Monkeys and Walks

Muhammad Waliji

August 12, 2006

1 Preliminaries

We will be dealing with outcomes resulting from random processes. An outcome of the process will sometimes be denoted ω. Note that we will identify the set A with the event that the random outcome ω belongs to A. Hence, A will mean 'A occurs', and $A \cup B$ means 'either A or B occurs', whereas $A \cap B$ means 'both A and B occur.'

For a sequence of events, $A_1, A_2, \ldots$, note that

$$\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

means that infinitely many of the A_n occur, or that A_n occurs infinitely often (A_n i.o.).

The events $A_1, \ldots, A_n$ are said to be independent if $P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$. An infinite collection of events is said to be independent if every finite subcollection is independent.

2 The Infinite-Monkey Theorem

Before we state the infinite-monkey theorem, we will prove a useful lemma. Consider the events $A_1, A_2, \ldots$.

Lemma 2.1 (Borel-Cantelli)**.** *If $\sum_n P(A_n) < \infty$, then $P(A_n$ i.o.$) = 0$. Furthermore, if the events A_n are independent, then if $\sum_n P(A_n) = \infty$, we have $P(A_n$ i.o.$) = 1$.*

Proof. Suppose that

$$\sum_{n=1}^{\infty} P(A_n) < \infty. \tag{2.1}$$

Then,

$$P\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n\right) = \lim_{m\to\infty} P\left(\bigcup_{n=m}^{\infty} A_n\right) \le \lim_{m\to\infty}\sum_{n=m}^{\infty} P(A_n) \tag{2.2}$$

and hence by (2.1), the limit in (2.2) must be 0.

For the converse, it is enough to show that

$$P\left(\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} A_n^c\right) = 0$$

and so it is also enough to show that

$$P\left(\bigcap_{n=m}^{\infty} A_n^c\right) = 0 \tag{2.3}$$

Note that since $1 - x \le e^{-x}$, and by independence,

$$\begin{aligned} P\left(\bigcap_{n=m}^{\infty} A_n^c\right) &\le P\left(\bigcap_{n=m}^{m+k} A_n^c\right) \\ &= \prod_{n=m}^{m+k}(1 - P(A_n)) \\ &\le \exp\left(-\sum_{n=m}^{m+k} P(A_n)\right). \end{aligned} \tag{2.4}$$

But since, taking $k \to \infty$, the last sum in (2.4) diverges, this establsihes (2.3), thereby completing the proof. □

Loosely speaking, the infinite-monkey theorem states that if a monkey hits keys on a typewriter at random for an infinite amount of time, he will almost surely produce the entire collected works of Shakespeare! Even better, he will almost surely do this infinitely often! Here is the precise statement:

Theorem 2.2 (Infinite Monkey)**.** *Consider an infinite-length string produced from a finite alphabet by picking each letter independently at random, uniformly from the alphabet (say the alphabet has n letters). Fix a string S of length m from the same alphabet. Let E_k be the event 'the m-substring starting at position k is the string S. Then, infinitely many of the E_k occur with probability 1.*

Proof. Note that the events E_{mj+1} are independent for $j = 0, 1, \ldots$. Furthermore, $P(E_k) = (\frac{1}{n})^m$. Hence,

$$\sum_{j=1}^{\infty} P(E_{mj+1}) = \sum_{j=1}^{\infty}\left(\frac{1}{n}\right)^m = \infty,$$

so by Lemma 2.1, $P(E_{mj+1}$ i.o.$) = 1$. □

3 Random Walks in $\mathbb{Z}^d$

Consider a Random Walk (a drunkard's walk) in the d dimensional lattice $\mathbb{Z}^d$. Suppose the that drunkard starts out at the origin, and at each step he moves to any adjacent point with equal probability. The question is, what is the chance that he will return to the origin? Actually, the question that we will consider is, what is the chance that he will return to the origin infinitely often?

Let us fix some notation. Let X_n be the position of the drunkard after n steps ($n = 0, 1, \ldots$). Let $P_i(\cdot)$ denote the probability that an event occurs for a random walk starting a position i. Let $p_{ij}^{(n)}$ denote the probability that after starting at position i, the walker is at position j after n steps. Let $f_{ij}^{(n)}$ denote the probability that after starting at position i, the walker reaches position j for the first time after n stepts. Let f_{ij} denote the probability that the walker eventually reaches j after starting out at i. So,

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}.$$

Definition 3.1. A position i is *recurrent* if $f_{ii} = 1$. It is *transient* if $f_{ii} < 1$.

Lemma 3.2. *Suppose a random walker starts out at position i. Then,*

$$P_i(X_n = i \ i.o.) = \begin{cases} 1 & \text{if } i \text{ is recurrent} \\ 0 & \text{if } i \text{ is transitive} \end{cases} \tag{3.1}$$

Proof. Let $1 \leq n_1 < \cdots < n_k$. Let $A_{n_1,\ldots,n_k}^{ij}$ be the event that $X_{n_1} = \cdots = X_{n_k} = j$ and $X_t \neq j$ for the other $t < n_k$. Then,

$$P_i(A_{n_1,\ldots,n_k}^{ij}) = f_{ij}^{(n_1)} f_{jj}^{(n_2-n_1)} \cdots f_{jj}^{(n_k-n_{k-1})}$$

Now,

$$A_k := \bigcup_{n_1,\ldots,n_k} A_{n_1,\ldots,n_k}^{ij}$$

is the event that $X_t = j$ at least k times. Also,

$$\begin{aligned}
P_i(A_k) &= \sum_{1 \leq n_1 < \cdots < n_k} P(A_{n_1,\ldots,n_k}^{ij}) \\
&= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=n_{k-1}+1}^{\infty} f_{ij}^{(n_1)} f_{jj}^{(n_2-n_1)} \cdots f_{jj}^{(n_k-n_{k-1})} \\
&= \sum_{n_1=1}^{\infty} f_{ij}^{(n_1)} \sum_{n_2=n_1+1}^{\infty} f_{jj}^{(n_2-n_1)} \cdots \sum_{n_k=n_{k-1}+1}^{\infty} f_{jj}^{(n_k-n_{k-1})} \\
&= \sum_{n_1=1}^{\infty} f_{ij}^{(n_1)} \sum_{n_2=1}^{\infty} f_{jj}^{(n_2)} \cdots \sum_{n_k=1}^{\infty} f_{jj}^{(n_k)} \\
&= f_{ij}(f_{jj})^{k-1} \qquad (3.2)
\end{aligned}$$

Taking the limit as $k \to \infty$ in (3.2), we get

$$\begin{aligned} P_i(X_n = j \text{ i.o.}) &= P_i\left(\bigcap_{k=1}^{\infty} A_k\right) \\ &= \lim_{k\to\infty} P_i(A_k) \\ &= \begin{cases} f_{ij} & \text{if } f_{jj} = 1 \\ 0 & \text{if } f_{jj} < 1 \end{cases} \end{aligned} \tag{3.3}$$

Replacing j with i in (3.3), we obtain (3.1). □

Before we answer the question posed at the beginning of the section, we would like to have a Borel-Cantelli like lemma for the random walk. However, the X_n are manifestly not independent. Still, the lemma will hold.

Lemma 3.3. *The following hold:*

(i) $P_i(X_n = i \text{ i.o.}) = 0 \Longleftrightarrow \sum_n p_{ii}^{(n)} < \infty$.

(ii) $P_i(X_n = i \text{ i.o.}) = 1 \Longleftrightarrow \sum_n p_{ii}^{(n)} = \infty$.

Proof. Note that under (3.1), (i) and (ii) are equivalent. We will prove (i). Note that $\Longleftarrow$ follows from Lemma 2.1. We will here prove $\Longrightarrow$.

We can condition $p_{ij}^{(n)}$ on the first t st $X_t = j$:

$$p_{ij}^{(n)} = \sum_{l=0}^{n-1} f_{ij}^{(n-l)} p_{jj}^{(l)}.$$

Hence,

$$\begin{aligned} \sum_{k=1}^{n} p_{ii}^{(k)} &= \sum_{k=1}^{n} \sum_{l=0}^{k-1} f_{ii}^{(k-l)} p_{ii}^{(l)} \\ &= \sum_{l=0}^{n-1} p_{ii}^{(l)} \sum_{k=l+1}^{n} f_{ii}^{(k-l)} \\ &\leq \sum_{l=0}^{n} p_{ii}^{(l)} f_{ii} \\ &= \left(1 + \sum_{l=1}^{n} p_{ii}^{(l)}\right) f_{ii} \end{aligned}$$

and therefore,

$$\left(\sum_{k=1}^{n} p_{ii}^{(k)}\right)(1 - f_{ii}) \leq f_{ii} \tag{3.4}$$

Now, since $f_{ii} < 1$, (3.4) gives an upper bound for the partial sums of

$$\sum_{n=1}^{\infty} p_{ii}^{(n)}$$

thereby completing the proof. $\square$

The question that we posed at the beginning of the section was answered by Polya in 1921.

Theorem 3.4 (Polya)**.** *If $d = 1, 2$, then the probability that the walker returns to the origin infinitly often is 1. For $d \geq 3$ the probability is 0.*

This is proved by calculating the sum

$$\sum_{n=1}^{\infty} p_{00}^{(n)}$$

and applying Lemma 3.3. We will give a proof for the case $d = 1, 2$. First, recall Stirling's formula:

$$n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n}$$

Therefore,

$$\binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}}.$$

Also, recall

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k} \binom{n}{n-k}$$

Proof when $d = 1$. Note that $p_{00}^{(m)} = 0$ if m is odd. If $m = 2n$, then every particular walk of length $2n$ has probability $(1/2)^{2n}$. If n of the steps are to the right, and n are to the left, the walker will have returned to the origin. There are $\binom{2n}{n}$ ways of doing this. Therefore, $p_{00}^{(2n)} = \binom{2n}{n}(1/2)^{2n}$. Summing over all n, we have

$$\sum_{n=1}^{\infty} p_{00}^{(2n)} = \sum_{n=1}^{\infty} \binom{2n}{n} (1/2)^{2n} \sim \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}}$$

which diverges, and so we are done by Lemma 3.3. $\square$

Proof when $d = 2$. Now, the probability of any particular walk of length n is $(1/4)^n$. In order that we return to the origin, there must be the same number of moves left as there are right, and there must be the same number of moves down as there are up. Again, we have $p_{00}^{(m)} = 0$ for odd m. If $m = 2n$, there are

$$\sum_{k=0}^{n} \binom{2n}{k \quad k \quad n-k \quad n-k}$$

ways of making the same number of moves left and right and the same number of moves up and down. Hence, we have

$$\begin{aligned}
p_{00}^{(2n)} &= \left(\frac{1}{4}\right)^{2n} \sum_{k=0}^{n} \binom{2n}{k \quad k \quad n-k \quad n-k} \\
&= \left(\frac{1}{4}\right)^{2n} \sum_{k=0}^{n} \frac{(2n)!}{k!k!(n-k)!(n-k)!} \\
&= \left(\frac{1}{4}\right)^{2n} \sum_{k=0}^{n} \frac{(2n)!}{n!n!} \frac{n!n!}{k!k!(n-k)!(n-k)!} \\
&= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \sum_{k=0}^{n} \binom{n}{k} \binom{n}{n-k} \\
&= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n}^2 \sim \frac{1}{\pi n}
\end{aligned}$$

Summing over n, we find that $\sum_n p_{00}^{(2n)}$ diverges, and so we are done by Lemma 3.3. $\square$

For $d \geq 3$, we will find that $p_{00}^{(2n)} = O\left(n^{-d/2}\right)$ by some sort of inductive proof.

4 Further Study

The following problem was proposed to me by Dr Laszlo Babai. Consider the random walk in $\mathbb{Z}$ starting at the origin, where at each step the walker moves either to the right or to the left, each with probability p, or he stays put with probability $1-2p$. It can easily be shown that the probabability of returning to the origin in n steps is $O(n^{-1/2})$, as above. Now, consider a 'phased' random walk on $\mathbb{Z}^d$, where at step n, we move in the direction of dimension n mod d, with the same probabilities as before. Then, it is trivial to prove that the probability of return to the origin in n steps is $O(n^{-d/2})$. Is it possible to find an alternate proof to Polya's Theorem by approximating the Random Walk on $\mathbb{Z}^d$ by the Phased Random Walk on $\mathbb{Z}^d$?