Distant parents in complete binary trees

S. P. GLASBY

Abstract. There is a unique path from the root of a tree to any other vertex. Every vertex, except the root, has a parent: the adjoining vertex on this unique path. This is the conventional definition of the parent vertex. For complete binary trees, however, we show that it is useful to define another parent vertex, called a *distant parent*. The study of distant parents leads to novel connections with dyadic rational numbers. Moreover, we apply the concepts of close and distant parent vertices to deduce an apparently new sense in which continued fractions are 'best' rational approximations.

2010 Mathematics subject classification: 05C05, 20E08, 11A55, 13-01
Keywords: Parent, children, infinite complete binary tree, string, inorder, continued fraction

1. Introduction

There is a unique path from the root of a tree to any other vertex. Hence each vertex in a tree with at most two children vertices can be associated with a string S of lefts and rights, an *LR-string*. We shall focus exclusively on (infinite) complete binary trees, where each vertex has precisely two children vertices. We henceforth abbreviate these as 'trees.' Rather than *identifying* the the vertices with LR-strings, it will be convenient to initially *label* vertices with LR-strings, see Fig. 1. The empty LR-string is denoted by ε. Thus the left and right children vertices of S are $C_L(S) := SL$ and $C_R(S) := SR$, respectively. If $S \neq \varepsilon$, then the conventional parent vertex is obtained by deleting the last symbol of S.

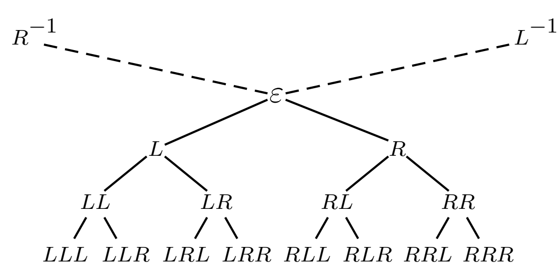

Figure 1. Infinite complete binary tree with vertices labeled by LR-strings.

In this paper we say that every string S has a left and right parent vertex denoted $P_L(S) = SR^{-1}$ and $P_R(S) = SL^{-1}$, respectively. The expressions SR^{-1} and SL^{-1} are evaluated recursively using the rules: $LL^{-1} = \varepsilon$, $RR^{-1} = \varepsilon$, $LR^{-1} = R^{-1}$, and $RL^{-1} = L^{-1}$. Thus when $S = L^2R^2$, for example, $P_L(S) = L^2R$ and $P_R(S) = L$. Every vertex $S \neq \varepsilon$ has a *close* and a *distant* parent vertex denoted $P_C(S)$ and $P_D(S)$, respectively. The former is the usual definition of parent, and the latter is studied in this note.

The aim of this note is to relate close and distant parents to dyadic[1] rationals via simple recurrence relations, or explicit formulae, see Theorems 1 and 2 in Section 2. Properties of distant parents are described using three metrics: the length $|S|$ of a string, its position $N(S)$ on a tree, and an order-preserving linear metric $r(S)$ defined later. Note that $|S|$ and $N(S)$ are natural numbers, while $r(S)$ is a dyadic rational number satisfying $0 \leqslant r(S) \leqslant 2$.

Infinite complete binary trees have strong connections with group theory [7, 8], with the theory of automata, and with the analysis of computer programs. However, this largely expository note focuses on elementary examples. An outline of this paper is as follows. Section 2 relates close and distant parents of a vertex S to the the numbers $|S|$, $N(S)$, and $r(S)$. In Section 3, an infinite complete binary tree whose vertices are continued fractions is considered. The children vertices are most naturally defined in terms of close and distant parents. Continued fractions are well-known to be associated with best rational approximations, see [6, 4.5.3. Ex. 42] and [3, p. 112]. For example, if $a = [a_0, a_1, \dots]$ and $b = [b_0, b_1, \dots]$ are irrational numbers, then the rational number c between a and b with smallest numerator or denominator is $c = [a_0, \dots, a_{k-1}, \min(a_k, b_k) + 1]$ where $a_i = b_i$ for $0 \leqslant i < k$ and $a_k \neq b_k$. We shall show in Theorem 4(d) that the close and distant parents to a continued fraction are the best lower-level rational approximations on a complete binary tree of all rationals.

2. The main results

In this section we define length $|S|$ of a string S, its position $N(S)$ on a tree, and an order-preserving linear function $r(S)$, see Fig. 2. These are related to the parent vertices of S. Let $S(k_0, k_1, \dots, k_m)$ denote the LR-string

$$S(k_0, k_1, \dots, k_m) := \begin{cases} R^{k_0}L^{k_1} \cdots L^{k_{m-1}}R^{k_m} & \text{if } m \text{ is even;} \\ R^{k_0}L^{k_1} \cdots R^{k_{m-1}}L^{k_m} & \text{if } m \text{ is odd;} \end{cases} \tag{1}$$

where $k_0 \in \mathbb{N} := \{0, 1, 2, \dots\}$ and $k_i \geqslant 1$ if $1 \leqslant i \leqslant m$. The *length* of $S = S(k_0, k_1, \dots, k_m)$ is defined to be $|S| = k_0 + k_1 + \cdots + k_m$. It counts the number of LR-symbols in S, and gives the *level* of S in the tree shown in Fig. 1. The *position* $N(S)$ of a string S is determined by Fig. 2(b), and a formula for $N(S)$ is given in Theorem 2 below.

The vertices of the tree in Fig. 1 can be ordered from left-to-right as the real numbers r in the interval $0 < r < 2$ are so ordered. Consider a vertical line through the vertex

[1] A dyadic rational is one whose denominator is a power of two.

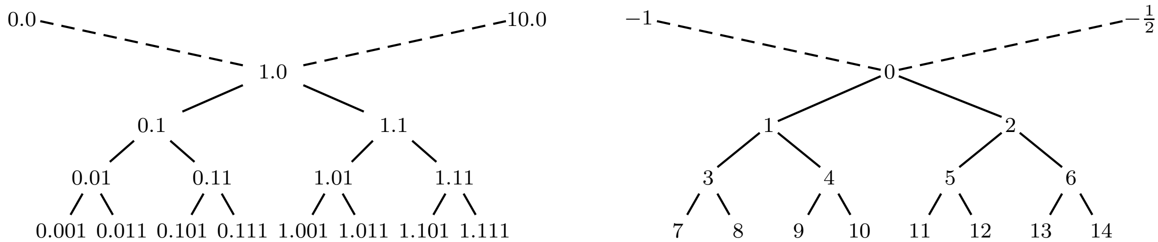

FIGURE 2. (a) binary expansions of $r(S)$; and (b) position values $N(S)$.

(string) S meeting the horizontal interval $0 < r < 2$ at the real number $r(S)$. The elements of the monoid $\{L, R\}^* := \{\varepsilon, L, R, L^2, LR, R^2, L^3, \dots\}$ will be called *strings*, and those of $\{L, R\}^\bigstar := \{L^{-1}, R^{-1}\} \cup \{L, R\}^*$ will be called *generalized string*. A convenient recursive definition of r is:

$$r(\varepsilon) = 1, \quad r(SL) = r(S) - 2^{-|SL|}, \quad r(SR) = r(S) + 2^{-|SR|} \quad \text{for } S \in \{L, R\}^*. \tag{2}$$

A simple induction shows that $r(S)$ is a dyadic rational. The value $r(R^{-1}) = 0$ is obtained from $r(SR) = r(S) + 2^{-|SR|}$ by substituting $S = R^{-1}$. Similarly, $r(L^{-1}) = 2$ is obtained from $r(SL) = r(S) - 2^{-|SL|}$ by substituting $S = L^{-1}$. Thus r extends to generalized strings. (Inorder traversal of a *finite* binary tree [5, §2.3.1] coincides with r-ordering.)

Moving left decreases the r-value, and moving right increases the r-value. However, a left move is not counteracted by any number of right moves; nor is a right move counteracted by any number of left moves. That is,

$$r(SL) < r(SLR) < r(SLR^2) < \cdots < r(S) < \cdots < r(SRL^2) < r(SRL) < r(SR). \tag{3}$$

Thus $r\colon \{L, R\}^\bigstar \to [0, 2]$ is an injective function which orders the generalized strings.

The generalized strings L^{-1} and R^{-1} have length (or level) -1, by definition. It follows from Theorem 1(c) below that the $S' \in \{L, R\}^\bigstar$ with $|S'| < |S|$ and $r(S') < r(S)$ which *maximizes* $r(S')$ is $P_L(S) = S'$. Similarly, the S' with $|S'| < |S|$ and $r(S') > r(S)$ which *minimizes* $r(S')$ is $P_R(S) = S'$. For this reason, the parents of a vertex in a tree are commonly 'best approximations' (in some sense) to the vertex.

Theorem 1. *Let $S \in \{L, R\}^*$ be a string, and let $m \geqslant 0$ be an integer. Then*

(a) $\{r(S) \mid |S| = m\}$ *equals* $\{\frac{2k-1}{2^m} \mid 1 \leqslant k \leqslant 2^m\}$;
(b) $\{r(S) \mid 0 \leqslant |S| \leqslant m\}$ *equals* $\{\frac{\ell}{2^m} \mid 1 \leqslant \ell \leqslant 2^{m+1} - 1\}$;
(c) *the following recurrences hold*

$$\begin{aligned} &P_L(\varepsilon) = R^{-1}, && P_L(SL) = P_L(S), && P_L(SR) = S; \\ &P_R(\varepsilon) = L^{-1}, && P_R(SL) = S, && P_R(SR) = P_R(S); \end{aligned}$$

(d) $\max(|P_L(S)|, |P_R(S)|) < |S|$, *and* $|P_L(S)| \neq |P_R(S)|$ *if* $S \neq \varepsilon$;
(e) $r(P_L(S)) = r(S) - 2^{-|S|}$ *and* $r(P_R(S)) = r(S) + 2^{-|S|}$;

(f) *if* $S = S(k_0, \dots, k_m)$*,* $X = S(k_0, \dots, k_{m-1}, k_m - 1)$*,* $Y = S(k_0, \dots, k_{m-2}, k_{m-1} - 1)$*, then the following formulas hold*

$$P_L(S) = \begin{cases} X & \text{if } m \text{ is even;} \\ Y & \text{if } m \text{ is odd;} \end{cases} \qquad P_R(S) = \begin{cases} Y & \text{if } m \text{ is even;} \\ X & \text{if } m \text{ is odd.} \end{cases} \tag{4}$$

Proof. (a) We use induction on m. The result is true when $m = 0$ as $r(\varepsilon) = 1$. Assume now that $|S| = m > 0$. Then S equals $S'L$ or $S'R$ where $|S'| = m - 1$. By induction, $r(S') = \frac{2k-1}{2^{m-1}}$ for a unique k with $1 \leqslant k \leqslant 2^{m-1}$. It follows from (2) that $r(S'L) = \frac{4k-3}{2^m}$ and $S'R = \frac{4k-1}{2^m}$, see Fig. 3. These r-values equal $\frac{2\ell-1}{2^m}$ for a unique ℓ with $1 \leqslant \ell \leqslant 2^m$. This proves part (a).

FIGURE 3. Children rules (a) for strings; (b) for r-values.

(b) By part (a), $\{r(S) \mid 0 \leqslant |S| \leqslant m\}$ equals $\bigcup_{i=0}^m \{\frac{2k-1}{2^i} \mid 1 \leqslant k \leqslant 2^i\}$. This is a disjoint union as the fractions are reduced, and hence distinct. The union has $\sum_{i=0}^m 2^i = 2^{m+1} - 1$ fractions, each of which is more than zero, and less than two. Since each fraction can be written with a denominator of 2^m, the union equals $\{\frac{\ell}{2^m} \mid 1 \leqslant \ell \leqslant 2^{m+1} - 1\}$, as desired.

(c) The initial conditions and the recurrences follow from the definition of $P_L(S)$ and $P_R(S)$ together with the rules for postmultiplying by L^{-1} and R^{-1} on page 2.

(d) Since $P_L(S) = SR^{-1}$ and $P_R(S) = SL^{-1}$, and $|S|$ counts the number of symbols (Ls and Rs) in S, it follows that $|P_L(S)| < |S|$ and $|P_R(S)| < |S|$. Suppose $S \neq \varepsilon$. One of $P_L(S)$ and $P_R(S)$, the close parent, has length $|S| - 1$ because precisely one symbol is canceled. For the distant parent, however, at least two symbols are canceled (this needs appropriate interpretation if $S = L$ or R). Hence $|P_L(S)| \neq |P_R(S)|$ if $S \neq \varepsilon$, and the parents of S lie on different levels.

(e) We use induction on $|S|$. The result is true for $|S| = 0$. Suppose now that $|S| > 0$. Then $S = S'L$, or $S = S'R$, for some $S' \in \{L, R\}^*$. Suppose $S = S'L$. Then $P_R(S) = S'$ and $r(S'L) = r(S') - 2^{-|S'L|}$ by (2). Thus $r(P_R(S)) = r(S) + 2^{-|S|}$. Since $S = S'L$, part (c) gives $P_L(S) = P_L(S')$, and induction gives $r(P_L(S')) = r(S') - 2^{-|S'|}$. Hence

$$r(P_L(S)) = r(P_L(S')) = r(S') - 2^{-|S'|} = r(S') - 2^{-|S|} - 2^{-|S|} = r(S) - 2^{-|S|}.$$

Similar arguments may be used to handle the case when $S = S'R$.

(f) Formula (4) needs interpretation when $m = 0$. In this case, $S = \varepsilon$, $k_0 = 0$, $X = R^{-1}$, and $Y = L^{-1}$. This agrees with $P_L(\varepsilon) = R^{-1}$ and $P_R(\varepsilon) = L^{-1}$. Suppose now that $m > 0$. The last symbol of S is R^{k_m} if m is even, and L^{k_m} if m is odd, where $k_m \geqslant 1$. Formula (4) now follows by canceling, as $P_L(S) = SR^{-1}$ and $P_R(S) = SL^{-1}$. $\square$

Counting the strings in $\{L, R\}^*$ in Fig. 1 from top down and then left-to-right gives the tree in Fig. 2b. Recursively define a bijective *position* function $N\colon \{L, R\}^* \to \mathbb{N}$ by

$$N(\varepsilon) = 0, \quad N(SL) = 2N(S) + 1, \text{ and } \quad N(SR) = 2N(S) + 2 \qquad \text{for } S \in \{L, R\}^*. \quad (5)$$

Substituting $S = L^{-1}$ into $N(SL) = 2N(S) + 1$ gives $N(L^{-1}) = -\frac{1}{2}$, and substituting $S = R^{-1}$ into $N(SR) = 2N(S) + 2$ gives $N(R^{-1}) = -1$.

As a consequence of Theorem 1(d), each $\varepsilon \neq S \in \{L, R\}^*$ has a *close parent*, denoted $P_C(S)$, and a *distant parent*, denoted $P_D(S)$. Set $n := N(S)$. Then $P_C(S)$ equals $P_L(S)$ if n is even, and $P_R(S)$ if n is odd. Similarly, $P_D(S)$ equals $P_L(S)$ if n is odd, and $P_R(S)$ if n is even. Table 1 suggests that $N(P_C(S)) = \left\lfloor \frac{N(S)-1}{2} \right\rfloor$ holds. This is easily proved. However, a formula for the numbers $N(P_D(S))$ is more mysterious. The integer-valued

$n = N(S)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
$N(P_C(S))$	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10
$N(P_D(S))$	-1	$-\frac{1}{2}$	-1	0	0	$-\frac{1}{2}$	-1	1	1	0	0	2	2	$-\frac{1}{2}$	-1	3	3	1	1	4	4	0

TABLE 1. Position numbers of close and distant parents

function $n \mapsto 2N(P_D(N^{-1}(n)))+1$ does not appear (at the time of writing) in the On-Line Encyclopedia of Integer Sequences, see http://oeis.org. A formula for $N(P_D(S))$ can be computed from (6a,b) below.

Theorem 2. *Set $S := S(k_0, \dots, k_{m-1}, k_m)$. The functions P_C, P_D, N, and r can be computed (nonrecursively) by the formulas*

$$P_C(S) = S(k_0, \dots, k_{m-1}, k_m - 1) \quad \text{and} \quad P_D(S) = S(k_0, \dots, k_{m-2}, k_{m-1} - 1); \tag{6a}$$

$$N(S) = \begin{cases} 2^{k_0+\cdots+k_m+1} - 2^{k_1+\cdots+k_m} + 2^{k_2+\cdots+k_m} - \cdots + 2^{k_m} - 2 & \text{if } m \text{ is even;} \\ 2^{k_0+\cdots+k_m+1} - 2^{k_1+\cdots+k_m} + 2^{k_2+\cdots+k_m} - \cdots - 2^{k_m} - 1 & \text{if } m \text{ is odd;} \end{cases} \tag{6b}$$

$$r(S) = 2\left(1 - 2^{-k_0} + 2^{-k_0-k_1} - \cdots + (-1)^m 2^{-k_0-k_1-\cdots-k_{m-1}} + (-1)^{m+1} 2^{-|S|-1}\right) \tag{6c}$$

$$= 2\left(\sum_{i=0}^{m+1} (-1)^i 2^{-\varepsilon_j}\right) + (-1)^{m+2} 2^{-|S|} \quad \text{where } \varepsilon_j = \sum_{j=0}^{i-1} k_j. \tag{6d}$$

Proof. Equation (6a) follows easily from the rules for postmultiplying by R^{-1} or L^{-1}. To verify formulas (6b) and (6c) it suffices to prove that they satisfy their respective recurrence relations. This is somewhat easier, paradoxically, than guessing the formulas in the first place. We begin by showing that (6b) satisfies the recurrence (5). First, $N(R^{k_0}) = 2^{k_0+1} - 2$ holds by (6b). Setting $k_0 = 0$ shows $N(\varepsilon) = 0$ which agrees with (5). Second, we must show $N(SL) = 2N(S) + 1$. When m is even, S ends in R^{k_m} and so $SL = S(k_0, \dots, k_{m-1}, k_m, 1)$. It follows from the second line of (6b) that $N(SL) = 2N(S) + 1$ holds. If m is odd, then S ends in L^{k_m} and so SL equals

$S(k_0, \ldots, k_{m-1}, k_m + 1)$. Again the second line of (6b) implies that $N(SL) = 2N(S) + 1$ holds. Similar reasoning involving the first line of (6b) shows that $N(SR) = 2N(S) + 2$ holds, independent of the parity of m. Hence (6b) is the (unique) solution to (5).

The last term of formula (6c) equals the last two terms of (6d) because

$$2(-1)^{m+1}2^{-|S|-1} = 2(-1)^{m+1}2^{-|S|} + (-1)^{m+2}2^{-|S|}.$$

Hence (6c) equals (6d). We now prove that the solution to the recurrence relation (2) is given by the formula (6c). The base case $r(\varepsilon) = 1$ accords with formula (6c). Suppose that $S = S(k_0, \ldots, k_m)$ where m is even. Then $SR = S(k_0, \ldots, k_{m-1}, k_m + 1)$. Comparing the expressions for $\frac{r(SR)}{2}$ and $\frac{r(S)}{2}$ given by (6c) yields

$$\frac{r(SR)}{2} = \frac{r(S)}{2} - (-1)^{m+1}2^{-|S|-1} + (-1)^{m+1}2^{-|SR|-1} \tag{7a}$$

$$= \frac{r(S)}{2} + (-1)^{m+2}2^{-|S|-2}(2 - 1) = \frac{r(S)}{2} + 2^{-|SR|-1}. \tag{7b}$$

If m is even, then $SL = S(k_0, \ldots, k_m, 1)$. By (6c), the last two terms of $\frac{r(SL)}{2}$ are

$$(-1)^{m+1}2^{-|S|} + (-1)^{m+2}2^{-|SL|-1} = (-1)^{m+2}2^{-|S|-2}(-2^2 + 1) = -3 \cdot 2^{-|S|-2}.$$

Comparing the expressions for $\frac{r(SL)}{2}$ and $\frac{r(S)}{2}$ given by (6c) yields

$$\frac{r(SL)}{2} = \frac{r(S)}{2} - (-1)^{m+1}2^{-|S|-1} - 3 \cdot 2^{-|S|-2} \tag{8a}$$

$$= \frac{r(S)}{2} + 2^{-|S|-2}(2 - 3) = \frac{r(S)}{2} - 2^{-|SL|-1}. \tag{8b}$$

Equations (7) and (8) accord with the recurrence relation (2). The proof when m is odd is similar. Hence the solution to the the recurrence relation (2) is (6c), as desired. $\square$

In Fig. 4 we compare the length function $N\colon \{L, R\}^* \to \mathbb{N}$ with another length function $M\colon \{L, R\}^* \to \mathbb{N}$ defined by $M(\varepsilon) = 0$ and $M(S(k_0, k_1, \ldots, k_m)) = m$ if $k_m \geqslant 1$.

FIGURE 4. Values for the length functions (a) M, and (b) N.

Corollary 3. (a) *With the above definition,* $M(S) \equiv N(S) \pmod 2$ *for all* $S \in \{L, R\}^*$.
(b) *Let* $S = R^{k_0}L^{k_1}R^{k_2}\cdots$ *and* $S' = R^{k'_0}L^{k'_1}R^{k'_2}\cdots$ *be finite strings. Then* $r(S) < r(S')$ *if and only if* $k_0 < k'_0$*, or* $k_0 = k'_0$ *and* $k_1 > k'_1$*, or* $(k_0, k_1) = (k'_0, k'_1)$ *and* $k_2 < k'_2$*, or* $(k_0, k_1, k_2) = (k'_0, k'_1, k'_2)$ *and* $k_3 > k'_3$*, …, using an 'alternating lexicographic' ordering.*

Proof. (a) Certainly $M(S) \equiv N(S) \pmod 2$ holds when $S = \varepsilon$, as $M(\varepsilon) = N(\varepsilon) = 0$. Assume $S := S(k_0, \dots, k_m)$ and $M(S) \equiv N(S) \pmod 2$ holds. If $m = M(S)$ is even, then S ends in R^{k_m}, so $SL = S(k_0, \dots, k_m, 1)$ and $M(SL) = M(S) + 1$ is odd. Hence, by induction, and Eq. 5

$$M(SL) = M(S) + 1 \equiv 2M(S) + 1 \equiv 2N(S) + 1 = N(SL) \pmod 2.$$

Similarly, if m is even, then $SR = S(k_0, \dots, k_{m-1}, k_m + 1)$ and $M(SR) = M(S)$. Hence

$$M(SR) = M(S) \equiv 2M(S) + 2 \equiv 2N(S) + 2 = N(SR) \pmod 2.$$

If $M(S) = m$ is odd, then S ends in L^{k_m} and so $M(SL) = M(S)$ and $M(SR) = M(S)+1$ hold. Hence $M(SL) \equiv N(SL) \pmod 2$ and $M(SR) \equiv N(SR) \pmod 2$ both hold.

(b) The formula (6c) for $r(S)$ implies that Eq. (3) holds. Hence moving left decreases the r-value, and moving right increase the r-value. A left move is not counteracted by any number of right moves; nor is a right move counteracted by any number of left moves. This implies that the r-values are ordered via the stated alternating lexicographic ordering. ☐

3. Continued fractions and the Stern-Brocot tree

In this section, we shall consider the tree T_{C} in Fig. 5 whose vertices are continued fractions. Parents of vertices in this tree are related to 'best approximations.' Recall that

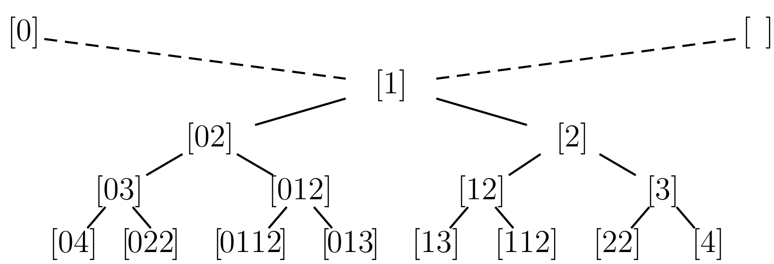

Figure 5. Binary tree T_{C} of continued fractions (with commas omitted).

a continued fraction is an expression of the form

$$[q_0, q_1, \dots, q_m] = q_0 + \cfrac{1}{q_1 + \cfrac{1}{\ddots + \cfrac{1}{q_m}}}.$$

Continued fractions can be computed recursively via the recurrence

$$[q_0] = q_0 \quad \text{and} \quad [q_0, \dots, q_{m-1}, q_m] = [q_0, \dots, q_{m-2}, q_{m-1} + 1/q_m] \quad \text{for } m > 0. \tag{9}$$

The tree T_{C} has root [1], and its children rules are described in Fig. 6 where '$\Diamond$' is an abbreviation for '$q_0, \dots, q_{m-1}$.' A simple induction proves that the continued fractions $[q_0, \dots, q_{m-1}, q_m]$ generated each have $q_0 \geqslant 0$, $q_1, \dots, q_{m-1} \geqslant 1$, and $q_m \geqslant 2$ if $m > 0$.

FIGURE 6. Children rules for the infinite complete binary tree T_{C} in Fig. 5.

(Incidentally, this ensures that when these continued fractions are evaluated using (9) that no denominators of zero are encountered.)

Let $\mathcal{C}$ be the set of vertices (i.e. continued fractions) of the infinite tree T_{C} in Fig. 5. When evaluated using (9), a continued fraction $[q_0, \dots, q_m]$ equals a positive rational $\frac{p}{q}$. Positive rationals have a natural ordering ($\frac{p_1}{q_1} < \frac{p_2}{q_2}$ if and only if $p_1 q_2 < p_2 q_1$), so the continued fractions in $\mathcal{C}$ are naturally ordered. We shall compare this ordering of $\mathcal{C}$ to the ordering of $\{L, R\}^*$ via the function r, see (2). Towards this end, we define a function $f\colon \mathcal{C} \to \{L, R\}^*$ by $f([q_0, \dots, q_{m-1}, q_m]) = S(q_0, \dots, q_{m-1}, q_m - 1)$.

Theorem 4. *Abbreviate $[q_0, \dots, q_{m-1}, q_m] \in \mathcal{C}$ by $[\Diamond, q_m]$ where '$\Diamond$' means '$q_0, \dots, q_{m-1}$'.*

(a) *The LR-location of $[\Diamond, q_m]$ in T_{C} is given by the string $f([\Diamond, q_m]) = S(\Diamond, q_m - 1)$.*
(b) *If $q_0, \dots, q_{m-1}$ is fixed, then $[\Diamond, q_m]$ is an increasing function of q_m if m is even, and a decreasing function of q_m if m is odd.*
(c) *The function $f\colon \mathcal{C} \to \{L, R\}^*$ is a bijection preserving level, children, and order.*
(d) *The parents of $[\Diamond, q_m]$ are the two closest smaller-level approximations to $[\Diamond, q_m]$. That is, of the $2^{q_0 + \cdots + q_m} - 1$ continued fractions $[q'_0, \dots, q'_n]$ with $\sum_{i=1}^n q'_i < \sum_{i=1}^m q_i$, the two closest to $[\Diamond, q_m]$ are the parent continued fractions of $[\Diamond, q_m]$.*

Proof. (a) Define the *length* of $[q_0, \dots, q_m]$ to be $q_0 + \cdots + q_m - 1$. Our proof uses induction on the length of $[q_0, \dots, q_m]$. The LR-location of $[q_0]$ in T_{C} is $R^{q_0 - 1}$, and $f([q_0]) = S(q_0 - 1)$. In particular, the base case of length 0 (when $q_0 = 1$) holds. Suppose now that $q_0 + \cdots + q_m > 1$. The length of the children of $[\Diamond, q_m]$ in Fig. 6 is one more than the length of $[\Diamond, q_m]$. If m is even, then induction gives

$$\begin{aligned} &f([\Diamond, q_m - 1, 2]) = f([\Diamond, q_m])L = S(\Diamond, q_m - 1)L = S(\Diamond, q_m - 1, 1), \quad \text{and} \\ &f([\Diamond, q_m + 1]) = f([\Diamond, q_m])R = S(\Diamond, q_m - 1)R = S(\Diamond, q_m). \end{aligned}$$

If m is odd, then induction gives

$$\begin{aligned} &f([\Diamond, q_m + 1]) = f([\Diamond, q_m])L = S(\Diamond, q_m - 1)L = S(\Diamond, q_m), \quad \text{and} \\ &f([\Diamond, q_m - 1, 2]) = f([\Diamond, q_m])R = S(\Diamond, q_m - 1)R = S(\Diamond, q_m - 1, 1). \end{aligned}$$

Thus the function f gives the LR-location of each continued fraction, as desired.

(b) We use induction on m. Certainly $[q_0] = q_0$ is an increasing function of q_0, and $[q_0, q_1] = q_0 + \frac{1}{q_1}$ is a decreasing function of q_1. Suppose now that $m \geqslant 2$. If m is even, then $[q_0, \dots, q_m] = [q_0, \dots, q_{m-2}, q_{m-1} + \frac{1}{q_m}]$ decreases (by induction) precisely when $q_{m-1} + \frac{1}{q_m}$ increases. Thus $[\Diamond, q_m]$ increases when q_m increases. If m is odd, then $[q_0, \dots, q_m] = [q_0, \dots, q_{m-2}, q_{m-1} + \frac{1}{q_m}]$ increases (by induction) precisely when $q_{m-1} + \frac{1}{q_m}$ increases. Thus $[\Diamond, q_m]$ decreases when q_m increases. This completes the induction.

(c) It is clear that f is surjective since $f([k_0, \dots, k_{m-1}, k_m + 1]) = S(k_0, \dots, k_{m-1}, k_m)$ is a typical LR-string in $\{L, R\}^*$. It is also clear that f is injective, and hence f is bijective. The level of $S(q_0, \dots, q_{m-1}, q_m - 1)$ is $q_0 + \cdots + q_m - 1$. A simple induction shows that the level of $[q_0, \dots, q_m]$ in $T_{\mathtt{C}}$ is also $q_0 + \cdots + q_m - 1$. (This is true for the root $[1]$ of $T_{\mathtt{C}}$. If $[q_0, \dots, q_m]$ has level $q_0 + \cdots + q_m - 1$, then by Fig. 6 its children have level $q_0 + \cdots + q_m$.) Part (a) and the children rules (Fig. 6) show that f preserves children, i.e. $f(C_L(v)) = C_L(f(v))$ and $f(C_R(v)) = C_R(f(v))$ for all $v \in \mathcal{C}$.

It remains to prove that f preserves order. This is true if f^{-1} preserves order. This, in turn, amounts to proving that the alternating lexicographic ordering of strings in Cor. 3(b) is the same as ordering of continued fractions (i.e. of the rational numbers). Suppose that $S := S(k_0, \dots, k_m - 1)$ and $S' := S(k'_0, \dots, k'_n - 1)$ where $r(S) < r(S')$. By Cor. 3(b) there exists an i for which $k_0 = k'_0, \dots, k_{i-1} = k'_{i-1}$, and $k_i < k'_i$ when i is even, and $k_i > k'_i$ when i is odd. We shall prove the rational $v := [k_0, \dots, k_m]$ is less than $v' := [k'_0, \dots, k'_n]$.

Before proving the base case when $i = 0$, we digress to prove $k_0 \leqslant [k_0, \dots, k_m] < k_0 + 1$. This is true when $m = 0$ as $[k_0] = k_0$. We prove that $k_0 < [k_0, \dots, k_m] < k_0 + 1$ holds for $m \geqslant 1$. The proof of this stronger statement uses induction on m. It is true when $m = 1$ as $k_0 < [k_0, k_1] = k_0 + \frac{1}{k_1} \leqslant k_0 + \frac{1}{2}$ because, in our context, $k_m \geqslant 2$ for $m \geqslant 1$. Suppose the stronger statement is true for $m - 1$ where $m > 1$. Then, by induction, $k_1 < [k_1, \dots, k_m] < k_1 + 1$ holds. Since $k_1 \geqslant 1$, taking inverses shows

$$\frac{1}{k_1 + 1} < [0, k_1, \dots, k_m] < \frac{1}{k_1},$$

and adding k_0 implies

$$k_0 < k_0 + \frac{1}{k_1 + 1} < [k_0, k_1, \dots, k_m] < k_0 + \frac{1}{k_1} \leqslant k_0 + 1.$$

Hence $k_0 < [k_0, k_1, \dots, k_m] < k_0 + 1$ holds for $m \geqslant 1$, establishing the digression.

Return now to the base case $i = 0$ of our induction. Certainly $k_0 < k'_0$ implies $v < v'$ when m or n is zero. If m and n are both positive, then $k_0 < k'_0$ implies

$$k_0 \leqslant v < k_0 + 1 \leqslant k'_0 \leqslant v' < k'_0 + 1.$$

In either case, $k_0 < k'_0$ implies $v < v'$. Suppose now $i \geqslant 1$ is even and $k_0 = k'_0, \dots, k_{i-1} = k'_{i-1}$, and $k_i < k'_i$. Abbreviate '$k_0, \dots, k_{i-1}$' by $\Diamond$. Then Eq. (9) implies

$$v = [\Diamond, k_i, \dots, k_m] = [\Diamond, [k_i, \dots, k_m]] \quad \text{and} \quad v' = [\Diamond, k'_i, \dots, k'_n] = [\Diamond, [k'_i, \dots, k'_n]].$$

Hence the base case says $k_i < k'_i$ implies $[k_i, \dots, k_m] < [k'_i, \dots, k'_n]$. Therefore part (b) implies $[\Diamond, [k_i, \dots, k_m]] < [\Diamond, [k'_i, \dots, k'_n]]$ and $v < v'$, as desired. The case when i is odd is proved similarly. This completes the proof that $S < S'$ implies $v < v'$.

(d) This follows from the definition of parent fractions and part (c). □

The continued fractions on the vertices of the tree in Fig. 5 give rise to a tree of rational numbers. This is the well-known Stern-Brocot tree [3, p. 117], which (remarkably) lists every positive rational number (in reduced form) precisely once. To each string $S \in \{L, R\}^*$ there corresponds a *reverse string* $\overline{S}$ defined by $\overline{\varepsilon} = \varepsilon$, $\overline{SL} = L\overline{S}$, and $\overline{SR} = R\overline{S}$. Reversing (or swapping $S \leftrightarrow \overline{S}$) vertices in a tree gives another tree, called the *reverse tree*. The reverse tree of the Stern-Brocot tree is another well-known tree called the Calkin-Wilf tree [1]. (In Fig. 7, $\frac{2}{3} \leftrightarrow \frac{3}{2}$ because $LR \leftrightarrow RL$.) The reader may better understand the connection between parent and children vertices by studying the Stern-Brocot and Calkin-Wilf trees, see [2].

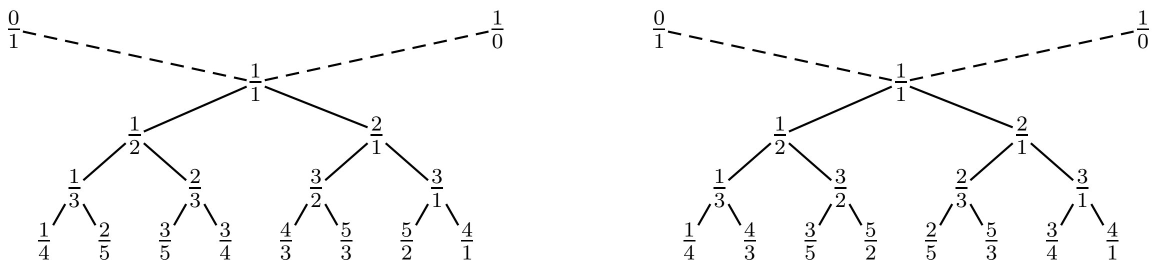

FIGURE 7. (a) Stern-Brocot tree; and the reversed (b) Calkin-Wilf tree.

There are further applications of parent vertices to a complete binary tree associated with the Cantor[2] set, however, exploring these goes beyond the scope of this note.

Acknowledgement: The author is grateful to the referee for his/her suggestions.

References

[1] Calkin, N, and Wilf, H. S (2000). Recounting the rationals, *Amer. Math. Monthly* **107**, 360–363.
[2] Glasby, S. P. (2011). Enumerating the rationals from left to right, *Amer. Math. Monthly* **118**, 830-835.
[3] Graham, R. L, Knuth, D. E, and Patashnik, O (1994). *Concrete Mathematics*, 2nd ed., Addison-Wesley, Reading, MA.
[4] Hawkins, T (1970). *Lebesgue's Theory of Integration. Its origins and development.*, University of Wisconsin Press, Madison, WI.
[5] Knuth, D. E (1997). *The Art of Computer Programming, vol. 1 Fundamental Algorithms*, 3rd ed., Addison-Wesley, Reading, MA.
[6] Knuth, D. E (1997). *The Art of Computer Programming, vol. 2 Seminumerical Algorithms*, 3rd ed., Addison-Wesley, Reading, MA.
[7] Schroeder, C (2006). Metric tree weight adjustment and infinite complete binary trees as groups, MSc Thesis, Drexel University.
[8] Serre, J.-P (1980). *Trees* (Translated from the French by John Stillwell), Springer-Verlag.
[9] Smith, H. J. S (1874). On the Integration of discontinuous functions, *Proc. London Math. Soc.* **s1-6**, 140–153. doi: 10.1112/plms/s1-6.1.140

[2]The original discoverer appears to be H.J.S. Smith [9, p. 147], see [4] for details.

(Glasby) Department of Mathematics, Central Washington University, WA 98926, USA. Also affiliated with The Faculty of Information Sciences and Engineering, University of Canberra and The Centre for Mathematics of Symmetry and Computation at the University of Western Australia. GlasbyS@gmail.com, http://www.cwu.edu/∼glasbys/