PRINCIPLES OF COMPLEX SYSTEMS

HW06 WRITE-UP

David W. Landay
University of Vermont
Graduate Student, Comp. Systems and Data Sci.

0.1 Problem 1

$1-d$ theoretical percolation problem:
Consider an infinite 1-d lattice forest with a tree present at any site with probability p.

(a) Find the distribution of forest sizes as a function of p. Do this by moving along the $1-d$ world and figuring out the probability that any forest you enter will extend for a total length l.

Imagine the $1-d$ lattice like in the example below:

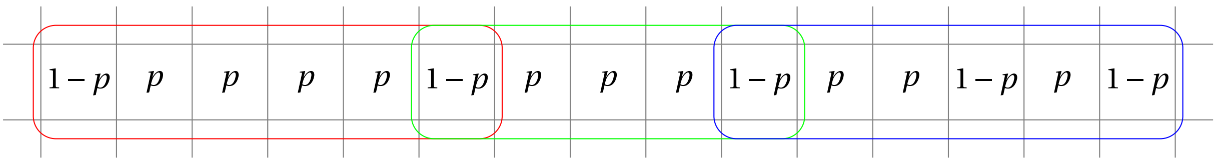


Figure 1: Trees on an infinite $1-d$ lattice: forests are defined as a string of consecutive trees ("p"s) separated by sites that contain no trees ("$1-p$"s). The colored boxes indicate how forests are separated in the lattice.

Sites on the lattice contain a tree with some probability p, and strings of consecutive trees represent forests. From the figure above, two things immediately jump out; that the probability of seeing a forest that extends for length l is the product of all of the probabilities, "p"s, where a tree exists is p^l, and that all forests must be preceded and followed by a site that contain no tree. This is because we ignore boundaries on an infinite lattice, and each site is independent of any other site. Hence, the number of forests of length l, is given by

$$n_l(l, p) = p^l (1 - p)^2$$

(b) Find p_c, the critical probability for which a giant component exists. Hint: One way to find critical points is to determine when certain average quantities explode. Compute $< l >$ and find p such that this expression goes boom (if it does).

In the case of an infinite $1-d$ lattice, a giant connected component would be a very large forest, or a large string of "p"s. In fact, the largest connected component on the lattice would be a tree at every site on the lattice; i.e. a forest of infinite length. We wish to find the critical probability p_c, that creates this scenario. Since the length of the lattice is infinite, we expect that all sites on the lattice must be trees in order to see a forest of length l for all poossible forests. Hence, we find that $n_l(l, p) = 1$ for an infinite $1-d$ lattice . Thus, we can rewrite our finding from the above as follows:

$$1 = p^l (1 - p)^2$$

$$\implies 0 = \ln(1 - p)^2 (p^l)$$

$$= \ln(1 - p)^2 + \ln(p^l)$$

$$= 2 \ln(1 - p) + l \ln(p)$$

$$\implies -2 \ln(1 - p) = l \ln(p)$$

$$\implies \frac{-2 \ln(1 - p)}{\ln(p)} = l$$

We can see that if l is infinite, then the only explanation is that $p \to 1$. Hence, the critical threshold for p, p_c, is equal to 1.

0.2 Problem 2

Show analytically that the critical probability for site percolation on a triangular lattice is $p_c = \frac{1}{2}$.

Imagine a triangular lattice like in the example below:

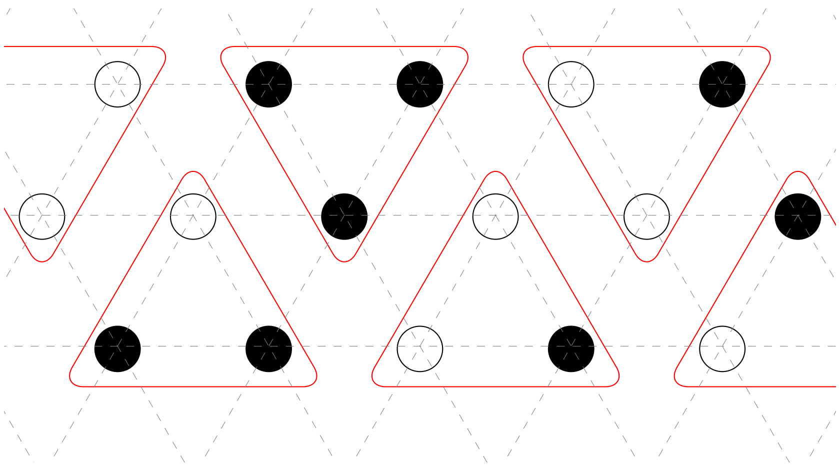


Figure 2: Percolation on a triangular lattice: Black nodes allow flow, red triangles represent groupings of nodes, or percolation sites on the space.

Sites where at least two of the three nodes are permeable, black nodes, allow material to percolate. As with the case of the $1-d$ lattice, individual nodes allow flow with some probability p. Hence, there are several configurations whereby percolation can occur; namely, the case when all of the nodes in a grouping are black, and all other cases where two of the three nodes are black. The probability that a site (a triangular grouping of nodes) is percolating is thus given by the sum of probabilities of each of the outcomes described above.

$$\begin{aligned}
\mathbf{P}' &= \mathbf{Pr}\begin{pmatrix} & \circ & \\ \bullet & & \bullet \end{pmatrix} + \mathbf{Pr}\begin{pmatrix} & \bullet & \\ \circ & & \bullet \end{pmatrix} + \mathbf{Pr}\begin{pmatrix} & \bullet & \\ \bullet & & \circ \end{pmatrix} + \mathbf{Pr}\begin{pmatrix} & \bullet & \\ \bullet & & \bullet \end{pmatrix} \\
&= p^2(1-p) + p^2(1-p) + p^2(1-p) + p^3 \\
& \text{where } \mathbf{P}' = 1 \text{ if } p = 1, \text{ and } \mathbf{P}' = 0 \text{ if } p = 0
\end{aligned} \tag{1}$$

So, we find that the probability of percolation at a site $\mathbf{P}'$ is a function of the probability of a node allowing flow p, given by $\mathbf{P}' = f(p)$. If we simplify the sum, we get

$$\mathbf{P}' = p^3 + 3p^2(1-p)$$

$$= p^3 + 3(p^2 - p^3)$$

$$\implies \mathbf{P}' = 3p^2 - 2p^3 ,$$

indicating $\mathbf{P}' = f(p) = 3p^2 - 2p^3$.

Thus, to determine the critical value at which percolation stops across the triangular lattice, we must find $\mathbf{P}' = f(p_c) = 3p_c^2 - 2p_c^3$. From 1, we know that $\mathbf{P}' = p_c$. Hence, we have that

$$p_c = 3p_c^2 - 2\,{}_c p^3$$

$$\implies 0 = p_c\left(3p_c - 2p_c^2 - 1\right)$$

$$\implies 0 = 3p_c - 2p_c^2 - 1$$

$$\implies 0 = (p_c - 1)(2p_c - 1)$$

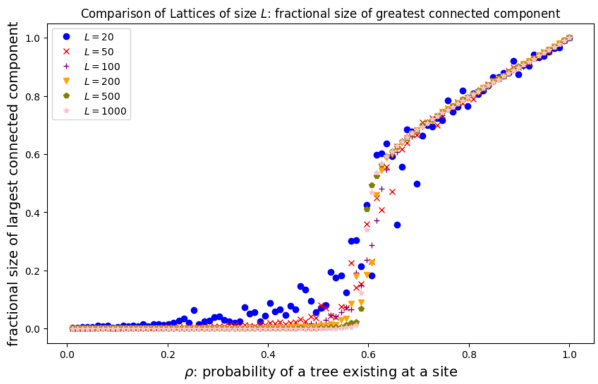
$$\implies p_c = \frac{1}{2} \therefore$$

Thus, $p_c = \frac{1}{2}$ over the triangle lattice

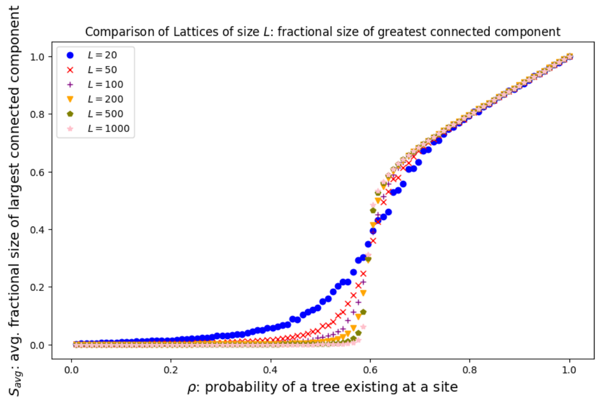
0.3 Problem 3

Percolation in two dimensions $(2-d)$ on a simple square lattice provides a classic example of a phase transition.

A site has a tree with probability p and a sheep grazing on what's left of a tree with probability $1-p$. Forests are defined as any connected component of trees bordered by sheep, where connections are possible with a site's four nearest neighbors on a lattice. Each square lattice is to be considered as a landscape on which forests and sheep co-exist.



(a)



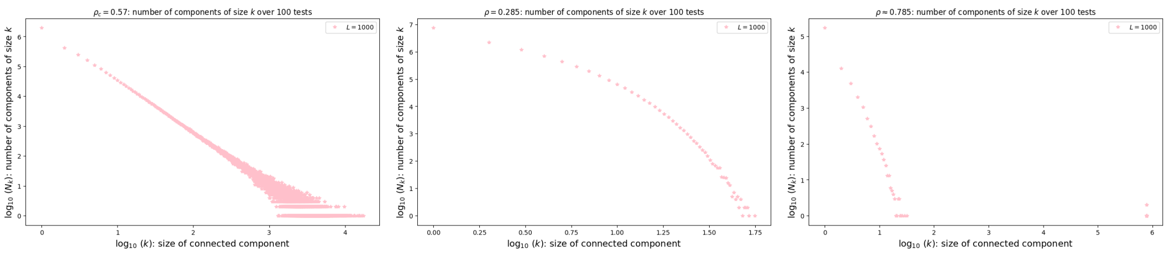
(b)

Figure 3: The two plots show the fractional size of the largest connected component in the site percolation model as a function of ρ; the probability of a tree existing at a site on the lattice. (a) shows the result of one simulation, while (b) describes S_{avg}; the average fractional size of the largest connected component over 100 simulations.

The plot reveals that on average, as the size of the lattice L increases, the probability of a tree existing at a site, ρ, approaches a clear threshold ρ_c, whereby the probability of a tree existing at any site on the lattice will either be very small, or will be present with increasing likelyhood. In other words, below the threshold, ρ_c, percolation almost never occurs, and almost always occurs above it. This is because a lower site probability will equate to smaller connected components, and for large L such connected components will occupy less of the overall lattice. As L increases, it becomes more obvious what ρ_c is equal too. Based on plot (b) in **Figure 3**, we estimate that $\rho_c \approx 0.57$.

0.4 Problem 4

If we estimate that $\rho_c = 0.57$, then we can get a sense of how the number of forests scales with forest size. Unsurprisingly, the relationship demonstrates a clear power-law size distribution around this value of ρ. The plots below demonstrate how N_k, the number of components of size k, behaves as a function of k, the size of connected components, at our estimated ρ_c, well below it, and above it respectively. **Figure 4** (a) indicates that we expect to find many more connected components of smaller size for $\rho \approx \rho_c$, than larger connected components.



(a) **(b)** **(c)**

Figure 4: Behavior of N_k, the number of connected components of size k, as a function of k for $\rho = \rho_c$, $\rho < \rho_c$, and $\rho > \rho_c$.s

Figure 4 (b) and (c) show that when $\rho < \rho_c$, we do not see any connected components over a certain size. Whereas in Plot (c) when $\rho > \rho_c$, we have fewer instances of large connected components, but we have a higher probability of seeing them. This notion is reminiscent of *Kolmogorov's zero-one law*, which says that the outcome of some event, called a *tail event*, will either "certainly" happen or "certainly" not happen ($p = 1$, or $p = 0$). It specifically states that for any infinite sequence of independent random variables, a tail event is independent of some finite subset of the random variables. We clearly see this in **Figure 3** and **Figure 4**, where outcomes around the threshold (a finite subset of our random variables) become increasingly uncertain, but where percolation will not occur for any $\rho < \rho_c$, and will occur for any $\rho > \rho_c$.

0.5 Problem 5

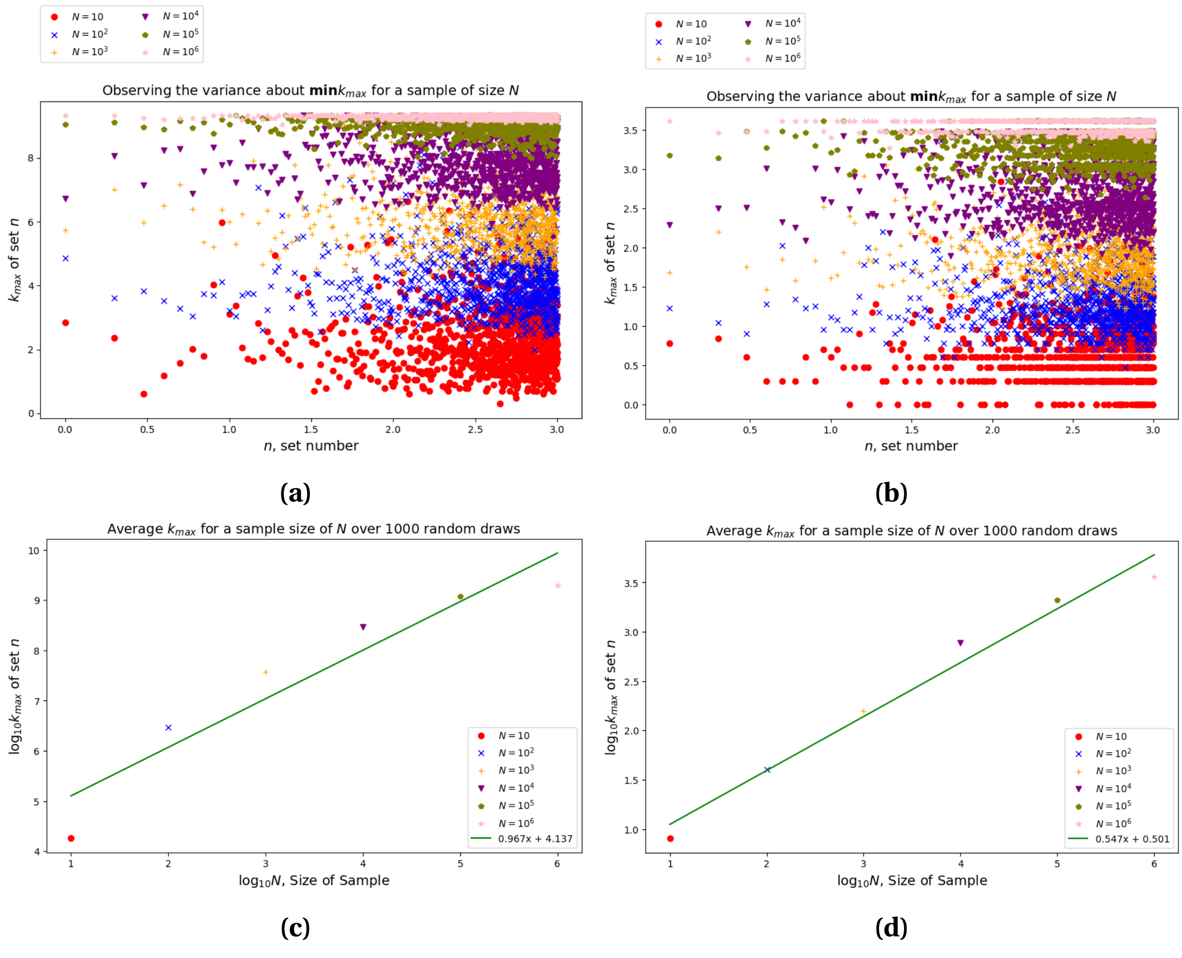


Figure 5: Plots (a) and (c) correspond to $\gamma = \frac{3}{2}$, while plots (b) and (d) correspond to $\gamma = \frac{5}{2}$. The top most plots describe the max size of a component, k_{max} as a function of set size n.

We can test our estimate of the expected value for **min** k_{max} for a fixed γ. Let $\gamma = \frac{3}{2}$. We can obtain a rough sense of the expected value of k_{max} as a function of N by generating many, $n = 1000$, samples for increasing size, $N = 10, 10^2, 10^3, 10^4, 10^5, 10^6$ using

$P_k = c\,k^{-\gamma}$. We will generate P_k using the Riemann Zeta function.
We see from the least squares regression, that our distribution of Avg.k_{max} fits our estimate.