Lecture 23: An Introduction to Percolation

François Blanchette

Department of Mathematics, MIT

May 19, 2001

1 The Model

We introduce a model for percolation which applies on any graph. However, the graphs usually considered are simple lattices (cubic, triangular). The process we are interested in is a random coloring of the graph. We simply occupy each site with probability p. We are then interested in clusters of points, particularly their shapes and sizes, and their statistics in general. Note that p plays the role of a concentration.

2 Properties

2.1 Spanning probability

We define $R_N(p)$ to be the spanning probability of a system of size N. By this we mean the probability that a single cluster bridges from one side to the other. For any finite N, this is an analytic function which starts at 0 when $p = 0$ and goes up to 1 when $p = 1$. However, as $N \to \infty$, the function sharpens to a step function, with jump at $p = p_c$, the critical point. This critical point depends on the graph in general. It is known exactly for only a few graphs, as will be discussed later.

2.2 Strength of the infinite cluster

We define $P_\infty(p)$ to be the strength of the infinite cluster. That is to say that $P_\infty(p)$ is the expected fraction of the number of sites which are part of the infinite cluster. This only makes sense when $N \to \infty$. Notice that we say here "the" infinite cluster as if it was clear that there was only one. In fact, there is only one infinite cluster only when $p > p_c$. By definition, when $p < p_c$ there is no infinite cluster. At the critical point, it is possible to have more than one infinite cluster. In fact, the probability to have m infinite clusters goes like e^{-m^α}. At $p = p_c$, the infinite cluster is a fractal object and it thus occupies a volume that tends to zero. The strength of the infinite cluster will thus be zero for $p \leq p_c$. This point physically corresponds to a phase transition. Because of the continuity of P_∞, it is

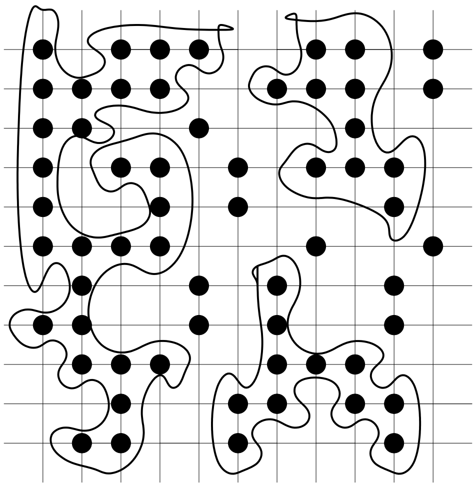

Figure 1: Percolation on a square grid

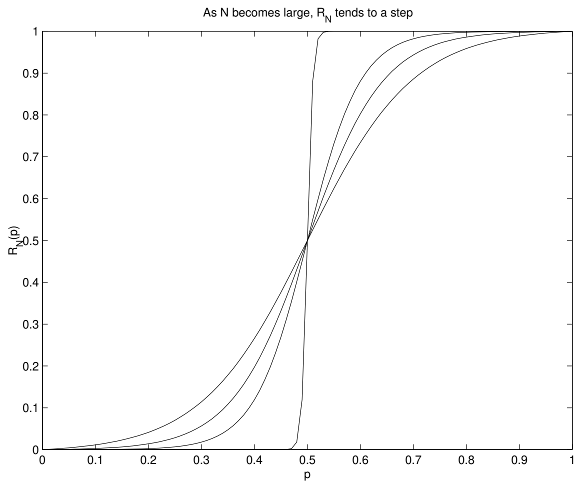

Figure 2: As N increases, the spanning probability tends to a step function

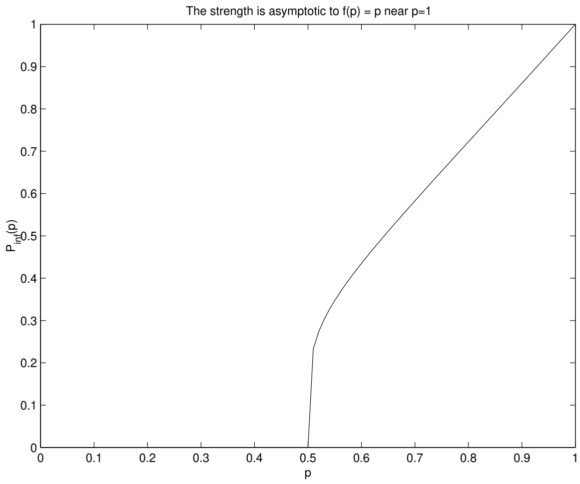

Figure 3: The strength is asymptotic to $f(p) = p$ near $p = 1$, and is 0 for $p \leq p_c$

called a second order phase transition (by opposition to first order ones which correspond to a discontinuity). Around p_c, the strength behaves like $(p - p_c)^\beta$ but it is not certain what β is. As p approaches 1, the strength asymptotically approaches the line $f(p) = p$, which corresponds to having all the occupied sites part of the infinite cluster. It is also possible to study P_∞/p and find equivalent results.

2.3 Correlation length

Let us consider the correlation length $\xi(p)$, that is to say the length over which the probability that two points are part of the same cluster is greater than some fixed probability. The correlation function $g(r)$ is the probability that two points which are a distance r apart are connected. For $p \leq p_c$, we find $g(r) \approx e^{-r/\xi}$ if we ignore boundary effects. As $p \to p_c$, $\xi \to \infty$. This is to be expected since two points arbitrarily far apart have a finite probability to be part of the same cluster for $p = p_c$. The exponent ν is define to be such that $\xi \approx (p - p_c)^{-\nu}$ as $p \to p_c$.

For larger values of p, the correlation length is meaningless, so we consider instead its counterpart ξ' which is the length over which to points have a finite probability to be connected, but not through the infinite cluster. Once again, if $h(r)$ is this probability for points that are a distance r apart, then $h(r) \approx e^{-r/\xi'}$. Also ξ' goes to infinity as p approaches p_c since the clusters which are not infinite get larger and larger in this limit.

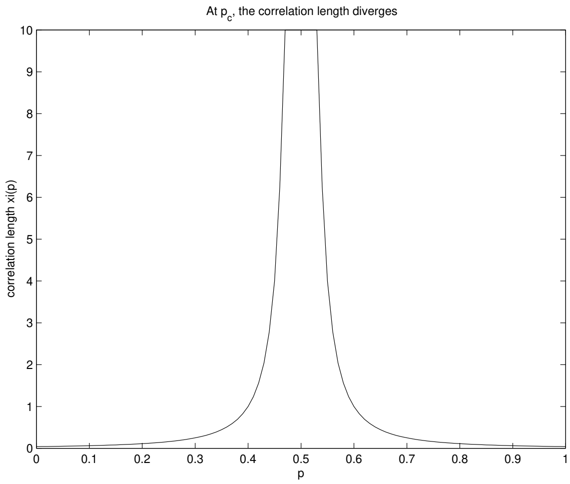

Figure 4: The correlation length goes to ∞ as $p \to p_c$

3 Applications

3.1 Polymer gelation

Polymers can be seen as clusters formed by the process described above. This application was first mentioned by Flory in 1941. As solvent is removed, usually by means of heating, the concentration of polymers increases, which corresponds to increasing p. The boiling of an egg is the most common example of such a phenomenon.

3.2 Flow in disordered media

This application was introduced by Hammersley in 1957. The point of interest here is to figure out how a fluid forced through a random media would flow. A similar problem is to predict the flow of electrons moving from atoms to atoms in a disorder media. This motion is similar to a continuous time random walk in the sense that electrons spend a random time on a given atom and then take a step. However, the possible steps are also constrained in space, and this restriction does not evolve with time. One can also associate the progression of forest fires to this class of applications, where the fire acts like the fluid. Social sciences also have some use for this model to describe the evolution of epidemics. Lastly, one can study the resilience of the Internet to failures using a similar model.

4 Invasion Percolation

This model differs from the previous ones in this that it assigns not simply 0 or 1, but any number between 0 and 1, uniformly. Each vertex is thus labelled with a value p, which can be seen as an effective porosity. One then grows a cluster from a starting point by "invading"

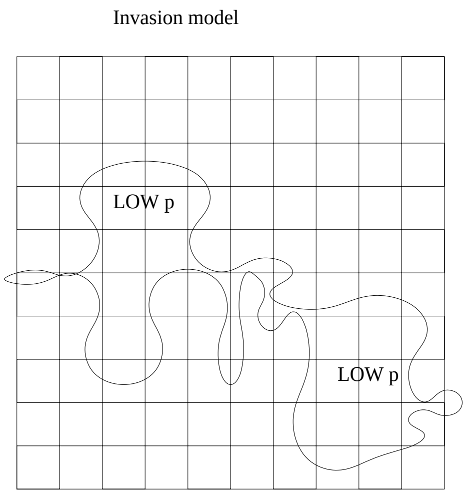

Figure 5: Basin of low p form, linked together through regions of higher p

the neighbouring point with the smallest value. This process is iterated and one ends up with basins of "low p" linked to each other through small regions of high p. This leads to self-organized criticality. The basins formed are self-similar, and if this process is repeated until the formation of an infinite cluster, then the resulting cluster is of the same type as the percolation clusters at $p = p_c$.

5 Percolation Thresholds

The value of p_c varies with the graph under consideration. As a general rule, the critical point decreases as the dimension of the graph increases because higher dimensionality offers more ways to get around an unoccupied region. Also, if the coordination number of the graph increases (the number of neighbours of each vertex) then p_c decreases for similar reason. Here is a table of various simple graph and the values of p_c found by simulations (except the triangular lattice and the Cayley tree which are analytical).

Graph	Dimension	Coord. number	p_c
Chain	1	2	1
Honeycomb	2	3	0.6962
Square lat.	2	4	0.592746
Triangular lat.	2	6	1/2
Tetrahedron	3	4	0.43
Cubic lat.	3	6	0.3116080
B.C.C.	3	8	0.246
F.C.C.	3	12	0.198
Cayley tree	∞	z	$1/(z-1)$

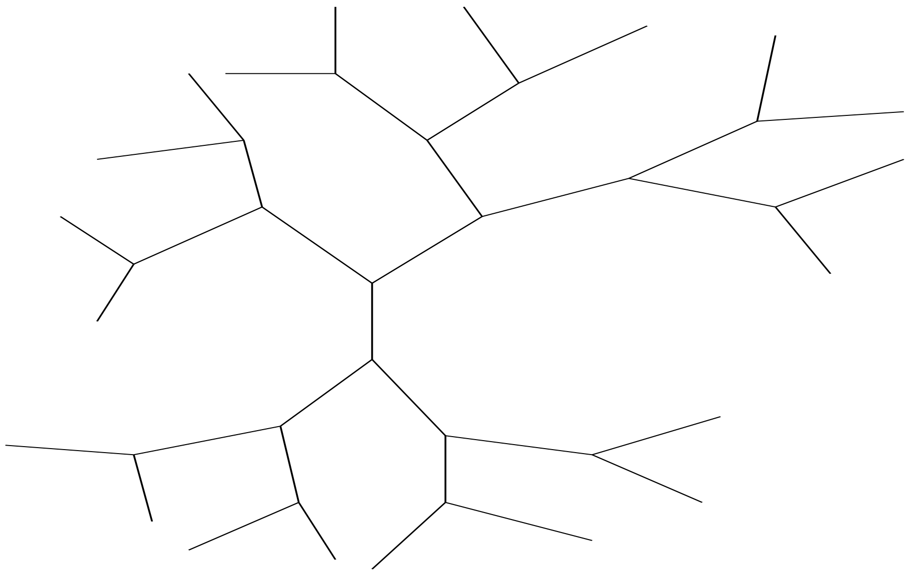

Figure 6: The Cayley tree with $z = 3$

Here the abbreviation B.C.C means bounds centered cubic, and F.C.C means face centered cubic.

5.1 Cayley Tree

The Cayley tree is simply a tree where all the vertices have the same coordination number, z. Because it is a tree, it has no loops and its analysis is very much simplified.

Let us consider the case $z = 3$. Why do we say that the Cayley tree has infinite dimension? Well if we compute the ratio of surface sites to the total number of sites, we get, at a distance m from the source:

$$\frac{\text{Surface sites}}{\text{Total sites}} = \frac{3 \cdot 2^{m-1}}{1 + 3(1 + 2 + 2^2 + .. + 2^{m-1})} \to \frac{1}{2}\text{as } m \to \infty$$

However, in general the volume scale like L^d and the surface like L^{d-1} so that their ratio is $1/L$ and tends to zero, for any finite d. Therefore, the Cayley tree is said to have infinite dimension. The use of studying such a graph comes from the "Bethe approximation" in which all loops are neglected, effectively replacing the graph with a tree.

The critical point of a Cayley tree can be found in the following way. To have an infinite cluster, we start from a pair of connected points. For either of these points, the expected number of sites where the cluster could grow is $p(z-1)$. It can thus be shown that if $p \leq 1/(z-1)$, the cluster will have finite length with probability 1 and if $p \geq 1/(z-1)$, there is a finite probability to have an infinite cluster. Therefore, $p_c = 1/(z-1)$.

Using a similar approach, we can also derive an expression for $P_\infty(p)$. Let Q be the probability that the origin is not part of an infinite cluster, given that one o its neighbour is occupied. Then we have the recursion formula:

$$Q = (1-p) + pQ^{z-1}.$$

Here we suppose that one of the neighbours of the origin is occupied and we want to see what is the probability that the cluster can grow to infinity. The first term is the probability that the origin itself is not occupied. The second term is the probability that all the other $(z-1)$ branches are not extending to infinity. This has to be multiplied by the probability that the origin is occupied, p. One can then solve for Q. For example, the case $z=3$ yields $Q=1$ or $Q=\frac{1-p}{p}$. The case $Q=1$ corresponds to $p \leq p_c$ where there is no infinite cluster possible. Using this, we can now compute $P_\infty(p)$ for $p \geq p_c$. The probability that a given site is part of an infinite cluster is thus

$$P_\infty(p) = p(1-Q^z) = p\left(1-\left(\frac{1-p}{p}\right)^3\right) \text{ for } z=3$$

5.2 Renormalization group estimates

The Renormalization group techniques, used in statistical physics, also have applications in percolation. At p_c, the system is self-similar (since $\xi \to \infty$, which allows us to do the following trick: replace a group of points by a "macro-point", and consider only the system made of macro-points. To decide if a macro-point is occupied, one looks at the group of original points. If they span the region, then the macro-point is occupied. The direction in which the points have to span the region can be taken arbitrarily since it makes no difference in the large size limit. Notice that this idea works as long as $\xi \geq N$.

If we now call $R_n(p)$ the probability that a macro-point of size n is occupied, we have that p_c is a fix point of $R_n(p)$.

$$R_n(p_c) = p_c$$

Consider for example the case of a two by two square lattice that we wish to convert to a macro-point. If we assume that the macro-point is occupied if and only if it is spanned vertically, we get:

$$R_4(p) = p^4 + 4p^3(1-p) + 2p^2(1-p)^2$$

Here the first term is the probability that all the sites are occupied. The second term is the probability that three sites are occupied, which can happen in four ways. The third term is the probability that two vertically aligned sites are occupied, which can happen in two ways. To figure out the critical point of such a system, we look for fix points of $R_4(p)$. Solving $R_4(p) = p$ gives four solutions:

$$\begin{aligned}
p &= 0 \\
p &= 1 \\
p &= -\frac{\sqrt{5}+1}{2} \\
p &= \frac{\sqrt{5}-1}{2}
\end{aligned}$$

The first two are attracting fix points to which the system will converge if it $p \neq p_c$. The third one is negative and thus has no probabilistic meaning. The last one corresponds to the critical point. It is an unstable fix point, as can be expected. This happens to be the golden mean, and it's numerical value is about $p_c^{(4)} = 0.618$. This is remarkably close, for such a simple estimate, to the best known numerical value of p_c for a square lattice, $p_c = 0.592726$. Performing the same calculations for larger macro-points gives estimates of p_c that are increasingly close to what is found numerically. However, this calculation becomes much harder very quickly.