Chapter 37

Random Walks I

By Sariel Har-Peled, February 17, 2022

“A drunk man will find his way home; a drunk bird may wander forever.”

Anonymous,

37.1. Definitions

Let $G = G(V, E)$ be an undirected connected graph. For $v \in V$, let $\Gamma(v)$ denote the set of neighbors of $v$ in $G$; that is, $\Gamma(v) = \{u \mid uv \in E(G)\}$. A random walk on $G$ is the following process: Starting from a vertex $v_0$, we randomly choose one of the neighbors of $v_0$, and set it to be $v_1$. We continue in this fashion, in the $i$th step choosing $v_i$, such that $v_i \in \Gamma(v_{i-1})$. It would be interesting to investigate the random walk process. Questions of interest include:

(A) How long does it take to arrive from a vertex $v$ to a vertex $u$ in $G$?

(B) How long does it take to visit all the vertices in the graph.

(C) If we start from an arbitrary vertex $v_0$, how long the random walk has to be such that the location of the random walk in the $i$th step is uniformly (or near uniformly) distributed on $V(G)$?

Example 37.1.1. In the complete graph $K_n$, visiting all the vertices takes in expectation $O(n \log n)$ time, as this is the coupon collector problem with $n - 1$ coupons. Indeed, the probability we did not visit a specific vertex $v$ by the $i$th step of the random walk is $\frac{1}{e^{(i-1)/n}} \leq 1/n^{10}$, for $i = \Omega(n \log n)$. As such, with high probability, the random walk visited all the vertex of $K_n$. Similarly, arriving from $u$ to $v$, takes in expectation $n - 1$ steps of a random walk, as the probability of visiting $v$ at every step of the walk is $p = 1/(n - 1)$, and the length of the walk till we visit $v$ is a geometric random variable with expectation $1/p$.

37.1.1. Walking on grids and lines

Lemma 37.1.2 (Stirling’s formula). For any integer $n \geq 1$, it holds $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$.

37.1.1.1. Walking on the line

Lemma 37.1.3. Consider the infinite random walk on the integer line, starting from 0. Here, the vertices are the integer numbers, and from a vertex $k$, one walks with probability $1/2$ either to $k - 1$ or $k + 1$. The expected number of times that such a walk visits 0 is unbounded.

Proof: The probability that in the $2i$th step we visit 0 is $\frac{1}{2^i} \binom{2i}{i}$. As such, the expected number of times we visit the origin is

$$\sum_{i=1}^{\infty} \frac{1}{2^i} \binom{2i}{i} \geq \sum_{i=1}^{\infty} \frac{1}{2\sqrt{i}} = \infty,$$

since $\frac{2^i}{2\sqrt{i}} \leq \binom{2i}{i} \leq \frac{2^i}{\sqrt{2i}}$ [MN98, p. 84]. This can also be verified using the Stirling formula, and the resulting sequence diverges. \hfill \Box

\section*{Notes}

\footnote{This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.}
37.1.1.2. Walking on two dimensional grid

A random walk on the integer grid \( \mathbb{Z}^d \), starts from a point of this integer grid, and at each step if it is at point \((i_1, i_2, \ldots, i_d)\), it chooses a coordinate and either increases it by one, or decreases it by one, with equal probability.

Lemma 37.1.4. Consider the infinite random walk on the two dimensional integer grid \( \mathbb{Z}^2 \), starting from \((0,0)\). The expected number of times that such a walk visits the origin is unbounded.

Proof: Rotate the grid by 45 degrees, and consider the two new axises \(X'\) and \(Y'\). Let \(x_i\) be the projection of the location of the \(i\)th step of the random walk on the \(X'\)-axis, and define \(y_i\) in a similar fashion. Clearly, \(x_i\) are of the form \(j/\sqrt{2}\), where \(j\) is an integer. By scaling by a factor of \(\sqrt{2}\), consider the resulting random walks \(x'_i = \sqrt{2}x_i\) and \(y'_i = \sqrt{2}y_i\). Clearly, \(x_i\) and \(y_i\) are random walks on the integer grid, and furthermore, they are independent. As such, the probability that we visit the origin at the \(2i\)th step is 
\[
\mathbb{P}[x'_i = 0 \cap y'_i = 0] = \mathbb{P}[x'_i = 0]^2 = \left(\frac{1}{2^{\frac{2i}{2}}}\right)^2 \geq \frac{1}{4i}. 
\]
We conclude, that the infinite random walk on the grid \(\mathbb{Z}^2\) visits the origin in expectation 
\[
\sum_{i=0}^{\infty} \mathbb{P}[x'_i = 0 \cap y'_i = 0] \geq \sum_{i=0}^{\infty} \frac{1}{4i} = \infty,
\]
as this sequence diverges.

37.1.1.3. Walking on three dimensional grid

In the following, let \(\left(\begin{array}{c} i \\ a \ b \ c \end{array}\right) = \frac{i!}{a! b! c!} \). 

Lemma 37.1.5. Consider the infinite random walk on the three dimensional integer grid \( \mathbb{Z}^3 \), starting from \((0,0,0)\). The expected number of times that such a walk visits the origin is bounded.

Proof: The probability of a neighbor of a point \((x,y,z)\) to be the next point in the walk is \(1/6\). Assume that we performed a walk for \(2i\) steps, and decided to perform \(2a\) steps parallel to the \(x\)-axis, \(2b\) steps parallel to the \(y\)-axis, and \(2c\) steps parallel to the \(z\)-axis, where \(a + b + c = i\). Furthermore, the walk on each dimension is balanced, that is we perform \(a\) steps to the left on the \(x\)-axis, and \(a\) steps to the right on the \(x\)-axis. Clearly, this corresponds to the only walks in \(2i\) steps that arrives to the origin.

Next, the number of different ways we can perform such a walk is \(\frac{(2i)!}{a! b! c!} \), and the probability to perform such a walk, summing over all possible values of \(a, b\) and \(c\), is 
\[
\alpha_i = \sum_{a+b+c=i \atop a,b,c \geq 0} \frac{(2i)!}{a! b! c!} \frac{1}{6^{2i}} \frac{1}{2^{2i}} \sum_{a+b+c=i \atop a,b,c \geq 0} \left(\frac{i!}{a! b! c!}\right)^2 \left(\frac{1}{3}\right)^{2i} = \left(\frac{i}{2i}\right) \frac{1}{2^{2i}} \sum_{a+b+c=i \atop a,b,c \geq 0} \left(\binom{i}{a b c}\right) \left(\frac{1}{3}\right)^{2i}. 
\]
Consider the case where \(i = 3m\). We have that \(\binom{i}{a b c} \leq \binom{m}{m m m}\). As such, 
\[
\alpha_i \leq \left(\frac{2i}{i}\right) \frac{1}{2^{2i}} \left(\frac{1}{3}\right)^{i} \sum_{a+b+c=i \atop a,b,c \geq 0} \left(\binom{i}{a b c}\right) \left(\frac{1}{3}\right)^{i} = \left(\frac{2i}{i}\right) \frac{1}{2^{2i}} \left(\frac{1}{3}\right)^{i} \binom{m}{m m m}. 
\]
By the Stirling formula, we have
\[
\binom{m}{i m m} \approx \frac{\sqrt{2\pi i} (i/e)^i}{(\sqrt{2\pi i/3 (i/3)^3})^3} = \frac{c^i}{i},
\]
for some constant \(c\). As such, \(\alpha_i = O\left(\frac{1}{\sqrt{i \cdot \frac{1}{3}}} \cdot \frac{i^{3/2}}{i}\right) = O\left(\frac{1}{i^{3/2}}\right)\). Thus,
\[
\sum_{m=1}^{\infty} \alpha_{6m} = \sum_{i} O\left(\frac{1}{i^{3/2}}\right) = O(1).
\]
Finally, observe that \(\alpha_{6m} \geq (1/6)^2 \alpha_{6m-2}\) and \(\alpha_{6m} \geq (1/6)^4 \alpha_{6m-4}\). Thus,
\[
\sum_{m=1}^{\infty} \alpha_m = O(1).
\]

37.2. Bibliographical notes

The presentation here follows [Nor98].

References