

Uniform Generalized Steinhaus Graphs

Neal Brand¹

Department of Mathematics
University of North Texas
Denton, TX 76203

and

Margaret Morton

Department of Mathematics
University of Auckland
Auckland, New Zealand

Abstract

In [1] it is shown that the first order theory of almost all generalized Steinhaus graphs is identical to the first order theory of almost all graphs where each generalized Steinhaus graph is given the same probability. A natural probability measure on generalized Steinhaus graphs is obtained by independently assigning a probability of p for each entry in the generating string of the graph. With this probability measure it is shown that the first order theory of almost all uniform generalized Steinhaus graphs is identical to the first order theory of almost all graphs.

1. Introduction.

The concept of a generalized Steinhaus graph was introduced in [1]. In this paper we consider uniform generalized Steinhaus graphs, these form a subset of the set of generalized Steinhaus graphs which includes the usual Steinhaus graphs. The definition follows the usual pattern of first defining a uniform generalized Steinhaus triangle and then using this to build the adjacency matrix of a uniform generalized Steinhaus graph.

We define a **uniform generalized Steinhaus triangle** of order n and type s to be the upper triangular portion of an n by n binary array $A = (a_{i,j})$ whose entries satisfy $a_{i,j} = \sum_{r=0}^{s-1} c_r a_{i-1,j-r} \pmod{2}$ where $2 \leq i \leq n-1$, $i+s-1 \leq j \leq n$, $c_r \in \{0,1\}$ and $c_{s-1} = 1$. Note that other than the condition $c_{s-1} = 1$ there are no conditions on the values of c_r . We will assume there is a fixed (but arbitrary) choice of values for the c_r . With this fixed choice we investigate the resulting collection of all uniform generalized Steinhaus graphs.

The associated uniform generalized Steinhaus graph is the labeled graph whose adjacency matrix is obtained from the uniform generalized Steinhaus triangle by making A symmetric with a zero main diagonal. We will represent the vertex set of a uniform generalized Steinhaus triangle and graph by $V_n = \{1, 2, \dots, n\}$. The generating string of the uniform generalized triangle and graph consists of $(a_{1,j})_{j=2}^n$

¹ Supported by a University of North Texas Faculty Research Grant.

plus the first $s - 2$ entries in rows 2 through n if $s > 2$. If there are fewer than $s - 2$ entries in a row then each entry in the row becomes part of the generating string. For convenience we will refer to the entries $(a_{1,j})_{j=2}^n$ in the generating string as the **top generators** and the remaining generators as the **diagonal generators**.

We define a probability measure on uniform generalized Stenhaus graphs of order n and type s by requiring that $Pr(g_r = 1) = p_{n,r}$ where $0 \leq p_{n,r} \leq 1$ and g_r is any element in the generating string. We then define $q_{n,r} = 1 - p_{n,r}$ and $m_{n,r} = \min(p_{n,r}, q_{n,r})$. Any function $f(n)$ with the properties that for each sufficiently large n , $0 < f(n) < 1$ and $m_{n,r} \geq f(n)$ for each g_r in the generating string is called a probability bound. We say that almost all uniform generalized Steinhaus graphs have a certain property if the probability that a uniform generalized Steinhaus graph has that property approaches one as n (the number of vertices) approaches infinity. Naturally the concept of almost all depends on the probability bound. The simplest case is the constant probability bound of $1/2$, this gives all generalized Steinhaus graphs the same probability and is discussed in [1]. The case for Steinhaus graphs, which corresponds to taking $s = 2$ and $c_0 = c_1 = 1$, was considered in [2].

Blass and Harary [3] and Fagin [4] showed that any first order property of graphs is either satisfied by almost all graphs or else almost all graphs do not satisfy the property. Furthermore, axioms are given, each of which almost all graphs satisfy, with the property that for any first order property of graphs either the property or its negation can be deduced from a finite number of axioms. If A represents the adjacency relation for a graph then the first order axiom scheme of [3] is

Axiom k : For any pair of disjoint sets of vertices $\{v_1, v_2, \dots, v_k\}$ and $\{w_1, w_2, \dots, w_k\}$ each with k elements, there is a vertex x with xAv_i and $\neg xAw_i$ for each i .

In [2] it is demonstrated that for each axiom almost all Steinhaus graphs (with the less restrictive probability measure) satisfy the axiom. A similar result is shown in [1] for generalized Stenhaus graphs with equal probability measure. We now extend these results to show that for each axiom almost all uniform generalized Stenhaus graphs satisfy the axiom.

Suppose we restrict our attention to uniform generalized Steinhaus graphs with fixed diagonal generator values. We give this set of graphs the conditional probability measure, conditioned on the choice of values for the diagonal generators. Equivalently, we fix the diagonal generators and define the probabilities for 1's and 0's on the first row as described above. Let $\epsilon > 0$ and k and s be positive integers. Let

$$0 < \gamma < \frac{\epsilon}{2k(2k+1)\log s} (\log \log s - \log s - \log \epsilon).$$

Note that the expression on the right is positive for ϵ sufficiently small. (Here and throughout this paper \log means \log_2 .)

Theorem 1.1 : Let k and s be positive integers and $\epsilon < \frac{\log s}{s}$. Also, let γ be as above. For any fixed choice of diagonal generators and any choice of probability measure using the probability bound $n^{-\gamma}$ almost all uniform generalized Steinhaus graphs satisfy Axiom k . Furthermore, there is a function $g(n)$ such that $g(n) \rightarrow 0$

Proof: In Y there is a one in the entry with position (r, c) if and only if the number of allowed paths from position $(0, 0)$ to position (r, c) is odd. An allowed path progresses down the table one row at a time so that with each step there is a j with $0 \leq j \leq s - 1$ and $c_j = 1$ so that one steps to the right j steps. This is equivalent to counting (mod 2) the number of strings of r numbers from the set $S = \{j | c_j = 1\}$ of step sizes so that the entries in the string sum to c .

The idea is to cancel out in pairs strings of length r having the same sum. The number of strings left after cancelling gives an upper bound on the number of entries in row r which are 1.

If the first two entries of a string are different then by reversing those entries we get another string with the same length r and the same sum. Therefore we can cancel these two strings. Thus we need consider only the strings whose first two entries are identical. Similarly we can pair up the third and fourth entries, fifth and sixth entries, and so on and consider only strings where the paired up entries are equal. Note that if r is odd then there is an entry at the end which is not paired with any other entry. In this case we say that there is a block of size 1.

Using the same argument as above we can cancel out in pairs strings where the common value of the first two entries are different from the common value of the third and fourth entries. So we only need consider strings where the first four entries are identical. Similarly we can consider only strings where the entries are identical within blocks of size 4. Note that at the end there will be an unmatched pair when in the base two representation of r there is a 1 in the 2^i 's place. In this case we say there is a block of size 2. Note that in this block of size 2 the entries are equal.

Continuing in this manner we need consider only strings which have a block of size 2^i if and only if in the base two representation of r there is a 1 in the 2^i place. But there are at most s^w such strings. \square

Generalized Pascal triangles are treated within the framework of a much more general theory on cellular automata developed by Willson in [5] and [6]. The particular result we wish to utilize later states that the number of ones in row $2^q j$, q a non-negative integer, equals the number of ones in row j , $j \geq 1$.

Lemma 2.2 : *Let $s > 1$ be a natural number, $0 < \epsilon < \log s$, and b a positive integer. The number of integers m between 1 and 2^b whose base two representation has at most w ones where $s^w \leq 2^{b\epsilon}$ is at least $\frac{K}{\sqrt{b \left(\frac{\epsilon}{\log s}\right)^{\frac{b\epsilon}{\log s}}}}$ for some $K > 0$ independent of b .*

Proof: Consider strings of 0's and 1's of length b . Let $w = bd$ where $d \leq \frac{\epsilon}{\log s} < 1$. There are clearly $\binom{b}{w}$ strings of this sort containing exactly w ones. Using Stirling's

formula we have

$$\begin{aligned}
 \binom{b}{w} &= \frac{b!}{w!(b-w)!} \\
 &\approx \frac{\sqrt{2\pi b} \left(\frac{b}{e}\right)^b}{\sqrt{2\pi w} \left(\frac{w}{e}\right)^w \sqrt{2\pi(b-w)} \left(\frac{b-w}{e}\right)^{b-w}} \\
 &= \frac{b^{b+\frac{1}{2}}}{\sqrt{2\pi w} w^{w+\frac{1}{2}} (b-w)^{(b-w)+\frac{1}{2}}} \\
 &= \frac{1}{\sqrt{2\pi b d} b^{bd+\frac{1}{2}} (1-d)^{b(1-d)+\frac{1}{2}}} \\
 &= \frac{K}{\sqrt{bd} b^{bd} (1-d)^{b(1-d)}} \\
 &> \frac{K}{\sqrt{bd} b^{bd}} \\
 &\geq \frac{K}{\sqrt{b} \left(\frac{\epsilon}{\log s}\right)^{\frac{be}{\log s}}}
 \end{aligned}$$

where $K = \frac{1}{\sqrt{2\pi d(1-d)}}$ is independent of b . (Of course, this value of K is valid for b sufficiently large. In general K may need to be decreased to account for a few small values for b .) □

Let v and w , $v < w$, be vertices in a uniform generalized Steinhaus graph of type s and Y be the generalized Pascal triangle. Define $B(v, w) = \{x \mid \max(1, w - (v-1)(s-1)) \leq x \leq w, y_{v-1, w-x} = 1\}$ and $H(v, w) = w - (v-1)(s-1)$. Observe that $H(v, w)$ is the smallest entry in $B(v, w)$ as long as $w > (v-1)(s-1)$. We now show that $B(v, w)$ is the set of positions in the top generating string which affect the entry in the (v, w) position of a uniform generalized Steinhaus graph of type s .

Note that we have defined $H(v, w)$ for $v < w$. We wish to define $H(v, w)$ to be the same as $H(w, v)$ in the case that $v > w$. In other words, $H(v, w) = \max(v, w) - (\min(v, w) - 1)(s-1)$.

Lemma 2.3 : *Suppose that G and G' are two uniform generalized Steinhaus graphs whose generating strings are identical except in one position in the first row where the entries are different. Then for any pair of vertices (v, w) with $w \geq (v-1)(s-1)$ the adjacency matrices for G and G' differ in position (v, w) if and only if the position where the generating strings for G and G' differ is in $B(v, w)$.*

Proof: Consider the uniform generalized Steinhaus graph H whose generating string consists of all zeros except in the position $(1, t)$, where it is 1. Let Y be the corresponding generalized Pascal triangle. Note that the adjacency matrix for H at position (v, w) is simply the entry $y_{v-1, w-t}$. Since the adjacency matrix of G added to the adjacency matrix of G' (mod 2) gives the adjacency matrix of H the lemma follows. □

3. First Order Properties.

The purpose of this section is to prove Theorem 1.1. First we give a few more definitions and develop some technical lemmas. For T contained in V_n define $B_T(v) = \cup_{t \in T} B(v, t)$ and $H_T(v) = \{H(v, t) | t \in T\}$. By Lemma 2.3 $B_T(v)$ is the set of all entries in the top generating string which when changed will change the entry in the (v, t) position, for some $t \in T$, in the adjacency matrix. We say a sequence v_1, v_2, \dots, v_r is T -independent if for each $1 \leq i \leq r$ and each $t \in T$, $H_T(v_i) \cap B_T(v_j) = \emptyset$ for $j < i$, $|H_T(v_i)| = |T|$ and $H(v_i, t) > 0$.

Lemma 3.1 : *Let G be a uniform generalized Steinhaus graph of order n and type s with fixed diagonal generators. Given $T \subseteq V$ with $|T| = 2k$ and $0 < \epsilon < \frac{\log s}{s}$ there is an integer N and a constant $C > 0$ such that if $n \geq N$ there is a T -independent sequence of length at least $\frac{C}{\sqrt{\log n} \left(\frac{\epsilon s}{\log s}\right)^{\frac{\epsilon \log n}{(2k+1) \log s}}}$. Note that C and N both depend on s, k , and ϵ , but are independent of n .*

Proof: Let t_1, \dots, t_{2k} be the elements of T listed in order and write each one as $t_i = n^{\alpha_i}$. We begin by showing that there must be two consecutive t_i in T with a 'large' number of elements from the top generating string of G between them. Let $\alpha_0 = 0$ and $\alpha_{2k+1} = 1$. Pick $0 \leq i \leq 2k$ so that $\alpha_{i+1} - \alpha_i \geq \frac{1}{2k+1}$ and consequently $\frac{t_{i+1}}{t_i} \geq n^{\frac{1}{2k+1}}$. Certainly there is at least one such i since the α_i 's break the unit interval into $2k+1$ intervals. Now there exist unique integers a', b' such that $2^{a'-1} < t_i \leq 2^{a'}$ and $2^{b'} \leq t_{i+1} < 2^{b'+1}$. Suppose $s < 2^\delta$, let $a = a' + \delta$ and $b = b' - \delta$, then $b - a = b' - a' - 2\delta$. Now $n^{\frac{1}{2k+1}} \leq \frac{t_{i+1}}{t_i} < \frac{2^{b'+1}}{2^{a'-1}} = 4 \left(\frac{2^{b'}}{2^{a'}}\right)$ so taking log base 2 we get $b' - a' \geq -2 + \frac{1}{2k+1} \log n$, that is $b - a \geq \frac{1}{2k+1} \log n - 2 - 2\delta$. We take N sufficiently large so that for $n \geq N$, $b - a > 0$.

Now we want to estimate the number of members x of the top generating string between 2^a and 2^b such that the entries (x, t) in the adjacency matrix for G , $t \in T$, are only affected by a 'small' number of members of the top generating string. As explained in Section 2 this is equivalent to showing that there are a 'large' number of rows between row 2^a and row 2^b in the generalized Pascal triangle of type s which contain a 'small' number of ones. For ease of computation we consider only those x of the form $2^a j$, $j \geq 1$. By Willson's result we can equivalently count the number of x in the top generating string between positions 1 and 2^{b-a} such that row $j-1$ of a generalized Pascal triangle of type s contains no more than s^w ones where $s^w < 2^{(b-a)\epsilon}$. By Lemma 2.2 in the first 2^{b-a} rows of a generalized Pascal triangle of type s there are at least $\frac{K}{\sqrt{b-a} \left(\frac{\epsilon}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$ rows containing no more than s^w ones. Thus there are at least $\frac{K}{\sqrt{b-a} \left(\frac{\epsilon}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$ values of j such that those x of the form $2^a j$ between 2^a and 2^b are candidates for the required T -independent set.

Let W be the set of all row numbers in a generalized Pascal table of type s which contain no more than s^w ones. Let $C(a, b) = \{x | 2^a \leq x \leq 2^b, x = 2^a j \text{ where } j \in W\}$, $x \in C(a, b)$ and $t \in T$. Note that $t \neq x$, so either $t \leq t_i \leq 2^{a'}$ or else $t \geq t_{i+1} \geq 2^{b'}$. For $t < x$, $H(x, t) = x - (t-1)(s-1) \leq x$. Moreover

$\frac{x+t+s}{ts+1} > \frac{2^s}{2^{s-1}} = 1$ so $x+t+s-ts-1 > 0$, hence $0 < H(x,t) \leq x$. For $t > x$, $H(x,t) = t - (x-1)(s-1)$. Since $\frac{t-1}{s} > \frac{2^{b'}-1}{2^{\delta}} = 2^b - 2^{-\delta} > x-1$, $t-1 > s(x-1)$ or $t-s(x-1) + (x-1) > x$. Hence, $H(x,t) = t - (x-1)(s-1) > x$. Combining these two results we see that for each x in $C(a,b)$ we have $|H_T(x)| = 2k = |T|$ and $H(x,t) > 0$ for each $t \in T$.

List the elements in $C(a,b)$ in order and label them h_1, h_2, \dots . We start a T -independent set by setting $x_1 = h_1$. Clearly $\{x_1\}$ is a T -independent sequence. We then assume inductively that $S_r = \{x_1, \dots, x_r\} \subset C(a,b)$ is a T -independent set and attempt to add another element $x_{r+1} \in C(a,b)$ to S_r by always choosing x_{r+1} the smallest element in $C(a,b)$ so that $S_{r+1} = S_r \cup \{x_{r+1}\}$ is T -independent.

There are two required conditions for S_{r+1} to be T -independent. First, $|H_T(x_{r+1})| = |T|$ has already been verified. The other condition is that $H_T(x_{r+1}) \cap B_T(x_i) = \emptyset$ for $i \leq r$. Since x_1, \dots, x_r have already been chosen then there are at most $4k^2rs^w$ forbidden values for x_{r+1} . This can be seen since for each x_i and each $t \in T$, $B(x_i, t)$ consists of at most s^w elements. So there are at most $2krs^w$ values which are forbidden for $H_T(x_{r+1})$. Since $|H_T(x_{r+1})| \leq 2k$, there are at most $4k^2rs^w$ forbidden values for x_{r+1} . We have already shown that $|C(a,b)| > \frac{K}{\sqrt{b-a} \left(\frac{\epsilon}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$.

Thus, in order to show that suitable values exist for x_{r+1} , it is sufficient to show that $4k^2rs^w < \frac{K}{\sqrt{b-a} \left(\frac{\epsilon}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$.

Recall that $s^w \leq 2^{(b-a)\epsilon}$, and let $C_1 = \frac{K}{4k^2}$. Then the following inequalities give sufficient conditions for a T -independent sequence of length r .

$$r < \frac{C_1}{\sqrt{b-a} 2^{(b-a)\epsilon} \left(\frac{\epsilon}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$$

$$r < \frac{C_1}{\sqrt{b-a} \left(\frac{\epsilon s}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$$

Thus there is a T -independent sequence of length at least $L = \frac{C_1}{\sqrt{b-a} \left(\frac{\epsilon s}{\log s}\right)^{\frac{(b-a)\epsilon}{\log s}}}$.

Now using the estimates $\frac{1}{2k+1} \log n - 2 - 2\delta \leq b-a \leq \log n$ we see that there is a T -independent sequence of length at least $L = \frac{C}{\sqrt{\log n} \left(\frac{\epsilon s}{\log s}\right)^{\frac{\epsilon \log n}{(2k+1) \log s}}}$ where

$$C = \frac{C_1}{\left(\frac{\epsilon s}{\log s}\right)^{-2(1+\delta)\epsilon/\log s}}.$$

□

Lemma 3.2 : Suppose v_1, \dots, v_r is a T -independent sequence for some set $T \subset V_n$ of order $2k$. Let G be any uniform generalized Steinhaus graph of type s with n vertices and fixed diagonal generators. Then by changing only the entries in the top generating string indexed by $H_T(v_i)$ it is possible to attain any combination of adjacencies between v_i and the vertices in T . Furthermore, making the changes does not change the adjacencies between v_j and the vertices of T for any $j < i$.

Proof: Let $(a_{1,j})_{j=2,\dots,n}$ be an arbitrary top generating string for a uniform generalized Steinhaus graph of type s with fixed diagonal generators which gives an adjacency matrix $(a_{i,j})$. Label the elements in T by t_1, \dots, t_r where $H(v_i, t_1) > \dots > H(v_i, t_r)$. It is then clear that by changing the value of $a_{1,H(v_i, t_i)}$, the value a_{v_i, t_i} changes. Furthermore, changing $a_{1,H(v_i, t_i)}$ does not change the value of a_{v_i, t_k} for $k < i$, nor does it change the value of a_{v_j, t_z} for any $j < i$ and any $1 \leq z \leq 2k$. \square

Proof of Theorem 1.1: Let $T = \{v_1, v_2, \dots, v_k\} \cup \{w_1, w_2, \dots, w_k\}$ be as in Axiom k . By Lemma 3.1 for n sufficiently large there is a T -independent sequence of length at least

$$L = \frac{C}{\sqrt{\log n} \left(\frac{\epsilon s}{\log s} \right)^{\frac{\epsilon \log n}{(2k+1) \log s}}}$$

Let this sequence be x_1, \dots, x_r , $r \geq L$. Note that the top generating string for a generalized Steinhaus graph of type s with fixed diagonal generators for which Axiom k fails for x_1, \dots, x_{j-1} can be partitioned into subsets each of size 2^{2k} by putting two strings in the same subset if the strings agree in each entry except for positions $(1, i)$ where $i \in H_T(x_j)$. By Lemma 3.2 in each subset there is a sequence whose generalized uniform Steinhaus graph of type s with fixed diagonal generators satisfies Axiom k using T and x_j . Therefore the probability that a generalized uniform Steinhaus graph of type s with fixed diagonal generators does satisfy Axiom k using T and any x_i with $i < j$ is at least m_n^{2k} . Consequently the probability of failure $\forall x_i, 1 \leq i \leq r$ is at most $(1 - m_n^{2k})^L$. Thus the probability of failure of Axiom k is at most $P_n = \binom{n}{k} \binom{n-k}{k} (1 - m_n^{2k})^L$.

We complete the proof by verifying that P_n approaches 0 as n approaches ∞ . Clearly $P_n < n^{2k}(1 - m_n^{2k})^L$ so the requirement P_n approaches 0 will be met provided $2k \log n + L \log(1 - m_n^{2k})$ approaches $-\infty$. Equivalently, it is sufficient to show $2k \log n - L m_n^{2k} \rightarrow -\infty$. Let $\epsilon_1 = \frac{\epsilon}{2k(2k+1) \log s} (\log \log s - \log s - \log \epsilon) - \gamma$. Then

$$\begin{aligned} 2k \log n - L m_n^{2k} &< 2k \log n - \frac{C}{\sqrt{\log n} \left(\frac{\epsilon s}{\log s} \right)^{\frac{\epsilon \log n}{(2k+1) \log s}}} n^{-2k\gamma} \\ &= 2k \log n - \frac{C}{\sqrt{\log n} \left(\frac{\epsilon s}{\log s} \right)^{\frac{\epsilon \log n}{(2k+1) \log s}}} n^{-2k} \left(\frac{\epsilon}{2k(2k+1) \log s} (\log \log s - \log s - \log \epsilon) - \epsilon_1 \right) \\ &= 2k \log n - C \frac{n^{2k\epsilon_1}}{\sqrt{\log n}} \\ &\rightarrow -\infty \end{aligned}$$

Note that the probability estimates used are independent of how the values of the diagonal generators are fixed. This implies there is a function g as stated in the theorem. \square

Note that Corollary 1.1 follows from Theorem 1.1 since the probability estimates given in the proof are independent of how the values of the diagonal generators are fixed.

References

- [1] N. Brand and M. Morton, Generalized Steinhaus graphs, J. Graph Theory, in press.
- [2] N. Brand, S. Jackson, Properties of classes of random graphs, submitted to Combinatorics, Probability and Computing.
- [3] A. Blass and F. Harary, Properties of almost all graphs and complexes, J. Graph Theory **3**(1976) 225-240.
- [4] R. Fagin, Probabilities on finite models, J. Symbolic Logic **41**(1976) 50-58.
- [5] S. Willson, Cellular automata can generate fractals, Discrete Applied Mathematics **8**(1984) 91-99.
- [6] S. Willson, Calculating growth rates and moments for additive cellular automata, Discrete Applied Mathematics **35**(1992) 47-65.

(Received 16/12/94; revised 8/6/95)