P(L) = P(I) = P(J) = P(k) = \frac{1}{4}*(2^{-2-k}+2^{-2-k}+2^{-2-k}) = 2 \cdot 2^{-3-k} = 2^{-2-k}$$

We note that the individual probabilities of each square correspond to those in our inductive hypothesis. By summing all the positions on the boundary, (A, B, C, D, E, F, G, H), we can calculate the probability of being absorbed by the border at $t = n + 1$:

$$P[Absorbtion] = 8 \cdot (2^{-4-k}) = 2^{-1-k}$$

We now consider odd time step $t = n - 1$ and even time step $t = n$. We note that $F(n - 1) = k$ and $F(n) = k + 1$. By our inductive hypothesis the board configuration at $t = n - 1$ is given by **Table 6.** The general configuration of the board at time $t = n$ is given by **Table 8.**:

We calculate:

$$P(I) = \frac{1}{4}(2^{-2-k} + 2^{-2-k} + 2^{-2-k} + 2^{-2-k}) = 4 \cdot 2^{-4-k} = 2^{-2-k} = 2^{-1-(k+1)}$$

$$P(E) = P(F) = P(G) = P(H) = \frac{1}{4}(2^{-2-k} + 2^{-2-k}) = 2^{-3-k} = 2^{-2-(k+1)}$$

$$P(C) = P(B) = P(A) = P(D) = \frac{1}{4}(2^{-2-k}) = 2^{-4-k} = 2^{-3-(k+1)}$$

TABLE 8. Possible positions of walker at even time step $t = n$

		A		
	E		F	
D		I		B
	H		G	
		C		

We note that the individual probabilities of each square correspond to those in our inductive hypothesis. By summing all the positions on the boundary, (A, B, C, D), we can calculate the probability of our walker being absorbed by the border:

$$P[Absorbtion] = 4 \cdot 2^{-4-k} = 2^{-1-(k+1)}$$

$\square$ $\square$

Now that we have a well defined function for the probability of absorption at any even-odd time step pair, we can expressly calculate the expected time to absorption.

Theorem 4. The expected time until absorbtion for a walker starting at the center of a 5x5 lattice is $\frac{9}{2}$ time steps.

Proof. Using our familiar formula for expectation, we know that the expected time to absorption is given by:

$$E[\text{Time until absorbtion}] = \sum_{n \geq 2}^{\infty} n \cdot P(\text{Absorbed at time } n)$$

Given that consecutive even-odd pairs map to the same k, we can simplify:

$$= \sum_{k=1}^{\infty} [(2k)P(\text{absorbed at time } k) + (2k+1)P(\text{absorbed at time } k)]$$

$$= \sum_{k=1}^{\infty} (4k+1)P(\text{absorbed at time } k)$$

$$= \sum_{k=1}^{\infty} (4k+1)2^{-1-k}$$

We can break down our above sum into familiar components that we know how to calculate:

$$= \frac{1}{2} \sum_{k=1}^{\infty} [2^{-1-k} + 4k \cdot 2^{-1-k}]$$

The first term, $\sum_{k=1}^{\infty} 2^{-1-k}$, is given by:

$$\sum_{k=1}^{\infty} 2^{-1-k} = \sum_{k=1}^{\infty} (\frac{1}{2})^k = \frac{1}{1 - \frac{1}{2}} - 1 = 1$$

And the second term, $4\sum_{k=1}^{\infty} k \cdot 2^{-1-k}$, is given by:

$$4\sum_{k=1}^{\infty} k \cdot 2^{-1-k} = \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = 4 \cdot 2 = 8$$

Multiplying our two previous sums by $\frac{1}{2}$, yields our desired value for expected time:

$$\frac{1}{2}(1+8) = \frac{9}{2}$$

$\square$
$\square$

3. Expected Size of Tree of Valence d

In this section we present the main result from the paper by Hochstadt, Padilla, and Zwiebach titled: **Percolation**. This result provided in **Section 5** of the aforementioned paper is an exciting result because using probability and combinatorics in conjunction with generating functions we are able to derive a closed form formula for the expected size of a tree with valence d and open edge probability p. This is another great example of how tools from the repertoire of discrete mathematics can be used to model and gain insights from natural processes. *Note*: For the purposes of this paper we have made minor edits to the original paper. The edits consist of eliminating all references to other sections in the orginal paper. This has been done for the ease of the reader.

Proof as presented in the reffered to paper:

3.1. **Definitions.** This section serves as a reference to rigorously define the symbols and terminology used throughout this paper.

Definition 3.1. $V_{d,p}$ is a percolation tree of valence d with open edge probability p.

Definition 3.2. $C_{d,p}$ is a cluster within the tree $V_{d,p}$. $|C_{d,p}|$ is the size of the cluster $C_{d,p}$.

Definition 3.3. $\Gamma_{d,p}$ is a d-ary percolation tree with open edge probability p. $|\Gamma_{d,p}|$ the size of the cluster within the tree $\Gamma_{d,p}$ which contains its root vertex.

3.2. **Trees of Valence d.** We note, that a tree of valence d has a useful relation to $(d-1)$-ary trees. A tree of valence d is comprised of a root vertex v, where v has degree d, and each edge attaches v to an infinite $(d-1)$-ary tree. Again, we denote the size of a cluster in a $V_{d,p}$ tree as $|C_{d,p}|$. Similarly, the size of a cluster in a d-ary tree is $|\Gamma_{d,p}|$. As such, the expected size of a cluster in a $V_{d,p}$ tree, $E[|C_{d,p}|]$, is given by

(3.1) $$E[|C_{d,p}|] = 1 + d \cdot p \cdot E[|\Gamma_{d-1,p}|].$$

3.3. **Number of Clusters of Size n in a d-Ary Tree.** We define a recursive function $t_d(n)$ that outputs the number of ways to generate a cluster of size n in a d-ary tree. Intuitively, the number of ways to generate a cluster of size n in a d-ary tree is the number of ways of generating clusters in each one of the d subtrees such that the sum of the sizes of the clusters in the sub-trees equals $n - 1$. More precisely, we can write this as

$$(3.2) \qquad t_d(n) = \sum_{0 \leq a_1, \ldots, a_d \text{ s.t. } a_1 + \ldots + a_d = n-1} t_d(a_1) \cdot \ldots \cdot t_d(a_d).$$

We utilize Equation 3.2 to define a generating function whose coefficient are the number of ways of generating a cluster of size n

$$F(x) = \sum_{n=0}^{\infty} t_d(n) \cdot x^n.$$

By inserting our expression for $t_d(n)$ into $F(x)$ we get that

$$F(x)^d = \sum_{a_1=0}^{\infty} \cdots \sum_{a_d=0}^{\infty} t_d(a_1) x^{a_1} \cdot t_d(a_2) x^{a_2} \cdots t_d(a_n) x^{a_n}$$

$$= \sum_{n=0}^{\infty} \Big(\sum_{0 \leq a_1, \ldots, a_d \text{ s.t. } a_1 + \ldots + a_d = n-1} t_d(a_1) \cdot \ldots \cdot t_d(a_d) \Big) \cdot x^{a_1 + \cdots + a_n}$$

$$= \sum_{n=0}^{\infty} t_d(n+1) \cdot x^n.$$

In order for the argument inside t_d and the exponent of x to agree we must shift our function over by one:

$$xF(x)^d = \sum_{n=0}^{\infty} t_d(n+1) \cdot x^{n+1} = \sum_{n=1}^{\infty} t_d(n) \cdot x^n = F(x) - 1$$

which implies:

$$(3.3) \qquad\qquad xF(x)^d - F(x) + 1 = 0.$$

When $d = 2$, we were able to explicitly solve for $F(x)$ by applying the quadratic formula. Solving for $F(x)$ in Equation 3.3 is more difficult. In the next section we use a relation between $F(x)$ and $G(y)$ to solve for the expected size of a cluster in a d-ary tree.

3.4. **Expected Cluster Size of a d-Ary Tree.** In order to derive a function for the expected size of of the d-ary tree we do not need to explicitly solve for $F(x)$. Rather we can use properties of a d-ary tree to derive a generating function for the probability of yielding a cluster of size n, that is dependent on $F(x)$.

A given cluster of size n is comprised of $n - 1$ open edges (the edges that connect the n vertices) and $n \cdot (d - 1) + 1$ closed edges. We can see this by considering what happens when we increase our cluster size by one vertex. When we add another vertex to our cluster we convert one *closed-edge* to an *open-edge*. At the same time, by only adding one vertex, we are adding

d closed-edges. This is because each vertex has d potential children, who in this case are impossible to reach because their edges are closed. As such, each time we add a vertex we also add $d - 1$ closed edges. So a cluster of size n "sees" $n \cdot (d - 1) + 1$ closed edges. (Where the "+1" term accounts for the $n = 1$ case.)

Utilizing this we can define a generating function $G(y)$ where the coefficient of the y^n term is the probability of having a cluster of size n in a d-ary tree:

$$G(y) = \sum_{n=1}^{\infty} p^{n-1} \cdot (1 - p)^{n \cdot (d-1)+1} \cdot t_d(n) \cdot y^n$$

$$= \frac{1-p}{p}(F(p \cdot (1 - p)^{d-1} \cdot y) - 1).$$

By rearranging the terms of our equality we can avoid having to explicitly derive $F(x)$. We do so by rearranging the terms so that

(3.4)
$$F(x) = \frac{p}{1-p} \cdot G(\frac{x}{p(1-p)^{d-1}}) + 1.$$

At this point we have a relational equation for $F(x)$ and $G(y)$, given by Equation 3.4. We can insert our new expression for $F(x)$ in terms of $G(y)$ into Equation 3.3 and simplify by letting $y = \frac{x}{p(1-p)^{d-1}}$:

$$p(1-p)^{d-1} \cdot y \cdot (\frac{p}{1-p}G(y) + 1)^d - \frac{p}{1-p}G(y) = 0.$$

Recall that if we take the derivative of a generating function for the probability of generating a cluster of size n and evaluate the derivative at the point $y = 1$, we get the total expected size of our cluster. Thus we are again looking for $G'(1)$, which is obtained by taking the derivative of $G(y)$ with respect to y and evaluate at the point $y = 1$. Note that since $G(y)$ is the generating function for the probability density function for cluster size, $G(1) = 1$ by the Law of Total Probability.

We begin by taking the derivative with respect to y:

$$-\frac{p}{1-p}G'(y)+p(1-p)^{d-1}((\frac{p}{1-p}G(y)+1)^d+y(d(\frac{p}{1-p}G(y)+1)^{d-1}\frac{p}{1-p}G'(y))) = 0$$

We evaluate at $y = 1$:

$$-\frac{p}{1-p}G'(1) + p(1-p)^{d-1}((\frac{p}{1-p}+1)^d + (d(\frac{p}{1-p}+1)^{d-1}\frac{p}{1-p}G'(1))) = 0$$

We simplify:

$$-\frac{p}{1-p}G'(1) + p(1-p)^{d-1}(\frac{p}{1-p}+1)^d + p^2(1-p)^{d-2}d(\frac{p}{1-p}+1)^{d-1}G'(y) = 0$$

We can simplify further: $(\frac{p}{1-p} + 1) = \frac{1}{1-p}$:

$$G'(1)(dp^2(1-p)^{d-2}(\frac{1}{1-p})^{d-1} - \frac{p}{1-p}) = G'(1)(\frac{p^2 d}{1-p} - \frac{p}{1-p})$$

$$G'(1) = \frac{1}{1-pd}$$

We assert that this is only true for values of p such that $G(1) = 1$, as the Law of Total Probability must still hold. This modifies our function $G'(1)$ slightly to become:

$$(5.5)\ G'(1) = E[|\Gamma_{d,p}|] = \begin{cases} \frac{1}{1-dp} & \text{if } p \in [0, \frac{1}{d}) \\ \infty & \text{if } p \in [\frac{1}{d}, 1] \\ 0 & \text{otherwise} \end{cases}$$

3.5. **Conclusion.** In the last section we derived an expression for the expected cluster size in a d-ary tree (Equation 3.5). At this point we can tie our result back to Equation 3.1 and state that

$$E[|C_{d,p}|] = \begin{cases} 1 + \frac{dp}{1-(d-1)p} & \text{if } p \in [0, \frac{1}{d-1}) \\ \infty & \text{if } p \in [\frac{1}{d-1}, 1] \\ 0 & \text{otherwise} \end{cases}$$

Slight simplification allows us to conclude that

Theorem 3 *In the graph $V_{d,p}$ the expected size of the cluster $C_{d,p}$ is:*

$$E[|C_{d,p}|] = \begin{cases} \frac{1+p}{1-dp+p} & \text{if } p \in [0, \frac{1}{d-1}) \\ \infty & \text{if } p \in [\frac{1}{d-1}, 1] \\ 0 & \text{otherwise} \end{cases}$$

4. CONCLUSION

In this paper we have examined two different natural processes. The first comprised of a random walker interacting with a series of boundaries, and the second was the process of percolation. There are a number of viable ways by which one could model a natural process mathematically. Given the probabilistic nature of these processes one natural inclination might have been to design a computational simulation that abided by the parameters of the process and run it enough times to gather statistically significant data.

The aforementioned method is particularly useful for more complicated structures and boundaries (in the case of the random walker) and for three dimensional porous materials (for the case of percolation.) While emperical analyses provide very important intuition for more complex scenarios, they lack the same mathematical definiteiveness that is provided by a combinatorial or probabilistic proof.

We chose these paper precisely because their main results are motivated by latter method of modeling and analysis. In the problem of the random walker, our authors use pattern tracking and induction to derive a function

for the probability that the walker is absorbed at a given time step. They use this result to further prove that an expected time until absorption for the same boundary. In the percolation paper our authors use the recursive characteristics of trees and generating functions to count clusters of a given size and leverage those tools to derive a closed form function for the expected size of a valence tree.

We invite the reader to explore this topic further. The two papers: **An Analysis of Random Walks in** $\mathbb{Z}^2$ and **Percolation** are attached for the reader to pursue the topic in further detail. The reader will find not only the same method of analysis applied to different shapes of boundaries and instances of valence trees, but also further exploration into the method of empirical analysis, a topic not explored in depth in this paper.

Happy Learning!

Massachusetts Institute of Technology, Department of Mathematics, Cambridge, Massachusetts, United States

E-mail address: aaronz@mit.edu