A variation of the Stern-Brocot tree

COLIN L. MALLows

We study a variation of the Stern-Brocot tree, in which not one but two fractions are inserted between each existing pair. Relating this tree to the original one gives rise to a permutation of the natural numbers.

KEYWORDS AND PHRASES: Stern-Brocot tree, permutations of \(\mathbb{N}\).

1. The Stern-Brocot tree, and a variation

The Stern-Brocot tree (or rather half of it) can be defined as follows. Start with two fractions 0/1 and 1/1, forming an ordered set \(S_0\). (Throughout this paper, “fraction” means “fraction in lowest terms”.) At stage \(k\), \((k = 1, 2, \ldots)\), form a new set \(S_k\) by inserting between each pair of adjacent fractions in \(S_{k-1}\), say \(p/q\) and \(r/s\), the fraction \((p+r)/(q+s)\). Name the (ordered) set of fractions that are introduced at this stage \(R_k\). Thus \(R_1 = \{1/2\}\), \(R_2 = \{1/3, 2/3\}\), \(R_3 = \{1/4, 2/5, 3/5, 5/4\}\), \(R_4 = \{1/5, 2/7, 3/8, 3/7, 4/7, 5/8, 5/7, 4/5\}\) etc. \(R_k\) has \(2^{k-1}\) elements. It is well known (see e.g. [1]) that every proper fraction appears (exactly once) in some \(R_k\), and that adjacent fractions \(p/q, r/s\) satisfy

\[(1) \quad |qr - ps| = 1.\]

We define a new tree (first noticed in [2, Section 9]) starting with \(S'_0 = S_0\). At the \(k\)-th stage insert two fractions between each existing adjacent pair in \(S'_{k-1}\), namely between \(p/q\) and \(r/s\) (where \(p\) is even and \(r\) is odd), insert \((p+r)/(q+s)\) and \((p+2r)/(q+2s)\). Notice that we may have either \(p/q < (p+r)/(q+s) < (p+2r)/(q+2s) < r/s\) or the same with all the inequalities reversed. It is easy to see that every adjacent pair of fractions in \(S'_k\) satisfy (1) and that the numerators of successive fractions in \(S'_k\) are alternately even and odd, so that the insertion rule is well-defined. Successive generations of insertions are denoted \(R'_1, R'_2, \ldots\). Thus \(R'_k\) has \(2.3^{k-1}\) elements.
Explicitly,

\[ R_1' = (1/2, 2/3), \quad R_2' = (1/3, 2/5, 4/7, 3/5, 3/4, 4/5), \]
\[ R_3' = (1/4, 2/7, 4/11, 3/8, 2/7, 4/9, 6/11, 5/9, 7/12, 10/17, 8/13, 5/8, 5/7, 8/11, 10/13, 7/9, 5/6, 6/7). \]

**Lemma 1.** For every proper fraction \( x \), there is a \( k \) such that \( x \) appears in \( R_k' \).

**Proof.** Define the “\( ndsum \)” of a fraction \( p/q \) to be \( p+q \). An easy induction shows that for \( k \geq 1 \) the minimum \( ndsum \) in the row \( R_k' \) is \( k+2 \) (attained by the first element, which is \( 1/(k+1) \)). The minimum \( ndsum \) in \( S_0' \) is 1, attained by 0/1. Suppose the fraction \( a/b \), where \( a+b \geq 2 \), does not appear in any \( R_k' \). Consider the row \( R_{a+b} \). There must be a fraction \( p/q \) in this row and a fraction \( r/s \) in \( S_{a+b} \) such that \( |qr-ps| = 1 \) and \( p/q < a/b < r/s \), or the same with both inequalities reversed. Suppose the inequalities are as shown. Then \( aq-bp > 0 \), so \( aq-bp \geq 1 \), and similarly \( br-as \geq 1 \). Thus

\[
(p+q)(br-as) + (r+s)(aq-bp) \geq p+q+r+s.
\]

But the l.h.s. of (2) equals \( (a+b)(qr-ps) = a+b \), and the r.h.s. is at least \( (a+b+2)+1 \), so \( a+b \geq a+b+3 \) which is a contradiction. When the inequalities are reversed, the argument is similar. \( \square \)

### 2. Relating the two trees

We study the relation between the sets \( \{R_k'\} \) and \( \{R_k\} \). We find that (as far as we have computed, namely \( R_6' \) and \( R_{12} \)) there is a sequence \( p \), starting

Sequence \( p \)

\[
1, 2, 5, 3, 4, 8, 17, 9, 10, 20, 11, 6, 7, 14, 29, 15, 16, 32, 65, 33, 34, 68, 35, 18, 19,
\]

such that for each \( k \), and for \( i = 1, 2, \ldots, 2.3^{k-1} \), the fraction \( R_k'(i) \) appears as \( R_{k'(i)}(p(i)) \) for some \( k'(i) \). We write \( k'(i) = k + r_k(i) \), and set \( n_k = 2^{k-1} \), which is the sequence of lengths of the rows \( k' \) of \( R \) in which these fractions appear. Thus for \( k = 3 \), the rows of the following matrix \( M_3 \) are

- the numerators of fractions in \( R_3' \)
- the corresponding denominators
- the \( m \) such that each such fraction appears in \( R_{m+3} \)
- the length of the row \( R_{m+3} \) (this is \( n_{m+3} \))
- the position of this fraction in \( R_{m+3} \) (this is a prefix of \( p \)).
Rows 3 and 5 of the first six columns of this matrix give the corresponding results for $R'_2$, while the fourth row is twice the fourth row for $R'_2$.

We have studied similar matrices through $k = 6$, finding that for each $k$, the fifth row of $M_k$ contains the first $2^{3k-1}$ elements of the sequence we have called $p$. The third row contains numbers in the range $(0, k)$, with successive entries equal or consecutive.

The following lemma shows how the sequence for $R'_{k+1}$ can be obtained from that for $R'_k$.

**Lemma 2.** Given the finite sequences $p_k$ and $n_k$ that describe the relation of $R'_k$ to the rows of $S$, the sequences for row $R'_{k+1}$ are as follows.

$$p_{k+1} = (p_k, \text{rev}(3n_k + 1 - p_k), 3n_k + p_k),$$

$$n_{k+1} = (2n_k, \text{rev}(4n_k), 4n_k)$$

where “rev” means “the reverse of”.

**Proof.** The $n$ and $p$ sequences for $R'_{k+1}$ are unchanged if we replace the starting fractions $S_0$ and $S'_0$ by $(0/1, 1/2)$. So the (finite) $p$ sequence for $R'_k$ is the same as the first third of the $p$ sequence for $R'_{k+1}$, while the rows for $R_{m+1}$ are twice as long as those for $R_m$. Similarly, the final third of the $p$-sequence for $R'_{k+1}$, which relate to the interval $(2/3, 1/1)$, are the same as the sequence for $R'_k$, translated by $3/4$ of the length, which is four times the length for $R_m$. Finally, for the middle third, which relates to the interval $(1/2, 2/3)$, we have to read the $R_k$ values backwards (because the numerator of $1/2$ is odd and the numerator of $2/3$ is even) and count backwards from $3/4$ of the lengths.

This lemma makes it easy to compute $p$ as far as desired. However it has not led us to a proof that the sequence $p$ is a permutation of the natural numbers. We will show that another sequence, $pp$, which we have checked agrees with $p$ through 354, 294 terms, is indeed a permutation.

### 3. The sequences $b$ and $pp$

To approach the sequence $pp$, we must first define another sequence $b(N)$. 
Algorithm B. \( b(1) = 1 \). For \( k \geq 1 \),
\[
(b(3k - 1), b(3k), b(3k + 1)) = (4i - 1, 2i, 4i + 1)
\]
where \( i = b(k) \). Thus the sequence \( b \) begins
\[
1, 3, 2, 5, 11, 6, 13, 7, 4, 9, 19, 10, 21, 43, 22, 45, 23, 12, 25, 51, 26, 53, 27, 14, 29, \ldots
\]

**Theorem 1.** The sequence \( b(N) \) is a permutation of \( N \).

*Proof.* Suppose \( m \) is the smallest integer that does not appear as an element of \( b(N) \). It is impossible that \( m \) is even, since \( m/2 \) does appear, and for some \( k \) we have \( b(k) = m/2 \). Then \( m \) must appear at \( b(3k) \). If \( m \) is odd, set \( i = \text{round}(m/4) \). Then \( i \) appears at some point \( k \), \( b(k) = i < m \), so that \( m \) appears as an element of the triad centered at \( 3k \). Thus all integers must appear. A similar argument shows that no integer can appear twice. Suppose \( m \) is the smallest integer that appears twice. If \( m \) is even, we have \( b(3k_1) = b(3k_2) = m \), with \( k_1 \neq k_2 \). Then \( b(k_1) = b(k_2) = m/2 \) so that the integer \( m/2 \) appears twice before \( m \) does. Thus \( m \) cannot be even. If \( m \) is odd, suppose first that the smallest violation is \( b(3k_1 - 1) = b(3k_2 - 1) = 4i - 1 \), with \( k_1 \neq k_2 \). Then \( b(3k_1) = b(3k_2) = 2i \) so that \( b(k_1) = b(k_2) = i \), and \( i \) appears twice before \( 4i - 1 \) does. Similarly if \( m = 4i + 1 \).

We define another sequence \( pp \) by:

**Algorithm PP.** \( pp(1) = 1, pp(2) = 2 \). For \( k = 1, 2, \ldots \)
\[
(pp(4k - 1), pp(4k), pp(4k + 1), pp(4k + 2)) = (6i - 1, 3i, 3i + 1, 6i + 2)
\]
where \( i = b(k) \).

**Theorem 2.** The sequence \( pp(N) \) is a permutation of \( N \).

*Proof.* Since the sequence \( b \) is a permutation of \( N \), it is clear that numbers of the form \( 3i \) and \( 3i + 1 \) appear just once in \( pp \), in positions \( 4k \) and \( 4k + 1 \), and numbers of the form \( 3i - 1 \) appear in positions \( 4k - 1 \) and \( 4k + 2 \), where \( i = b(k) \).

We have verified that the sequences \( p \) and \( pp \) agree through their first 354,294 terms. Of course, this result does not prove anything about the sequence \( p \), merely that it agrees with the facts as far as we have computed them. We have not been able to prove that Algorithms P and PP generate the same sequence.
We think it remarkable that (it appears) Algorithm B and Algorithm P are so closely related, since $b$ generates blocks of length 4 in PP, while Algorithm P generates the sequence $p$ in blocks of length 2, 4, 12, 36, ... with the first half of each block involving reading previous blocks backwards.

4. Generalizations

Once we have the sequence $b$ in hand, we can generate many permutations of $\mathcal{N}$ by constructions similar to that in Algorithm PP. For example,

**Algorithm AA.** $aa(1, 2, 3) = (1, 2, 3)$. For $k \geq 1$,

$$
(aa(6k - 2), aa(6k - 1), aa(6k), aa(6k + 1), aa(6k + 2), aa(6k + 3))
= (8i - 2, 8i - 1, 4i, 4i + 1, 8i + 2, 8i + 3)
$$

where $i = b(k)$.

This particular sequence happens to be identical to one that makes no reference to the sequence $b$, but is generated by the following

**Algorithm A.** Set $a(1) = 1$. For $n \geq 1$:

\begin{align*}
(3) & \quad a(2n) = (1 + a(2n - 1))/2 \text{ if this value has not yet appeared} \\
(4) & \quad = 2a(2n - 1) \text{ else} \\
(5) & \quad a(2n + 1) = 1 + a(2n).
\end{align*}

The proof of this equality is left for another occasion.

There are other ways of defining a modified Stern-Brocot tree, but we have not found any as elegant as the one we have presented.

References


COLIN L. MALLOWS
AVAYA LABS
BASKING RIDGE
NJ 07920
USA
E-mail address: colinm@research.avayalabs.com

Received October 10, 2010