Infinitely Long Walks on 2-colored Graphs Which Don't Cover the Graph

Pete Gemmell*

December 14, 1992

Abstract

Suppose we have a undirected graph $G = (V, E)$ where V is the set of vertices and E is the set of edges. Suppose E consists of red colored edges and blue colored edges. Suppose we have an infinite sequence S of characters R and B.

We take a random walk starting at vertex v on G based on the sequence S as follows:

At the ith step, if S has an R at position i the walk traverses a random red edge out of the current vertex (chosen uniformly from the outgoing edges). If S has a B the walk traverses a random blue edge out of the current vertex.

We say S covers G starting at vertex v when a random walk using S starting at v covers every vertex of G.

Theorem 1 *If G is a red-blue colored undirected graph which is connected in red and connected in blue and there exists an RB-sequence S such that starting at some vertex v,*

$$Pr[S \text{ covers } G] < 1$$

then G contains a proper subgraph H such that H's vertices can be divided into two sets: U and W where there are no red edges between U and $V - W$ and no blue edges between U and $V - U$.

1 Notation

In this paper, we consider random walks on graphs with undirected edges. The edges are always one of two colors – red or blue. Furthermore, we will always consider graphs which are connected, both by blue edges and by red edges.

Definition 1 *Let S be a fixed infinite sequence of symbols "R" and "B".*

Let G be a two colored graph as described above.

Let v be a vertex of G.

A random walk, (W, S) on G starting at v is an infinite sequence of vertices $\{v_i\}_{i=1}^{\infty}$ of G. If the ith entry of S is a "B", then at the ith step of the walk traverses an edge chosen uniformly at random

*Computer Science Division, UC Berkeley, Berkeley, CA 94720., supported by NSF grant number CCR-9201092

from those blue edges adjacent to v_{i-1}. v_i is the vertex on the other end of this edge. If the ith entry of S is an "R", then at the ith step of the walk traverses an edge chosen uniformly at random from those red edges adjacent to v_{i-1}.

In this report, we will refer to finite continuous subsets of the infinite R and B sequences as *blocks* or *substrings*. Also, by $(RB)^k$ we mean a block of the form $RBRB\ldots$ repeated k times.

2 The Proof

Lemma 2 *If $\exists$ sequence S, vertex $v \in V$ such that $Pr[S$ covers G starting at $v] < 1$ then there is some start vertex $v' \in V$ such that for the sequence $T = (RB)^\infty$ we have :*

$$Pr[T \text{ covers } G \text{ starting at } v'] = 0$$

Proof.

The structure of the proof is as follows:

1. We divide the sequence S up into an infinite number of blocks, $b_i, i = 1, \ldots\infty$ of length L where $S = b_1 b_2 \ldots$ and $L >> n$.
2. We argue that, with probability 1, for a random walk (W, S), for an infinite number of blocks, $b_{i_j}, j = 1, \ldots\infty$, adjacent R's (and adjacent B's) will effectively cancel each other in a way that allows the walk to be simulated by a walk (W',S') where S' has long substrings of the form $(RB)^M$.
3. Because, with probability 1, these substrings $(RB)^M$ will occur an infinite number of times for a random walk W and because there is no upper bound on M, we know that there exists a vertex $v' \in V$ such that $\forall M, Pr_W[(RB)^M$ covers G starting at $v'] = 0$ and hence $Pr_W[(RB)^\infty$ covers G starting at $v'] = 0$.

We will use a transformation of strings based on cancelling adjacent R's and adjacent B's.

Definition 2 *Let b be a block of R's and B's.*

We say b' is a **transform** *of b if one can obtain b' by recursively cancelling adjacent R's and B's in b.*

For example, if $b = RBBRBRRBB$ then both B and RRB are tranforms of b.

Definition 3 *For sequence $S = b_1 b_2 \ldots$, we say sequence $S' = b'_1 b'_2 \ldots$ is a* **transform** *of S if, for all i, b'_i is a (possibly trivial) transform of b_i.*

Definition 4 *Let sequence S' be a transform of sequence S.*

Let (W, S) be a walk on graph G. Suppose that, at every place in S where two R's or two B's can be compressed in transforming S to S', the walk W backtracks, first transversing an edge in one direction and then traversing it in the other.

Let W' be the walk on G using S' such that $W' = W$ except that all such backtracks are deleted.

We say (W', S') is a **transform** *of (W, S).*

Note that the set of vertices covered by a transformation of a walk (W, S) is a subset of the set of vertices covered by the walk (W, S).

Claim 3 *Let $S = b_1 b_2 \ldots$ where each block has L symbols.*

Let (W, S) be a random walk on G starting at vertex v.

For each i, let b'_i be an arbitrary transformation of b_i.

Then, with probability 1 over all random walks (W, S) on G, there exists sequence S' and a walk (W', S') such that:

1. *(W', S') is a transformation of (W, S)*
2. *$S' = c_1 c_2 \ldots$ where, for an infinite number of values of i, $c_i = b'_i$ and for the rest, $c_i = b_i$.*
3. *The probability distribution on the random walk (W', S') is identical to the uniform distribution on random walks using S'.*

That (W', S') is uniformly distributed is important because we will use the above claim to show that uniformly distributed random walks using sequences with long $(RB)^M$ substrings are unlikely to cover the graph.

Proof.

Let (W, S) be a random walk on G using S.

For each $i > 0$, consider one way of cancelling R's and B's to transform b_i to b'_i.

Let x_i denote the number of cancellations needed for this transformation.

If every backtrack implied by this transformation occured in the random walk (W, S), then we could conceivably include the block b'_i in the transformed string S'.

However, in order to achieve a uniform distribution on the walk (W', S'), we include the block b'_i in the transformed string S' with some probability possibly less than 1. We normalize the probability of including b'_i in S' so that the event that b'_i is included in S' yields no information about the actual path that (W', S') follows while using block b'_i of S'.

For the jth backtrack needed by the transformation, let d_i^j be the degree of the vertex from which the walk (W, S) backtracks.

If all the backtracks required by the transformation occur, then we include the block b'_i in the sequence S' with probability: $\Pi_{j=1}^{x_i} \frac{d_i^j}{n}$.

Therefore, the probability (over random walks (W, S)) that b'_i will be included in S' is:

$$(\Pi_{j=1}^{x_i} \frac{1}{d_i^j})(\Pi_{j=1}^{x_i} \frac{d_i^j}{n}) = \frac{1}{n}^{x_i} \geq \frac{1}{n}^{\frac{L}{2}}$$

Because there are an infinite number of blocks b_i, with probability 1, we can transform walk (W, S) into a walk (W', S') which satisfies conditions (1) and (2) of the claim.

The transformations of the blocks b_i to b_i' yield no information about the walk (W', S') and therefore condition (3) is satisfied.

■

Claim 4 *Let $S = b_1 b_2 \dots$ be any sequence of R's and B's such that the length of each block b_i is L.*

Then for all i, $\exists b_i'$ such that b_i' is a transform of b_i and where b_i' contains a substring of the form $(RB)^M$ or of the form R^{n^4} or of the form B^{n^4} where M is any number such that $L \geq (4M^2n^4)^{n^4}$.

Proof.

The transformation of each block goes as follows:

1. We can use cancellation to reduce each block b_i to a string of $R's$ and $B's$ such that no more than two $R's$ or $B's$ occur in a row. If the block has no consecutive string of more than n^4 R's or n^4 B's, then this will not reduce the length of the block more than a factor of n^4.
2. Cancel all adjacent $R's$. This reduces the string by at most a factor of 3.
3. The string now looks like this:

$$(RB)^a B (RB)^b B (RB)^c B (RB)^d B (RB)^e B (RB)^f B (RB)^g B (RB)^h B ...$$

   and is of length $L_0 \geq \frac{L}{3n^4}$.

   We now repeat the following steps.

   (a) In iteration j, we start with a string of the form:

$$(RB)^a B (RB)^b B^{2j+1} (RB)^c B (RB)^d B^{2j+1} (RB)^e B (RB)^f B^{2j+1} (RB)^g B (RB)^h B^{2j+1} ...$$

   and with length at least $\frac{L}{(4n^4)^j}$.

   (b) We cancel adjacent groupings to get:

$$[(RB)^{a-b} B \text{ or } B(RB)^{b-a}] B^{2j+1} [(RB)^{c-d} B \text{ or } B(RB)^{d-c}] B^{2j+1} [(RB)^{e-f} B \text{ or } B(RB)^{f-e}] \dots$$

   We call terms of the form $(RB)^{a-b}B$ **type 1** and we call terms of the form $B(RB)^{b-a}$ **type 2**. If $a = b$, then we will call the term $B = (RB)^{a-b}B = B(RB)^{b-a}$ **type 3**.

   (c) We eliminate all terms of type 3 by cancelling them with an adjacent B^{2j+1} grouping. We note that this can decrease the length of the string by at most a factor of n^4 unless we have more than n^4 adjacent B symbols in a row.

(d) If we have two consecutive terms of type 1 $(RB)^{a-b}BB^{2j+1}(RB)^{c-d}B$ then we cancel the middle $B's$ to get a longer term of type 1: $(RB)^{a+c-(b+d)}B$ If we have more than M consecutive terms which are all of type 1, then we're done because this yields a string of the form $(RB)^M$.

(e) We do the same for two consecutive terms of type 2. If we have more than M consecutive terms which are all of type 2, then we're done because this yields a string of the form $(RB)^M$. This step, combined with the previous step, decreases the length of the string by at most a factor of M

(f) Our string now has many alternations between type 1 and type 2.
If we have a term of type 1 followed by a term of type 2 then we have: $(RB)^{a-b}BB^{2j+1}B(RB)^{d-c} = (RB)^{a'}B^{2j+3}(RB)^{c'}$.
If we have a term of type 2 followed by a term of type 1 then we have: $B(RB)^{b-a}B^{2j+1}(RB)^{c-d}B = B(RB)^{a'}B^{2j+1}(RB)^{c'}B$.
In any event (getting rid of the '), our new string looks like this:

$$...(RB)^aB^{2j+1}(RB)^bB^{2j+3}(RB)^cB^{2j+1}(RB)^dB^{2j+3}(RB)^eB^{2j+1}(RB)^fB^{2j+3}\dots$$

which we can reduce to:

$$...(RB)^aB(RB)^bB^{2j+3}(RB)^cB(RB)^dB^{2j+3}(RB)^eB(RB)^fB^{2j+3}\dots$$

This new string has length $L_{j+1} \geq \frac{L_j}{4M^2n^4}$.

Each time we are able to repeat the process, the intervening string of B's becomes longer. Because we have assumed that $L \geq (4M^2n^4)^{n^4}$ we can repeat the process for at least $\frac{j}{2}$ steps. If we repeat the process for $\frac{j}{2}$ steps, we would have the n^4 B's in a row.

Therefore, we can transform every block into a string which contains B^{n^4} or R^{n^4} or $(RB)^M$ as a substring.

■

We can now complete the proof of Lemma 2.

Let $S = b_1b_2\dots$ be a sequence such that $Pr_W[S$ covers G starting at $v] < 1$.

Let $b'_1b'_2\dots$ be such that each b'_i is a transform of b_i and contains a substring of the form B^{n^4} or R^{n^4} or $(RB)^M$.

Let $S' = c_1c_2\dots$ be the sequence and (W', S') be the walk guaranteed to exist by Claim 3 such that $c_i = b'_i$ for an infinite number of values of i and $c_i = b_i$ for the rest.

By Claim 3, we have $Pr_W[S$ covers G starting at $v] < 1$.

Because S' contains an infinite number of substrings of the form B^{n^4} or R^{n^4} or $(RB)^M$ and because an infinite number of B^{n^4} or R^{n^4} substrings would implied that S' would cover G with probability 1, we know that S' has to contain an infinite number of substrings of the form $(RB)^M$.

Because there is no upper bound on L, there is no upper bound on M and therefore we know that there exists a vertex $v' \in V$ such that $\forall M, Pr_W[(RB)^M$ covers G starting at $v'] = 0$ and hence $Pr_W[T = (RB)^\infty$ covers G starting at $v'] = 0$.

■

Now we show that G contains a subgraph of the appropriate form.

We will be considering be a graph G which satisfies the conditions of the main theorem and for which we have proven: $\exists$ vertex v' of G such that

$$Pr_W[T = (RB)^\infty \text{ covers } G \text{ starting at } v'] = 0$$

We need the following definitions and claim:

Definition 5 *Let (W, T) be a random walk on G starting at vertex v.*

Then W_i denotes the ith edge which traversed by the walk (W, T).

$W[1, i]$ denotes the first i edges of the walk.

Definition 6 *Let H be a subgraph of G and let $W[1, i]$ be a path of i edges in G starting at v'.*

$P_i(H, W[1, i]) = Pr[$ we will ever leave H after i steps $|$ the first steps of the walk have been $W_1 \dots W_i]$.

Claim 5 *If G satifies the conditions of the theorem, then there exists an path $W[1, i]$ starting at v', a subgraph H of G, and a positive integer I_0 such that*

1. $\forall i \geq I_0, P_i(H, W[1, i]) = 0$
2. $\forall H'$ *proper subgraph of* $H, \forall i, P_i(H', W[1, i]) > 0$.

Proof.

(1): Because $Pr[S'$ covers G starting at $v] < 1$, we know that $\exists H, \forall W, P_0(H, W[1, 0]) < 1$.

Note that $P_i(H, W)$ can have at most $2n$ different values, depending only on what vertex, v_i, we are visiting and whether S'_i is an R or a B.

We will construct W as follows:

Let $u_1 \dots u_k$ be the neighbors of v_i joined to v_i by an edge of color S'_i. We will leave v_i via one of these edges and possibly exit H in the process. So $P_i(H, W[1, i])$ is less than or equal to the weighted average of $P_{i+1}(H, W[1, i]\{v_i, u_j\})$ for $j \in [1, k]$.

Choose $W_{i+1} = \{v_i, u_j\}$ such that $P_{i+1}(H, W[1, i]\{v_i, u_j\})$ is minimized. If any of the u'_js are in $\overline{H}$, the inequality will be strict. If not, we are guaranteed that the quantity $P_{i+1}(H, W[1, i + 1])$ will not increase. Because it can adopt only a finite number of values, it must reach 0 eventually.

(2): Let H be a minimal subgraph such that (1) holds.

Note that if $P_i(H, W[1, i]) = 0$ then $\forall m > i, P_m(H, W[1, m]) = 0$. ■

Now we classify the vertices of H and the edges between adjacent to them.

Let $afterRed = \{v \in H | Pr[$ we reach v after a blue edge at any step $m > I_0 | W[1, I_0]] = 0\}$.

Let $afterBlue = \{v \in H | Pr[$ we reach v after a red edge at any step $m > I_0 | W[1, I_0]] = 0\}$.

Let $afterBoth = H - (afterRed \cup afterBlue)$

Claim 6 *1. There are no blue edges between between $afterRed$ and $\overline{H}$. and there are no red edges between between $afterBlue$ and $\overline{H}$.*

2. There are no blue edges internal to $afterRed$ and no red edges internal to $afterBlue$.

3. There are no blue edges between $afterRed$ and $afterBoth$ and no red edges between $afterBlue$ and $afterBoth$.

4. There are no edges between $afterBoth$ and $\overline{H}$. And therefore both $afterRed$ and $afterBlue$ are non-empty.

Note this claim implies of Theorem 1 where $H_0 = afterRed \cup afterBlue$.

Proof.

(1) If there is a blue edge between between $v \in afterRed$ and $u \in \overline{H}$ then $Pr[$ we reach v after step $I_0] > 0 \Rightarrow Pr[$ we escape from H after step $I_0 | W[1, I_0]] > 0$ which is a contradiction.

(2) If there is a blue edge $\{v, u\}$ internal to $afterRed$, then $Pr[$ we reach u after a blue edge at some step $m > I_0 | W[1, I_0]] > 0$ – a contradiction.

(3) If there is a blue edge between $v \in afterRed$ and $u \in afterBoth$, then $Pr[$ we reach v after a blue edge some step $m > I_0 | W[1, I_0]] > 0$.

(4) If there is a blue edge between $v \in afterBoth$ and $u \in \overline{H}$ then $Pr[$ we escape H after step $I_0] > 0$

■