Lecture Notes # 15

Daniel M. Romero (dmr239) and Cameron Allen (cpa9)

February 20, 2009

1 The Fibonacci Numbers

The Fibonacci numbers are a sequence of natural numbers starting with $0, 1, 1, 2, 3, 5, 8, \cdots$. In general, each new number is obtained by adding the previous 2 numbers in the sequence.
We can define the Fibonacci Sequence f_n by the following recurrence relation:

$$f_0 = 0$$
$$f_1 = 1$$
$$f_n = f_{n-1} + f_{n-2} \text{ for } n \geqslant 2$$

We would like to get the corresponding generating function:

$$f(x) = \sum_{i=0}^{\infty} f_i x^i \tag{1}$$

Since we have $f_i = f_{i-1} + f_{i-2}$ for $i \geqslant 2$ then we can sum over all $i \geqslant 2$ and multiply times x^i to get:

$$\sum_{i=2}^{\infty} f_i x^i = \sum_{i=2}^{\infty} f_{i-1} x^i + \sum_{i=2}^{\infty} f_{i-2} x^i$$

We can write each term of the above equation in terms of $f(x)$:

$$f(x) - x = x^2 f(x) + x f(x)$$

We can now solve for $f(x)$ in the above equation to get the generating function in closed form:

$$f(x) = \frac{x}{1 - x - x^2} \tag{2}$$

It is easy to verify by long division that equation 1 in fact equals equation 2. We can also use partial fraction decomposition to write equation 2 as the sum of two fractions with linear denominator. The result of this procedure is:

$$f(x) = \frac{\sqrt{5}/5}{1 - \phi_1 x} + \frac{-\sqrt{5}/5}{1 - \phi_2 x}$$

where $\phi_1 = \frac{1+\sqrt{(5)}}{2}$ and $\phi_1 = \frac{1-\sqrt{(5)}}{2}$ The Taylor expansion of $f(x)$ gives:

$$f(x) = \frac{\sqrt{5}}{5}(1+\phi_1 x+(\phi_1 x)^2+(\phi x)^3+\cdots)-\frac{\sqrt{5}}{5}(1+\phi_2 x+(\phi_2 x)^2+(\phi_2 x)^3+\cdots)$$

Since $f(x)$ is the generating function of the Fibonacci sequence then the above equation implies that:

$$f_n = \tfrac{\sqrt{5}}{5}(\phi_1^n - \phi_2^n) \ \forall n \geqslant 1$$

Note that since $|\phi_2| < 1$ and $\phi_1 > 1$ then ϕ_1 tells you how fast the Fibonacci sequence grows. Also, since $|\phi_2^n| < 1 \ \forall n \geqslant 1$ and the Fibonacci only contains integers then $f_n = \lfloor \frac{\sqrt{5}}{5} \phi_1^n \rfloor \ \forall n \geqslant 1$

2 Model for a Growing Graph

In many realistic situations where networks are a useful tool we find that in the beginning the graph has only a few vertices and it grows as time goes by. The model $G(n,p)$ fails to capture this characteristic since the number of vertices, n, is fixed for the very beginning. Because of this, it is important to study a different model that allows for the total number of vertices to vary. The following procedure yield a very simple model in which the total number of vertices increases as time increases:

- Choose $\delta \in [0,1]$
- At $t = 0$ there are no vertices.
- At every unit of time add a new vertex.
- Choose two nodes uniformly at random and add an edge between them with probability δ.[1]

2.1 Degree Distribution

Since the total number of vertices and edges changes as time increases then the expected number of vertices with a certain degree will be a function of t. Let $d_k(t)$ be the expected number of vertices of degree k.
First we calculate $d_0(t+1)$. The expected number of vertices of degree 0 at time $t+1$ will be the number of them we had at time t plus one vertex we add at time $t+1$. But since we will also add an edge at time $t+1$ with probability δ then we might lose some. Therefore $d_0(t+1) = d_0(t) + 1 - 2\delta\frac{d_0(t)}{V(t)}$, where $V(t)$ is the total number of vertices at time t. But by the construction of the graph $V(t) = t$. Hence we get[2]:

$$d_0(t+1) = d_0(t) + 1 - 2\delta\frac{d_0(t)}{t} \tag{3}$$

[1]Note that this model allows for self-loops since at $t = 1$ there is only one vertex. However when the total number of vertices is large the probability of choosing the same vertex is very small, therefore we will assume that there are no self-loops

[2]A more precise calculation would use the total number of vertices of degree 0 at time t instead of $d_0(t)$ but this would make the analysis of the model much more complicated. However, because the actual number of vertices of degree k is concentrated around the expected value this approximation is, in fact, very accurate

By a similar analysis we get the formula for $d_k(t+1)$. Note that this time we can lose or gain more vertices of degree k by adding an edge at time $t+1$, hence the extra term, Also, we no longer gain a vertex of degree k when we add a vertex to the graph, hence we don't add 1:

$$d_k(t+1) = d_k(t) + 2\delta\frac{d_{k-1}(t)}{t} - 2\delta\frac{d_k(t)}{t} \tag{4}$$

Next we try to get a solution for $d_k(t)$ of the form $d_k(t) = p_k t$. Note that $d_0(t+1) = (t+1)p_0$, $d_0(t) = (t)p_0$, and $\frac{d_0(t)}{t} = p_0$, and substituting these in equation 3 we get:

$$(t+1)p_0 = p_0 t - 2\delta p_0 + 1 \iff p_0 = \tfrac{1}{1+2\delta}$$

Similarly, we can repeat the above to equation 4 and obtain p_k:

$$(t+1)p_k = p_k t + 2\delta p_{k-1} - 2\delta p_k \iff p_k = \tfrac{2\delta}{1+2\delta}p_{k-1}$$

Applying the above formulas to $p_{k-1}, p_{k-2}, \cdots, p_0$ we obtain:

$$p_k = (\frac{2\delta}{1+2\delta})^k \frac{1}{1+2\delta}$$

2.2 Distribution of Components

We would like to find the generation function for the size of the component containing a randomly chosen vertex. It is important to note the following: Assume we have a graph that contains 100 components, 50 are of size 4 and 50 of which are of size 2. If we choose a component at random then the probability that it is of size 4 is $\frac{1}{2}$. However, if we choose a vertex and random, the probability that it belongs to a component of size 4 is not $\frac{1}{2}$. Since there are 50 components of size 4 and 50 components of size 2 then there are a total of 300 vertices, 200 of which belong to components of size 4 and 100 of which belong to components of size 2. Therefore the probability that a randomly chosen vertex belongs to component of size 4 is $\frac{2}{3}$. This example shows that choosing a random vertex and asking what size component it belongs to is not the same as choosing a random component and asking what size it is. Let $N_k(t)$ be the expected number of components of size k. Then we get the following equations:

$$N_1(t+1) = N_1(t) + 1 - 2\delta \frac{N_1(t)}{t} \tag{5}$$

$$N_k(t+1) = N_k(t) + \sum_{j=1}^{k-1} \left(\frac{jN_j(t)}{t} \frac{(k-j)N_{k-j}(t)}{t} \right) - 2\delta \frac{kN_k(t)}{t} \tag{6}$$

The second term of the above equation accounts for the components of size j that can be connected to components to a component of size $k-j$. The third term accounts for components of size k that were connected to another components.[3]

We will now conduct the same procedure from the previous section. We will get a solution of the form $N_k(t) = a_k t$. As in the previous section, equations 6 and 5 yield:

$$a_1 = \frac{1}{1+2\delta}$$

$$a_k = \frac{\delta}{1+2\delta k} \sum_{j=1}^{k-1} (j(k-j)a_j a_{k-j})$$

The number a_k tells you the fraction of components of a given size therefore it is proportional to the probability that if a component is selected at random it is of size k. However, one can check that $\sum_{k=1}^{\infty} a_k \neq 1$, hence a_k is not a probability. However, we note the following:

$$\sum_{k=1}^{\infty} k a_k = \sum_{k=1}^{\infty} k \frac{N_k(t)}{t} = \frac{1}{t} V(t) = \frac{1}{t} t = 1$$

[3]Note the implicit assumption we are making: a vertex in a component will not be connected to another vertex in the same component. This implies that this calculation only works before the a giant component is formed.

Therefore ka_k is the probability that if a vertex is chosen at random it will belong to component of size k. Consider the generating function $g(x) = \sum_{k=1}^{\infty} ka_k x^k$. We will show the following equality:

$$g = -2\delta x g' + 2\delta x g g' + x \tag{7}$$

To do so we will follow a similar procedure to the one used for the Fibonacci sequence. Recall that $a_1(1+\delta) - 1 = 0$ and $a_k(1+2\delta k) = \delta \sum_{j=1}^{k-1} j(k-j)a_j a_{k-j}$ so multiplying through by kx^k and adding we obtain:

$$-x + \sum_{k=1}^{\infty} ka_k x^k + 2\delta x \sum_{k=1}^{\infty} a_k k^2 x^{k-1} = \delta \sum_{k=1}^{\infty} \left(kx^k \sum_{j=1}^{k-1} j(k-j)a_j a_{k-j} \right)$$

Notice that the right hand side of the above equation is $-x + g(x) + 2\delta x g'(x)$ so we start seeing the resemblance with equation 7.

TO BE CONTINUED...