Scan add to 4
A LATTICE PATH PROBLEM

H.L. Abbott and D. Hanson

1. Introduction.

The following problem was brought to our attention a few years ago by Pavol Hell. A certain rectangular voting district, m miles by n miles, is to be divided into two connected constituencies, where the division is to take place only along lines determined by the mileage markers. In how many ways can this be done? We consider the following problem which offers a partial solution to the above problem. Consider the integer lattice \( I_{mn} = \{(x,y): 0 \leq x \leq m, 0 \leq y \leq n, \ x,y \text{ integers}\} \). Let \( f(m,n) \) denote the number of lattice paths on \( I_{mn} \) from \((0,0)\) to \((m,n)\) that have the property that no vertex occurs more than once on any given path. We shall call such a path a non-intersecting path. It is trivial that \( f(m,0) = 1 \) and it is easy to check that \( f(m,1) = 2^m \) for all \( m \geq 1 \).

In general, the problem of evaluating \( f(m,n) \) appears to be difficult. In this paper we evaluate \( f(m,2) \) and obtain bounds for \( f(m,n) \) for \( n \geq 3 \).

2. A Matrix Setting for the Problem.

Consider any non-intersecting path from \((0,0)\) to \((m,n)\). Label each unit square inside \( I_{mn} \) that lies to the "right" of the path by a 1 and those that lie to the "left" by a 0. This labelling then
defines a one-to-one correspondence between the set of all non-intersecting paths on \( I_{mn} \) and a subset of the set of 0-1 \( n \times m \) matrices. For example we have the following correspondence:

\[
\begin{bmatrix}
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 1
\end{bmatrix}
\]

An \( n \times m \) 0-1 matrix will be called admissible if it corresponds to a non-intersecting path on \( I_{mn} \) and inadmissible otherwise. Since the total number of \( n \times m \) 0-1 matrices is \( 2^{mn} \) we have the following trivial upper bound

\[(1) \quad f(m, n) \leq 2^{mn}.
\]

We say that two entries \( a_{ij} \) and \( a_{kl} \) of an admissible matrix are 0-adjacent (respectively, 1-adjacent) if and only if \( i = k \) and \( |j - l| = 1 \) or \( j = l \) and \( |i - k| = 1 \) and \( a_{ij} = a_{kl} = 0 \) (respectively, \( a_{ij} = a_{kl} = 1 \)). Further, we say that \( a_{ij} \) and \( a_{kl} \) are 0-path (respectively, 1-path) adjacent if there is a sequence of entries \( a_{ij}, \ldots, a_{kl} \) each of which is 0-adjacent (respectively, 1-adjacent) to its neighbors in the sequence. For convenience, we agree that a zero (one) entry is 0-path (1-path) adjacent to itself.

The following two observations are immediate:

(a) Any zero entry of an admissible matrix is 0-path adjacent to some entry in the first row or first column.

(b) Any one entry of an admissible matrix is 1-path adjacent to some entry in the last row or last column.

We now define for each \( n \geq 2, 2^{n} \times 2^{n} \) matrices \( B_n, \overline{B}_n \) and \( \overline{B}_n \).

Associate with a column of 0-1 matrix, reading from top to bottom, a binary number in the natural way. The matrix \( B_n = [b_{ij}] \), \( 0 \leq i \leq 2^{n}-1, 0 \leq j \leq 2^{n}-1 \) is now defined as follows:

\[ b_{ij} = 1 \text{ if every admissible } n \times (m-1) \text{ matrix whose (m-1)st column is given by the binary representation of } j \text{ can be extended to an admissible } n \times m \text{ matrix by adjoining an } m \text{th column given by the binary representation of } i. \]

\[ b_{ij} = 0 \text{ if every admissible } n \times (m-1) \text{ matrix whose (m-1)st column is given by the binary representation of } j \text{ yields an inadmissible } n \times m \text{ matrix when we adjoin an } m \text{th column given by the binary representation of } i. \]

\[ b_{ij} = * \text{ (unspecified) if there are admissible } n \times m \text{ matrices whose } m\text{th and } (m-1)\text{st columns are given by the binary representation of } i \text{ and } j \text{ respectively, but which are such that deleting the } m\text{th column does not result in an admissible } n \times (m-1) \text{ matrix.} \]

The matrix \( B_n \) is thus well defined for every \( n \geq 2 \). \( B_n \) is the matrix obtained from \( B_n \) by replacing each * entry by 0 and \( \overline{B}_n \)
3. Evaluation of $f(m, 2)$.

The following additional observations follow easily from observations (a) and (b) in Section 2. $A_{nm} = [a_{ij}]$ denotes an $n \times m$ matrix and we suppose $m \geq 2$ throughout.

(c) Any admissible $A_{2,m-1}$ can be extended to an admissible $A_{2,m}$ by adjoining an $m$th column $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(d) Any admissible $A_{2,m-1}$ whose $(m-1)$st column is either $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ can be extended to an admissible $A_{2,m}$ by adjoining an $m$th column $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(e) Any admissible $A_{2,m-1}$ whose $(m-1)$st column is not $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ can be extended to an admissible $A_{2,m}$ by adjoining an $m$th column which is either $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(f) If $A_{2,m-1}$ is inadmissible but can be extended to an admissible $A_{2,m}$, then it is inadmissible because $a_{2,m-1} = 0$ and there is no 0-path adjacent entry $a_{1,p}$, $1 \leq p \leq n$ or a 0-path adjacent entry $a_{q,1}$, $1 \leq q \leq 2$. That is, the only inadmissible $A_{2,m-1}$ matrices that can be extended to an admissible $A_{2,m}$ are some of those whose $(m-1)$st column is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and these can be extended by adjoining an $m$th column $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Further, all such matrices must have as their last $k+1$ columns, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ followed by $k$ columns $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and the first $m-k-2$ columns must form an admissible $A_{2,m-k-2}$ for some $k$, $1 \leq k \leq m-2$.

From observations (c), (d), (e) and (f) we have that $B_2$, as defined in Section 2, is given by

$$B_2 = \begin{bmatrix} 1 & 1 & * & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

**Lemma 1.** For $m \geq 3$,

$$f(m, 2) = 4f(m-1, 2) - 3f(m-2, 2) + 2f(m-3, 2) + f(m-4, 2).$$

**Proof.** Let $a_j^m$ be the number of admissible matrices $A_{2,m}$ whose last column is given by the binary representation of $j$. Then, since $f(m, 2) = a_j^m$, we must show, for $m \geq 3$,

$$a_j^{m+1} = 4a_j^m - 3a_j^{m-1} + 2a_j^{m-2} + a_j^{m-3}.$$  \hspace{1cm} (2)

It follows from observations (c), (d), (e) and (f) above that

$$a_j^m = a_j^{m-1} + a_j^{m-1} + a_j^{m-1} + \sum_{i=1}^{m-1} a_j^{m-1}.$$  \hspace{1cm} (3)

We have, by repeatedly appealing to the various equations given by (3),
It now follows immediately from the lemma that the following theorem holds.

**THEOREM 1.** For \( m > 0 \),

\[
f(m, 2) = \frac{4+\sqrt{13}}{2} \left( \frac{3+\sqrt{13}}{2} \right)^{m} + \frac{\sqrt{13}-4}{2} \left( \frac{3-\sqrt{13}}{2} \right)^{m} + \frac{1}{2^{m/2}} \cdot \left( \frac{1+i\sqrt{3}}{2} \right)^{m}.
\]

Remark: Theorem 1 shows that \( f(m, 2) \sim c(1.817...)^{2m} \) for some constant \( c \). It also follows from the lemma that \( f(m, 2) \) has the following rational generating function

\[
\frac{1-x}{1-4x+3x^{2}-2x^{3}} = \sum_{m=0}^{\infty} f(m, 2)x^{m}.
\]

A similar approach may be used to find the number of non-intersecting paths in \( I_{m,2} \) from \((0,0)\) to a point of \( I_{m,2} \) other than \((m,2)\). The complexity of the solution will vary with the choice of the terminal point. For example, we count the number of non-intersecting paths, say \( g(m) \), from \((0,0)\) to \((0,2)\). The matrix analogous to \( B_{2} \) is

\[
\begin{bmatrix}
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 \\
1 & 1 & 1 & 1
\end{bmatrix}
\]

Thus, corresponding to (3), defining the \( a_{3}^{m} \) in the natural way, we have

\[
G(87) 6190 6191 6192
\]
The matrices $B_n$ and $\bar{B}_n$ may be used to define sequences $a(m,n,j)$ and $\bar{a}(m,n,j)$ as follows:

$$a(m,n,j) = \sum_{k=0}^{2^{n-1}} a(m-1,n,k) b_{jk}$$

$$\bar{a}(m,n,j) = \sum_{k=0}^{2^{n-1}} \bar{a}(m-1,n,k) \bar{b}_{jk}$$

where $a(1,n,j) = \bar{a}(1,n,j) = 1$ for all $n$ and all $j$, $0 \leq j \leq 2^{n-1}$. If we further define $f(m,n) = a(m-1,n,2^{n-1})$ and $\bar{f}(m,n) = \bar{a}(m-1,n,2^{n-1})$ we then have

$$\bar{f}(m,n) \leq f(m,n) \leq \bar{f}(m,n),$$

and

**THEOREM 2.** For $m \geq 2$,

$$f(m,n) = \text{sum of the entries in } B_n^{m-1}$$

$$\bar{f}(m,n) = \text{sum of the entries in } \bar{B}_n^{m-1}.$$

**Proof.** It will suffice to prove the result for $f(m,n)$.

Let $e_m = [a(m,n,0), a(m,n,1), \ldots, a(m,n,2^{n-1})]$. It follows from (4) that for $m \geq 2$

$$a(m,n,j) = e_{m-1} \cdot \text{row } j \text{ of } B_n^T$$

and hence that

$$e_m = e_{m-1} \cdot B_n^T = \ldots = e_1 \cdot (B_n^T)^{m-1}.$$

Since $e_1 = [1,1,\ldots,1]$ we find that

\[ a(m,n,j) = e_{m-1} \cdot \text{row } j \text{ of } B_n^T = \ldots = e_1 \cdot (B_n^T)^{m-1}. \]
where \( \lambda \) and \( \overline{\lambda} \) are, respectively, the largest eigenvalues of \( B_n \) and \( \overline{B}_n \) and \( c_n \) and \( \overline{c}_n \) are constants depending solely on \( n \). (Note that \( b_{4,1}, i = 0,1,\ldots,2^n-1, b_{2^n-1,1}, i = 0,1,\ldots,2^n-1, \) and \( b_{1,2^n-1} \) are all 1's and therefore, for all \( n > 1 \), \( B_n \) and \( \overline{B}_n \) are indeed primitive matrices.)

In the case \( n = 3 \) one may verify that

\[
\begin{bmatrix}
1 & 1 & * & 1 & * & * & * & 1 \\
1 & 1 & 0 & 1 & * & * & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

Thus \( B_3 \) and \( \overline{B}_3 \) have the characteristic polynomials

\[
P_3(\lambda) = \lambda^2 (\lambda - 8\lambda^5 + 91\lambda^3 - 18\lambda^4 + 6\lambda^2 + 8\lambda - 4)
\]

and

\[
P_{\overline{B}}_3(\lambda) = \lambda(\lambda - 8\lambda^6 + 91\lambda^5 - 18\lambda^4 + 6\lambda^3 - 6\lambda^2 + 2\lambda + 1)
\]

respectively. One then finds that the largest eigenvalues of \( B_3 \) and \( \overline{B}_3 \) are \( \lambda_3 = 4.78283 \ldots \) and \( \overline{\lambda}_3 = 5.56165 \ldots \) respectively and therefore

\[
c_3 (1.684 \ldots)^3 < f(m,3) < \overline{c}_3 (1.772 \ldots)^3.
\]

It is clear that the above method may be used in principle to estimate \( f(m,n) \) for \( n > 3 \). However, calculation of the associated characteristic polynomials becomes an extremely tedious task. One may, however, use the following alternate approach: It follows from (6)
that if \( A \) is a primitive 0-1 matrix then
\[
\lim_{k \to \infty} e^{-A} \cdot e^{T} = \lambda.
\]

Thus \( \lambda \) may be approximated by calculating successive ratios of sums of entries of the powers of \( A \). Of course it is difficult to measure the rate of convergence to \( \lambda \). However our calculations suggest that in the case \( n = 4 \)
\[
c_4(1.661...)^{4m} < f(m,4) < c_4(1.757...)^{4m}.
\]

We do not give all of the details here but we record the matrix \( B_4 \) so that the interested reader may verify matters for himself.

\[
\begin{bmatrix}
1 & 1 & * & 1 & * & 1 & * & 1 & * & 1 & * & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & * & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

We now prove a theorem which enables us to obtain estimates of \( f(m,n) \) in terms of \( f(m,k) \), \( k < n \).

**Theorem 3.** Let \( t \geq 0 \), \( k \geq 1 \) and let \( n = 2t(k+1)+k \). Then
\[
f(m,n) \geq f(m,k)^{2t+1}.
\]

**Proof.** Divide the integer lattice \( I_{m,n} \) into \( 2t+1 \) horizontal rectangles \( I_1, I_2, \ldots, I_{2t+1} \) of width \( k+1 \). That this is possible follows from the definition of \( n \) in terms of \( t \) and \( k \). In fact we have, for \( j = 1, 2, \ldots, 2t+1 \)
\[
I_j = \{(x,y) : 0 \leq x \leq m \text{, } (k+1)(j-1) \leq y \leq (k+1)(j-1)+k\}.
\]

For \( j \) odd let \( p_j \) denote any one of the \( f(m,k) \) non-intersecting paths in \( I_j \) from \((0,(k+1)(j-1))\) to \((m,(k+1)(j-1)+k)\) and for \( j \) even let \( p_j \) denote any one of the \( f(m,k) \) non-intersecting paths in \( I_j \) from \((m,(k+1)(j-1))\) to \((0,(k+1)(j-1)+k)\). These paths may then be juxtaposed to give a non-intersecting path in \( I_{m,n} \). The number of non-intersecting paths in \( I_{m,n} \) is thus at least \( f(m,k)^{2t+1} \).

**Corollary.** For each fixed \( n \),
\[
\lim_{m \to \infty} \frac{1}{f(m,n)^{mn}} = \beta \text{ exists.}
\]

**Proof.** Let \( a = \lim_{m \to \infty} \inf f(m,n)^{mn} \leq \lim_{m \to \infty} \sup f(m,n)^{mn} = \gamma \). Let \( \varepsilon > 0 \) and let \( k \) be such that
\[
\frac{1}{f(k,n)} > \gamma - \varepsilon \text{ and } (\gamma - \varepsilon) - \frac{1}{k} > \gamma - 2\varepsilon.
\]

Henceforth \( k \) is fixed. Let \( m \) be given and let \( t \) be defined by \( 2t(k+1) + k \leq 2(t+1)(k+1)+k \). We may suppose that \( m \) is so large that
Then, by Theorem 3 we have

\[ f(m, n) = f(n, m) \geq f(n, 2t(k+1)+k) \geq f(n, k)^{2t+1} \]

so

\[ f(k, n)^{2t+1} > \frac{1}{(\gamma-\varepsilon)^{m}} (\gamma-\varepsilon)^{k(2t+1) - \frac{1}{k}} \]

Thus

\[ \lim_{m \to \infty} \inf_{k} \frac{1}{m} \left( \frac{1}{m} \right)^{k(2t+1) - \frac{1}{k}} > \frac{1}{m} \left( \frac{1}{m} \right)^{\gamma - 2\varepsilon} \]

from which it follows that \( \alpha = \gamma \) as required.

In a similar manner, it follows that \( \beta = \lim_{m \to \infty} f(m, m)^{2m} \) exists.

In fact, it follows easily from Theorem 3 that

\[ \beta > \frac{1}{\frac{k}{k+1}} \]

for all \( k \). Our calculations, based on (7), indicate that \( \beta > 1.653 \)...

and hence that

\[ f(m, m) > c(1.583...)^{2m} \]

Finally, we indicate how upper bounds for \( f(m, n) \) for \( n \leq m \) may be obtained. Note that of the 16 ways of partitioning the unit squares of \( I_{2,2} \) into two sets, at most 14 can correspond to portions of non-intersecting lattice paths. Specifically, the patterns \([0, 1] \) and \([1, 0] \) are not possible. Thus, if \( m \) and \( n \) are both even, say \( m = 2r, n = 2s \),

\[ f(m, n) \leq 14^{rs} \sim (1.935...)^{mn} \]

and a slightly weaker bound follows if \( m \) and \( n \) are not both even. Similarly, one may show that of the \( 2^{16} = 65536 \) ways of partitioning the unit squares of \( I_{4,4} \) into two sets, at most 22662 can correspond to portions of non-intersecting paths. Then, if \( m \) and \( n \) both divisible by 4 one has

\[ f(m, n) \leq 22662 \sim (1.872...)^{mn} \]

with slightly weaker results in the remaining cases. No doubt further such improvements are possible.

5. Some Open Questions.

Clearly there are a number of unanswered questions which remain. We summarize some of these.

1. Evaluate \( \beta \) for \( n \geq 3 \) and evaluate \( \frac{1}{n} \).

2. Is the sequence \( f(m, n)^{mn} \) decreasing in both \( m \) and \( n \)? The first few values of \( f(m, n) \) and \( f(m, n)^{mn} \) are given in the following tables. If \( f(m, n)^{mn} \) is decreasing, it would follow that \( f(m, n) < c(1.758...)^{mn} \) for \( m \geq 5, n \geq 4 \).

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>64</td>
<td>128</td>
<td>256</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>12</td>
<td>38</td>
<td>125</td>
<td>414</td>
<td>1369</td>
<td>4522</td>
<td>14934</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>184</td>
<td>976</td>
<td>5382</td>
<td>29739</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>8512</td>
<td>79384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. For even \( n \), let \( h(n) \) denote the number of non-intersecting paths in \( T_{mn} \) which are Hamiltonian. We can show that \( a = \lim_{n \to \infty} \frac{1}{n} h(n)^{\frac{1}{2}} \) exists and that \( \frac{1}{3} \leq a \leq \frac{1}{4} \). We believe \( a < 1 \).

4. For \( k \) even, \( 2n \leq k \leq n^2 + 2n \) let \( f_k(n) \) be the number of non-intersecting paths from \((0,0)\) to \((n,n)\) which have exactly \( k \) edges, so that

\[
 f(n,n) = \sum_{k=2n}^{n^2+2n} f_k(n).
\]

\( f_2n(n) = \binom{2n}{n} \) is the classical ballot problem result. The evaluation of \( f_k(n) \) for \( k > 2n \) appears difficult.

REFERENCE


University of Alberta, University of Regina,
Edmonton, Alberta Regina, Saskatchewan.

A BALANCED HYPERGRAPH DEFINED BY CERTAIN SUBTREES OF A TREE

by

Rick Giles

ABSTRACT

Let \( T \) be a tree with vertex set \( V = \{v_1, v_2, \ldots, v_p\} \). For each \( i, 1 \leq i \leq p \), let \( a_i \) be a nonnegative integer and define

\[ E_i = \{v \in V : d(v, v_i) \leq a_i\}. \]

We show that the hypergraph \( (V, E) \), where \( E = \{E_1, E_2, \ldots, E_p\} \), is balanced. This result generalizes two previously known min-max relations for this hypergraph.

Let \( T \) be a tree with vertex set \( V = \{v_1, v_2, \ldots, v_p\} \). For each \( i, 1 \leq i \leq p \), let \( a_i \) be a nonnegative integer and define

\[ E_i = \{v \in V : d(v, v_i) \leq a_i\}. \]

Thus each \( E_i \) is the vertex set of a subtree of \( T \) "centred" at \( v_i \). Let \( R = \{(v_1, a_1), (v_2, a_2), \ldots, (v_p, a_p)\} \) and associate a hypergraph \( H_{T,R} = (V, E) \) with \( T \) and \( R \), where \( E = \{E_1, E_2, \ldots, E_p\} \). The hypergraph terms we will use can be found in Berge [2].

Min-max theorems for \( H_{T,R} \) have been obtained in [3] and [5]. One result from [3] implies that the maximum cardinality of matching in \( H_{T,R} \) is equal to the minimum size of a transversal of \( H_{T,R} \). The main result of [5] is that the maximum cardinality of a strongly stable set of vertices of \( H_{T,R} \) is equal to the minimum size of a cover in \( H_{T,R} \). These min-max results are also consequences of a more general property of \( H_{T,R} \).