Journal of Universal Computer Science, vol. 1, no. 7 (1995), 484-503
submitted: 14/6/95, accepted: 14/6/95, appeared: 28/7/95©Springer Pub. Co.

LCF: A Lexicographic Binary Representation of the Rationals

Peter Kornerup
(Dept. of Mathematics and Computer Science
Odense University
DK-5230 Odense, Denmark
kornerup@imada.ou.dk)

David W. Matula
(Dept. of Computer Science and Engineering
Southern Methodist University
Dallas, TX 75275
matula@seas.smu.edu)

Abstract: A binary representation of the rationals derived from their continued fraction expansions is described and analysed. The concepts "adjacency", "mediant" and "convergent" from the literature on Farey fractions and continued fractions are suitably extended to provide a foundation for this new binary representation system. Worst case representation-induced precision loss for any real number by a fixed length representable number of the system is shown to be at most 19% of bit word length, with no precision loss whatsoever induced in the representation of any reasonably sized rational number. The representation is supported by a computer arithmetic system implementing exact rational and approximate real computations in an on-line fashion.
Category: G.1.0 [Numerical Analysis]: Computer Arithmetic. B.5.1 [Register Transfer Level Implementation]: Arithmetic and logic units. E.2 [Data Storage Representations].
Key Words: Computer arithmetic, continued fractions, lexicographic, number systems, number theory, rational numbers.

1 Introduction.

The foundations of a binary representation of the rationals are presented, and many of the representation system's features are described. Evidence is provided indicating that a computer arithmetic system employing this representation would provide a facility for exact rational and approximate real arithmetic not currently available in any single system.

Our proposed binary representation system derives from the continued fraction representation of the rationals. A self delimiting bitstring encoding of the integers is employed to represent each partial quotient. Particular features of the integer encoding and the subsequent concatenation process allow us to obtain bit string representations of the rationals, which are shown lexicographically order preserving over real order. Our bitstring representation is thus termed the lexicographic continued fraction (LCF) representation of a rational number.

The LCF representation can be considered an encoding of the individual steps of the Euclidean algorithm performed in binary, where the determination of the individual remainders are computed using a non-restoring division algorithm. As such, it is derived from algorithms performing arithmetic operations upon rational operands in fraction form, i.e. a numerator/denominator representation, [see Kornerup and Matula 83] where the LCF representation was first described. However, as a number representation it naturally leads to a kind of on-line arithmetic where operands are consumed bit-sequential, and the result is produced bit-sequential, most significant bit first. Such an on-line arithmetic unit has been described in [Kornerup and Matula 88], capable of performing all the basic arithmetic operations in a unified manner as cases of the bihomographic function

$$z(x,y) = \frac{axy + bx + cy + d}{exy + fx + gy + h}$$

specified by eight integer coefficients $a, b, \cdots, h$. By factoring certain transformations (matrices corresponding to the individual partial quotients of a continued fraction) into simple "binary" matrices, the algorithm can be realized by simple shift-and-add operations. However, we shall not further pursue the arithmetic here, but concentrate on properties of the LCF representation.

In [Section 2] we formally define the LCF expansion as a bitstring. We introduce background material from the theory of continued fractions to guide the development of a theory for LCF expansions. In particular we extend the notion of the sequence of convergents (often termed "best rational approximations") of a real number to a super-sequence of biconvergents (binary convergents) determined by the LCF expansion of x. The biconvergents are shown to form a somewhat base dependent sequence of rational approximations to x. The biconvergent sequence is shown to contain, on average, about 3.51 times the number of terms of the subsequence of canonically defined convergents.

In [Section 3] we study the hierarchy of rational numbers as determined bitwise by their LCF expansions through the construct of the LCF binary tree. The LCF tree provides for enumerating biconvergent sequences as paths down the tree, and also provides for enumerating all fixed length LCF expansion values by traversal across the LCF tree truncated at fixed depth. One can visualize in the structure of the tree the order preserving property of LCF expansions, and the fact that LCF representation is one-to-one between finite bit strings and positive rationals. We develop tools for investigating the set Q_k of irreducible fractions in [0,1] whose LCF expansions have order k (equivalently: length $k+1$ bits or depth at most k in the LCF tree). Our principal results are that the fundamental properties from the theory of Farey fractions [Hardy and Wright 79] regarding adjacency, mediant and recursive construction of the tree of Farey fractions, can be extended to comparable concepts of bijacency, binary mediant and recursive construction of the LCF tree. Properties of the LCF tree are then available as tools for both the investigation of the rate of convergence of biconvergent sequences, and for the study of the gap sizes between successive members of the sets Q_k. The latter result dictates the precision obtainable for arithmetic employing such fixed length representations.

Utilizing these tools the extremes of gap size variability over Q_k are then discussed in [Section 4]. The main result is that the maximum gap size in Q_k is of the order 2^{-ak} for $a = 0.814\ldots$. This implies at most a 19% precision loss (storage space loss) in the worst case approximation error by fixed length LCF bit strings, being the price to be able to accommodate the exact representation of a set of simple rationals at a fairly regular spacing, and supporting a rational arithmetic in an on-line fashion. More detailed results of exhaustive and sampled distribution of gap sizes are available in [Kornerup and Matula 85].

2 Continued Fraction and Lexicographic Continued Fraction Expansions.

The lexicographic continued fraction expansion of a rational number is a bitstring whose interpretation will be based on some fundamental properties of continued fraction expansions of rationals. For these purposes our notation should make explicit the particular numerator and denominator components of a fraction, as well as the particular sequence of partial quotient values of a continued fraction expansion, as these terms are not necessarily uniquely determined. For our purposes it is sufficient to treat only representation of finite nonnegative rational numbers, as signs can be appended for the negative values.

Formally, a fraction, denoted p/q or $\frac{p}{q}$, is herein an ordered pair composed of a nonnegative integer *numerator* p, and a positive integer *denominator* q. The quotient of p/q is the rational number determined by the ratio of p to q. The numerator and denominator of an *irreducible* fraction must have a greatest common divisor (gcd) of unity, other fractions being termed *reducible*.

Employing the equality symbol between various forms of rational representation will herein denote the weaker interpretation of equality between their rational values with the following exception. Equality between fractions denoted with the horizontal bar format shall imply equal numerator and denominator values. Thus

$$\frac{p}{q} = \frac{r}{s} \text{ iff } p = r \text{ and } q = s,$$

whereas

$$p/q = r/s \text{ iff } qr = ps.$$

Notationally, the symbol $\preccurlyeq$ is used to denote the *simpler than* relation between fractions, and is defined over all pairs of fractions by

$$\frac{p}{q} \preccurlyeq \frac{r}{s} \text{ iff } \frac{p}{q} \neq \frac{r}{s} \text{ and both } p \leq r, q \leq s.$$

For example $\frac{0}{1} \preccurlyeq \frac{1}{1}$, and $\frac{1}{2} \preccurlyeq \frac{2}{3} \preccurlyeq \frac{2}{4}$.

We utilize the notation $[a_0/a_1/a_2/\cdots/a_n]$ for the *n-th order* (simple) *continued fraction* expansion

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

where the *partial quotients* a_i are assumed to be integral with $a_0 \geq 0$, $a_i \geq 1$ for $1 \leq i \leq n$. It is known from the theory of continued fractions [Khinchin 35, Hardy and Wright 79] that any positive rational number, denoted by the irreducible fraction $\frac{p}{q}$, has exactly two finite expansions herein termed the *canonical* and *long expansions*, as given and related by:

$$\frac{p}{q} = \begin{cases} [a_0/a_1/\cdots/a_{n-1}/a_n] & \textit{canonical} \\ [a_0/a_1/\cdots/a_{n-1}/a_n - 1/1] & \textit{long} \end{cases} \quad \text{where } a_n \geq \begin{cases} 2 \text{ for } n \geq 1 \\ 1 \text{ for } n = 0 \end{cases}$$

with 0 having the unique and canonical expansion [0]. It follows that any positive rational number has both a unique *even order* continued fraction expansion $[a_0/a_1/\cdots/a_{2m}]$ together with a unique *odd order* continued fraction expansion $[a_0/a_1/\cdots/a_{2m+1}]$. It is the unique even order expansion $[a_0/a_1/\cdots/a_{2m}]$ that will later be employed for the definition of the lexicographic continued fraction expansion.

The irreducible fractions $\frac{p_i}{q_i} = [a_0/a_1/\cdots/a_i]$ determined for $0 \le i \le n$ by truncating the continued fraction $[a_0/a_1/\cdots/a_n]$ constitute a sequence of irreducible fraction approximations to $\frac{p}{q} = \frac{p_n}{q_n} = [a_0/a_1/\cdots/a_n]$ termed *convergents* to $\frac{p}{q}$. The *preconvergent* shall denote the convergent $\frac{p_{n-1}}{q_{n-1}}$ immediately preceding $\frac{p_n}{q_n}$ in the canonical expansion for $\frac{p}{q}$. Note that the long expansion of $\frac{p}{q} \neq 0$ includes one additional convergent termed the *parent* in addition to those for the canonical expansion of $\frac{p}{q}$. For example, the canonical expansion $\frac{277}{642} = \frac{p}{q} = [0/2/3/6/1/3/3]$ has the convergents $\frac{0}{1}, \frac{1}{2}, \frac{3}{7}, \frac{19}{44}, \frac{22}{51}, \frac{85}{197}, \frac{277}{642}$ with preconvergent $\frac{85}{197}$. The parent of $\frac{277}{642}$ is then $\frac{192}{445} = [0/2/3/6/1/3/2]$.

The convergents have many important properties, some of which are cited here for reference from [Hardy and Wright 79, Khinchin 35]:

Theorem 1. *The convergents $\frac{p_i}{q_i} = [a_0/a_1/\cdots/a_i]$ of any (canonical or long) continued fraction $\frac{p}{q} = [a_1/a_2/\cdots/a_n]$ for $i = 0, 1, \cdots, n$ satisfy the following properties:*

i) *Recursive ancestry:*
With $p_{-2} = 0, p_{-1} = 1, q_{-2} = 1$, and $q_{-1} = 0$,

$$p_i = a_i p_{i-1} + p_{i-2},$$
$$q_i = a_i q_{i-1} + q_{i-2},$$

or in matrix form:

$$\begin{Bmatrix} p_{i-2} & p_{i-1} \\ q_{i-2} & q_{i-1} \end{Bmatrix} \begin{Bmatrix} 0 & 1 \\ 1 & a_i \end{Bmatrix} = \begin{Bmatrix} p_{i-1} & p_i \\ q_{i-1} & q_i \end{Bmatrix} \text{ and } \begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix} \prod_{j=0}^{i} \begin{Bmatrix} 0 & 1 \\ 1 & a_i \end{Bmatrix} = \begin{Bmatrix} p_{i-1} & p_i \\ q_{i-1} & q_i \end{Bmatrix},$$

ii) *Irreducibility:*

$$gcd(p_i, q_i) = 1,$$

iii) *Adjacency:*

$$q_i p_{i-1} - p_i q_{i-1} = (-1)^i,$$

iv) *Simplicity:*

$$\frac{p_i}{q_i} \lessgtr \frac{p_{i+1}}{q_{i+1}} \text{ for } i \le n-1,$$

v) *Alternating convergence:*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots \frac{p_{2i}}{q_{2i}} < \cdots \le \frac{p}{q} \le \cdots < \frac{p_{2i-1}}{q_{2i-1}} < \cdots < \frac{p_1}{q_1},$$

vi) *Best rational approximation:*

$$\frac{r}{s} \lll \frac{p_i}{q_i} \Longrightarrow \left|\frac{r}{s} - \frac{p}{q}\right| > \left|\frac{p_i}{q_i} - \frac{p}{q}\right|,$$

vii) *Quadratic convergence:*

$$\frac{1}{q_i(q_{i+1} + q_i)} < \left|\frac{p_i}{q_i} - \frac{p}{q}\right| \leq \frac{1}{q_i q_{i+1}} \text{ for } i \leq n-1,$$

viii) *Real approximation:*

$\left|x - \frac{p}{q}\right| < \frac{1}{2q^2}$ *for irreducible* $\frac{p}{q}$ *implies that* $\frac{p}{q}$ *is a convergent of a (possibly infinite) continued fraction expansion of* x.

From Theorem 1(v) we see that the even order convergents approach $\frac{p}{q}$ from below and the odd order convergents approach from above, with the interval between any two successive convergents containing $\frac{p}{q}$.

[Tab. 1] illustrates the rate at which the sequence of convergent values for $\frac{277}{642} = [0/2/3/6/1/3/3]$ give better approximations to 277/642. Note from the table that the accuracy attained by successive convergents has a larger incremental improvement when the next partial quotient is large, as anticipated by Theorem 1 (vii).

Continued fraction	*Fraction*	*Decimal representation*	*Relative error*
[0]	0/1	$0.0\cdots$	1
[0/1]	1/2	$0.50\cdots$	0.15
[0/2/3]	3/7	$0.428\cdots$	0.0067
[0/2/3/6]	19/44	$0.4318\cdots$	0.00082
[0/2/3/6/1]	22/51	$0.43137\cdots$	0.00021
[0/2/3/6/1/3]	85/197	$0.431472\cdots$	0.000018
[0/2/3/6/1/3/3]	277/642	$0.4314641\cdots$	0

Table 1: The canonical convergents to $\frac{277}{642}$ in continued fraction, fraction, and decimal form, and the relative errors of these convergents as best rational approximations.

Consider from Theorem 1 and the example of [Tab. 1] that the notion of an i'th order best rational approximation by itself is not useful in finite precision computational practice as the resulting accuracy depends without bound on the size of the particular partial quotients involved. The lexicographic continued fraction expansion we now introduce will be shown quite analogously to identify a sequence of rational approximations termed "biconvergents". The biconvergents contain and supplement the sequence of convergents of the canonical continued fraction suitably granularized at the bit level to allow a measure of accuracy in terms of bit length.

For the purpose of defining a binary based lexicographic continued fraction expansion, we will employ a binary representation of the positive integers which is both "self delimiting", i.e. which implicitly contains an end-marker when read from left to right, and lexicographically order preserving over the integers.

Formally, if the integer $p \geq 1$ has the $(n+1)$-bit binary radix representation $1b_{n-1}\cdots b_1b_0$, with $n \geq 0$ and $\circ$ denoting string concatenation, the $(2n+1)$-bit bitstring

$$\ell(p) \equiv 1^n \circ 0 \circ b_{n-1}b_{n-2}\cdots b_1b_0 \tag{1}$$

will be termed the *lexibinary* form of p. $\ell(p)$ is thus composed on a (possibly vacuous) *unary part* 1^n delimited by the *switch-bit* 0, followed by the (possibly vacuous) *binary part* $b_{n-1}b_{n-2}\cdots b_0$. The value of the lexibinary integer $1^n0b_{n-1}b_{n-2}\cdots b_0$ is then $2^n + \sum_{i=0}^{n-1} b_i2^i$. This representation is *order preserving* in that the lexicographic ordering (leftmost bit first) of lexibinary bitstrings is seen to correspond to the numeric ordering of their values. [Tab. 2] illustrates the lexibinary form of several integers. For a discussion of alternative lexicographic order preserving binary encodings of the integers [see Knuth 82], where similar representations are analysed.

Integer	*Binary*	*Lexibinary*
1	1.	0
2	10.	100
3	11.	101
4	100.	11000
5	101.	11001
6	110.	11010
7	111.	11011
8	1000.	1110000
16	10000.	111100000
32	100000.	11111000000
100	1100100.	1111110100100
200	11001000.	111111101001000
1000	1111101000.	1111111110111101000

Table 2: Right-adjusted standard binary representation and left-adjusted lexibinary bitstring representation of certain integers.

Note from the definition of a continued fraction expansion that $[a_0/a_1/\cdots/a_n]$ is an increasing function of any even order partial quotient, and a decreasing function of any odd order partial quotient. Thus to obtain an order preserving representation of the rationals we simply represent the odd order quotients in complemented lexibinary integer form before concatenation. To be able to compare bit strings lexicographically from left to right it is assumed that any (finite length) representation is extended to the right with an arbitrary number of extra zeroes. This corresponds with the observation that, suitably interpreted and extending the even order continued fraction expansion

$$p/q = [a_0/a_1/\cdots/a_{2m}] = [a_0/a_1/\cdots/a_{2m}/\infty]. \tag{2}$$

If we define $\ell(\infty)$ to be an infinite string of ones, then since ∞ occurs in an odd order position in (2), $\ell(\infty)$ will always appear in complemented form yielding $\overline{\ell(\infty)} = 00\cdots$. This provides a terminal infinite string of zeroes, which may either be denoted by 0^∞ or taken as assumed.

Formally, implicitly handling the case $a_0 = 0$ ($0 \le p/q < 1$) by a leading zero bit and $a_0 \ge 1$ ($1 \le \frac{p}{q}$) by a leading unit bit, the *lexicographic continued fraction* (LCF) expansion of $\frac{p}{q} \ge 0$ is the (infinite) bitstring determined employing (1) and (2) by:

$$LCF\left(\frac{p}{q}\right) = \begin{cases} 1 \circ \ell(a_0) \circ \overline{\ell(a_1)} \circ \cdots \circ \ell(a_{2m}) \circ \overline{\ell(\infty)} & \text{for } 1 \le \frac{p}{q} \\ 0 \circ \overline{\ell(a_1)} \circ \cdots \circ \overline{\ell(a_{2m-1})} \circ \ell(a_{2m}) \circ \overline{\ell(\infty)} & \text{for } 0 \le \frac{p}{q} < 1 \end{cases} \tag{3}$$

The positive valued LCF expansion $b_0 b_1 \cdots b_{k-1} 1 0^\infty$ is said to have *order k* (the index of the least significant unit), with $\frac{0}{1} = 0^\infty$ having order zero. The LCF expansion denoted with the concatenation symbol $\circ$ at each corresponding partial quotient boundary as in (3) is said to be in *parsed form.* The leading bit b_0 of the LCF expansion $b_0 b_1 \cdots b_{k-1} 1 0^\infty$ is termed the *reciprocal* bit. Note that all subsequent bits may uniquely be identified as members of the unary, switch, or binary portions of the i'th order partial quotient in the even order continued fraction expansion in (3).

The finite bit string $b_0 b_1 \cdots b_n$, with or without trailing zeros, is taken as an alternative finite LCF expansion equivalent to $b_0 b_1 \cdots b_n 0^\infty$. The *minimal* LCF expansion $b_0 b_1 \cdots b_{k-1} 1$ of $\frac{p}{q} > 0$ is truncated at the last unit bit, and thus including the reciprocal bit has length one greater than the order of the LCF expansion. Zero is taken to have its minimal LCF expansion composed of the single 0 reciprocal bit.

Example 1.

$$\begin{aligned} \tfrac{22}{7} &= \tfrac{22}{7} && \text{irreducible fraction form} \\ &= [3/6/1] && \text{even order continued fraction form} \\ &= [3/6/1/\infty] && \text{infinite extension} \\ &\left.\begin{aligned} &= 1 \circ \ell(3) \circ \overline{\ell(6)} \circ \ell(1) \circ \overline{\ell(\infty)} \\ &= 1 \circ 101 \circ \overline{11010} \circ 0 \circ \overline{1^\infty} \\ &= 1 \circ 101 \circ 00101 \circ 0 \circ 0^\infty \end{aligned}\right\} && \text{equivalent parsed LCF expansions} \\ &= 110100101 && \text{minimal LCF expansion (order 8)} \end{aligned}$$

The irreducible fraction $\frac{r_k}{s_k} = b_0 b_1 \cdots b_{k-1} 1$ for $0 \le k \le n$ determined by truncating the lexicographic continued fraction $\frac{r}{s} = \frac{r_n}{s_n} = b_0 b_1 \cdots b_{n-1} 1$ at index $k-1$ and appending a unit bit, is termed the k'th order *biconvergent* (*binary convergent*) of $\frac{r}{s}$. Each biconvergent $\frac{r_0}{s_0}, \frac{r_1}{s_1} \cdots \frac{r_n}{s_n}$ in sequence provides then either an improved upper or lower bound on $\frac{r}{s}$ determined by

$$\frac{r_k}{s_k} = b_0 b_1 \cdots b_{k-1} 1 \begin{cases} \ge \frac{r}{s} \text{ if } b_k = 0, \\ \le \frac{r}{s} \text{ if } b_k = 1. \end{cases}$$

To compare biconvergent approximation with the k-bit binary radix approximation, we can compute the "precision" of the k'th order biconvergent approximation in bits by the negative base two logarithm of the bounding intervals, as illustrated by the following table of biconvergents of $\frac{277}{642}$ through order 12:

k	0	1	2	3	4	5	6	7	8	9	10	11	12
r_k/s_k	1/1	1/2	1/4	1/3	2/5	4/9	3/7	7/16	13/30	25/58	19/44	22/51	41/95
$-\log_2(gap)$	0.00	1.00	2.00	2.58	3.32	4.49	5.97	6.80	7.71	8.76	10.31	11.13	12.24

Some useful facts about biconvergents follow from the definition and certain properties of convergents, and we will summarize these in some observations. Noting that the reciprocal of $\frac{p}{q} = [a_0/a_1/\cdots/a_n] \geq 1$ is $\frac{q}{p} = [0/a_0/a_1/\cdots/a_n]$, so then the convergents to $\frac{p}{q}$ are $\frac{0}{1}$ and the reciprocals of the convergents to $\frac{p}{q}$. From (3) then

Observation 2. *The LCF expansion of the reciprocal of* $\frac{p}{q} = b_0 b_1 \cdots b_{k-1} 1$ *is the 2's complement,* $\frac{q}{p} = \overline{b_0 b_1 \cdots b_{k-1}} 1$. *Thus the reciprocal of a k'th order LCF number is also of k'th order.*

Observation 3. *The biconvergents to the reciprocal* $\frac{q}{p} = \overline{b_0 b_1 \cdots b_{k-1}} 1$ *of* $\frac{p}{q} = b_0 b_1 b_{k-1} 1$ *are the reciprocals of the biconvergents to* $\frac{p}{q}$.

It is immediate from (3) that the even order convergents of the canonical continued fraction for $\frac{p}{q}$ are also biconvergents to $\frac{p}{q}$. The odd order convergents to $\frac{p}{q}$ have reciprocals that are even order convergents to $\frac{q}{p}$ and thus are biconvergents to $\frac{q}{p}$. Using [Observations 2,3] it follows that

Observation 4. *Every convergent of the canonical continued fraction expansion of* $\frac{p}{q}$ *is also a biconvergent to* $\frac{p}{q}$.

Although the order of the lexicographic continued fraction can be arbitrarily large compared to the order of the ordinary continued fraction for a given rational, on the average the orders can be related. From classical material on continued fractions it is known that the partial quotients in the continued fraction expansion of a randomly chosen $\frac{r}{s} \in [0,1]$ (see [Knuth 81] or [Blachman 84] for details) will have value i with probability essentially given by

$$p_i = \log_2 \left(1 + \frac{1}{i(i+2)}\right), \tag{4}$$

where then

$$p_1 = 0.415,\ p_2 = 0.170,\ p_3 = 0.093,\ p_4 = 0.059, \cdots.$$

With the distribution of partial quotient size given by (4), we note that 41.5% of all partial quotients are unity and are encoded by a single bit in the LCF expansion. Another 26.3% of the partial quotients have values two or three and contribute 3 bits each to the LCF expansion, and an average partial quotient from (4) has expected length

$$\sum_i (2\lfloor \log_2 i \rfloor + 1) \log_2 \left(1 + \frac{1}{i(i+2)}\right) = 3.51 \cdots.$$

Thus in summary,

Observation 5. *From the known distribution (4) of partial quotient size, it follows that the canonical continued fraction expansion and LCF expansion of a rational* $\frac{p}{q} = [a_1/a_2/\cdots/a_n] = b_0 b_1 \cdots b_{k-1}1$ *yield an expected biconvergent to convergent ratio of* $Exp(\frac{k}{n}) = 3.51\cdots$.

Observation 6. *From the known distribution (4) of partial quotient size, it follows that* $1/3.51\cdots = 28.5\%$ *of the biconvergents to* $\frac{p}{q}$ *will also be convergents to* $\frac{p}{q}$*, i.e. the so-called best rational approximations which are characterized without any dependence on the binary representation employed for the LCF expansion.*

LCF expansions and biconvergent approximations have many properties of theoretical interest and/or which find use in employing LCF representation as a basis of computer arithmetic unit design. We shall particularly pursue herein issues related to assessing the accuracy of finite precision computation employing fixed length LCF expansions. A summary of related topics [see Kornerup and Matula 83, 88] that will not be further pursued in this paper are listed here for reference:

Regarding uniqueness of LCF representation of the nonnegative rationals:

- There is a one-to-one correspondence between all minimal LCF expansions and the nonnegative rational numbers.

Regarding arithmetic with LCF represented numbers:

- LCF expansions may be used bit-by-bit in a left-to-right scan as input to on-line algorithms for the direct computation of arithmetic expressions upon LCF operands yielding LCF results [Kornerup and Matula 88]. In this context the LCF bitstring may be interpreted as an encoding of the individual steps (transitions) in a finite automaton performing the Euclidean gcd algorithm (in binary) on p and q.

Regarding the efficiency in bit-length of LCF expansions:

- The minimum redundancy encoding, Huffman encoding, and LCF expansion average bit-length per partial quotient can be computed employing classical results [Knuth 81] on the distribution of the size (4) of partial quotients, yielding:

Average Bits per Partial Quotient	
Minimum redundancy encoding	3.43... ,
Huffman encoding	3.47... ,
LCF expansion	3.51... .

Thus the computationally useful format of the LCF expansion is achieved with an encoding length only about 2% greater than that which could be obtained by any minimal redundancy encoding, and only about 1% greater than that achievable by a Huffman encoding, where the latter two encodings would most likely be of no practical value for arithmetic computation.

Regarding multiplicative and additive inverses:

– LCF representation may be extended with a sign as follows:

$$SLCF\left(\frac{p}{q}\right) = \begin{cases} 1 \circ LCF(\frac{p}{q}) \text{ for } \frac{p}{q} \geq 0, \\ 0 \circ \widetilde{LCF}(\frac{p}{q}) \text{ for } \frac{p}{q} < 0. \end{cases}$$

where $\widetilde{\mathrm{LCF}}$ denotes the 2's complement of any minimal finite LCF bitstring, and the 1's complement of any infinite bitstring. SLCF representation is then order preserving over the reals. Note then that SLCF representation has leading sign and reciprocal bits that treat both the additive and multiplicative inverses of real numbers in an analogous manner.

3 The LCF Rational Number Hierarchy

A hierarchy is imposed on the rationals by the order (length) of their LCF expansions. The enumeration of this rational hierarchy is conveniently illustrated by associating the positive irreducible fractions with the nodes of an infinite binary tree, termed the LCF tree, where the LCF bitstring denotes the path to the node containing the associated irreducible fraction. The fraction $\frac{p}{q} = b_0 b_1 \cdots b_{k-1} 1$ of order k is assigned to the node at depth k reached by proceeding to the left child when $b_i = 0$ and to the right child when $b_i = 1$ for $i = 0, 1, 2, \cdots, k-1$.

The left half of the LCF tree truncated at depth 5 is illustrated in [Fig. 1], where we note that the values in the nodes of the right half of the LCF tree are simply the reciprocals of those in the left half reached by the complemented bitstring.

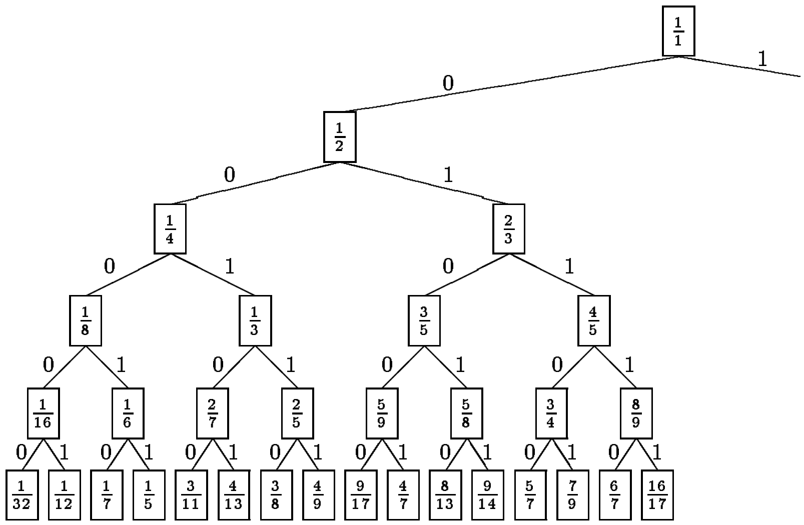

Figure 1: The left half of the LCF tree through depth five.

The LCF tree provides a convenient reference for interpreting the accuracy of finite precision LCF representation both "vertically" and "horizontally":

Vertically: The path from the root down to any node enumerates the biconvergents to the irreducible fraction at that node, e.g. $\frac{1}{1}, \frac{1}{2}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{4}{9}$ is the sequence of biconvergents to $\frac{4}{9}$.

Horizontally: The (nonzero) elements of the set Q_k of irreducible fractions of $[0,1]$ whose LCF expansions have order at most k can be enumerated by an inorder traversal of the left half of the LCF tree truncated at depth k, e.g. including $\frac{0}{1}, \frac{1}{1}$:

$$Q_2 = \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1} \right\},$$

and

$$Q_4 = \left\{ \frac{0}{1}, \frac{1}{16}, \frac{1}{8}, \frac{1}{6}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{5}{9}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{8}{9}, \frac{1}{1}, \right\}.$$

There are then $2^k + 1$ members of Q_k in the interval $[0,1]$ yielding an average gap size of 2^{-k}. The definition of LCF representation provides no immediate clue as to the extent of variation of these gap sizes for fixed k, knowledge of which is essential to assess the precision obtainable by k-bit LCF approximation. In this section we develop tools for investigating gap sizes in Q_k and their relation to gap sizes in Q_{k+1}. For purposes of analysis we shall be particularly concerned with characterizing relations between LCF represented numbers in terms of their more familiar irreducible fraction and continued fraction representations.

For the fractions $p/q < r/s$, the size of the interval $[p/q, r/s]$ is given by the expression $(rq - ps)/qs$ and will be a minimum relative to the size of the denominator qs when $|rq - ps| = 1$. We say the fractions $\frac{p}{q}$ and $\frac{r}{s}$ are *adjacent* whenever $|rq - ps| = 1$. For $\frac{p}{q}$ adjacent to $\frac{r}{s}$ it follows immediately that both $\frac{p}{q}$ and $\frac{r}{s}$ are irreducible, and that either $\frac{p}{q} \lessdot \frac{r}{s}$ or $\frac{r}{s} \lessdot \frac{p}{q}$. Note that some, but not all, successive pairs of fractions of Q_k are adjacent.

The notion of adjacency also identifies an important relation among continued fractions that will form a bridge to understanding the neighbor relations for members of Q_k. The following theorem provides alternative characterizations.

Theorem 7. *For the fractions $\frac{p}{q} \lessdot \frac{r}{s}$ (i.e. $\frac{p}{q} \neq \frac{r}{s}$ and $p \leq r$, $q \leq s$) each of the following four properties implies the other three and serves as an equivalent definition of adjacency.*

i) *Determinant form: $|rq - ps| = 1$,*

ii) *Interval form: $\frac{p}{q}$, $\frac{r}{s}$ are both irreducible and both simpler than any other fraction within the interval bounded by $\frac{p}{q}$ and $\frac{r}{s}$*

iii) *Continued fraction form: $\frac{p}{q}$ and $\frac{r}{s}$ are both irreducible and related by*

$$\frac{r}{s} = [a_0/a_1/\cdots/a_{n-1}/a_n] \qquad \textit{(canonical form of } \tfrac{r}{s}\textit{)}$$

and

$$\frac{p}{q} = \begin{cases} [a_0/a_1/\cdots/a_{n-1}] & \textit{(preconvergent of } \tfrac{r}{s} \textit{ for } n \geq 1\textit{)} \\ \textit{or} & \\ [a_0/a_1/\cdots/a_{n-1}/a_n - 1] & \textit{(parent of } \tfrac{r}{s} \textit{ for any } n\textit{).} \end{cases}$$

iv) *Convergent form: $\frac{r}{s}$ is irreducible and $\frac{p}{q}$ is either its preconvergent or parent.*

Proof. The equivalence of (i) and (ii) is known from the classical theory of Farey Fractions, e.g. [Hardy and Wright 79]. The fact that (iii) implies (i) follows from Theorem 1. (iv) is essentially a restatement of (iii). To complete the proof we need only show that the two alternatives given for $\frac{p}{q}$ in (iii) yield the only fractions simpler than and adjacent to $\frac{r}{s}$.

Let $\frac{r_i}{s_i} = [a_0/a_1/\cdots/a_i]$ for $i = 0, 1, \cdots, n$ denote the convergents to the canonical continued fraction for $\frac{r}{s}$. The set of linear equations

$$xr + yp = r_{n-1}$$

$$xs + yq = s_{n-1}$$

will have a unique solution (x, y) since $|ps - qr| = 1$. Note that $|y| = |y||ps - qr| = |r_{n-1}s - s_{n-1}r| = 1$. So $y = \pm 1$, and x must be integral since $\frac{r}{s}$ is irreducible.

For $y = 1$, $0 \leq p = r_{n-1} - xr$, with $r_{n-1} < r$, implies $x = 0$, hence

$$\frac{p}{q} = \frac{r_{n-1}}{s_{n-1}} = [a_0/a_1/\cdots/a_{n-1}].$$

For the case $y = -1$ then $p = xr - r_{n-1}$ implies $x = 1$ since $p \leq r$. Hence $p = x(a_n r_{n-1} + r_{n-2}) - r_{n-1} = (a_n - 1)r_{n-1} + r_{n-2}$ and similarly we find $q = (a_n - 1)s_{n-1} + s_{n-2}$, thus

$$\frac{p}{q} = [a_0/a_1/\cdots/a_n - 1].$$

$\square$

From the LCF tree it is clear that membership of a fraction in Q_k depends in some manner on a truncated binary representation of the final partial quotient of a continued fraction representation of the fraction. This introduces a base dependence phenomenon similar to that obtained in finite length binary radix representation. We seek both to model and understand the ramifications of this base dependency in LCF representation. Importantly, the adjacency relation itself as interpreted on fractions and/or continued fractions suffers no base dependence on the representation of the individual partial quotients. The following extension of adjacency introduces a dependency on the binary representation only in the last partial quotient. This "binary adjacency" relation will be shown sufficient to characterize all neighbor pairs in Q_k.

We first note that two finite bitstrings α, β are termed *lexicographically adjacent* (with α *lexicographically preceding* β) if

$$\alpha = \sigma \circ 0 \circ 1^j$$

$$\beta = \sigma \circ 1 \circ 0^j$$

for some prefix σ and some $j \geq 0$. Lexicographically adjacent bitstrings thus may equivalently be said to differ by a unit in the last place (ulp). This is precisely the relation between neighbors in Q_k we want to express in terms of equivalent relations among rationals in fraction or continued fraction form. However, it turns out to be most convenient to define the relation wanted in terms of continued fraction expansions.

The fractions $\frac{p}{q} \lessdot \frac{r}{s}$ are *bijacent* (of positive, zero or negative degree i) iff $\frac{p}{q}$ and $\frac{r}{s}$ are both irreducible, and

$$\begin{aligned} \frac{p}{q} &= [a_0/a_1/\cdots/a_n] \\ \frac{r}{s} &= [a_0/a_1/\cdots/a_n + 2^i], \end{aligned} \tag{5}$$

where $a_n = k2^i \geq 2$ when $i \geq 1$.

The definition includes situations corresponding to i being positive, zero, or negative which may be separately interpreted as follows:

$i \geq 0$: $\frac{p}{q}$ and $\frac{r}{s}$ have canonical expansions differing only in the last partial quotient

$$\begin{aligned} \frac{p}{q} &= [a_0/a_1/\cdots/a_{n-1}/k2^i] \\ \frac{r}{s} &= [a_0/a_1/\cdots/a_{n-1}/(k+1)2^i]. \end{aligned}$$

$i \leq 0$: In this case we interpret $a_n + 2^i$ as two partial quotients

$$\frac{r}{s} = [a_0/a_1/\cdots/a_n + 2^i] = [a_0/a_1/\cdots/a_n/2^{-i}],$$

and further for i negative, $\frac{r}{s}$ must have a canonical expansion with last partial quotient a positive power of two, i.e. $2^{-i} \geq 2$.

Note that for i being zero both interpretations apply and $\frac{p}{q}$ is the parent of $\frac{r}{s}$. For i being negative, $\frac{p}{q}$ is its preconvergent. When i is positive, $\frac{p}{q}$ is not a convergent to $\frac{r}{s}$, but is a biconvergent to $\frac{r}{s}$.

The notion of bijacency may be equivalently characterized in terms of either LCF bitstrings or irreducible fractions as summarized in the following theorem.

Theorem 8. *Each of the following three properties implies the other two and serves as an equivalent definition of bijacency.*

i) *Continued fraction form: $\frac{p}{q} \lessdot \frac{r}{s}$ are bijacent of degree i as specified in (5),*

ii) *LCF form: The irreducible fractions $\frac{p}{q}$ and $\frac{r}{s}$ have lexicographically adjacent LCF expansions (and thus are neighbors in Q_k for some k),*

iii) *Fraction form: $\frac{p}{q} \lessdot \frac{r}{s}$ are bijacent of degree i iff for*

$i \geq 0$: $|ps - qr| = 2^i = gcd(r-p, s-q)$ and $\frac{r}{s} \lessdot \frac{2p}{2q}$,

$i \leq 0$: $|ps - qr| = 1$ and $\frac{2^{-i}p}{2^{-i}q} \lessdot \frac{r}{s} \lessdot \frac{(2^{-i}+1)p}{(2^{-i}+1)q}$ when $\frac{p}{q} \neq \frac{0}{1}$,

and $\frac{r}{s} = \frac{1}{2^{-i}}$ when $\frac{p}{q} = \frac{0}{1}$.

Proof. It is straightforward to prove ii) and iii) from i). We first prove i) given iii). Let us start with the case where $i \leq -1$ for $p \neq 0$, then there exists $\frac{a}{b} \lessdot \frac{p}{q}$ such that $r = 2^{-i}p + a$ and $s = 2^{-i}q + b$. Since $|\frac{p}{q} - \frac{r}{s}| = \frac{1}{qs} < \frac{1}{2q^2}$ it follows from Theorem 1 (viii) that $\frac{p}{q}$ is a convergent of $\frac{r}{s}$. Since $r \geq 2p$, $\frac{p}{q}$ cannot be the parent of $\frac{r}{s}$, so $\frac{r}{s} = [a_0/a_1/\cdots/a_n/2^{-i}]$ and $\frac{p}{q} = [a_0/a_1/\cdots/a_n]$. For $i = 0$

the same argument applies, except that now $\frac{p}{q} = [a_0/a_1/\cdots/a_n]$ is the parent of $\frac{r}{s} = [a_0/a_1/\cdots/a_n + 2^0]$.

Now, assume $i \geq 1$. Then there exists a and b such that $r - p = 2^i a$ and $s - q = 2^i b$. From

$$2^i |aq - bp| = |(r-p)q - (s-q)p| = |rq - sp| = 2^i$$

we obtain $|aq - bp| = 1$ and similarly $|as - br| = 1$. Hence $\frac{a}{b}$ is irreducible and adjacent to $\frac{p}{q}$ as well as to $\frac{r}{s}$. Now $\frac{r}{s} \lessdot \frac{2p}{2q}$ implies $\frac{r-p}{s-q} = \frac{2^i a}{2^i b} \lessdot \frac{p}{q}$ hence $\frac{a}{b} \lessdot \frac{p}{q}$ and as above we find that $\frac{a}{b}$ is a convergent of $\frac{p}{q}$ as well as of $\frac{r}{s}$. With $\frac{p}{q} = [a_0/a_1/\cdots/a_n]$, $a_n \geq 2$, either $\frac{a}{b}$ is the preconvergent or the parent of $\frac{p}{q}$. Assume $\frac{a}{b}$ is the parent, then

$$\frac{p}{q} = \frac{1 \cdot a + p_{n-2}}{1 \cdot b + q_{n-2}}$$

where $\frac{p_{n-2}}{q_{n-2}}$ is the preconvergent of $\frac{p}{q}$, hence $\frac{p}{q} \lessdot \frac{2a}{2b}$ which contradicts $\frac{2^i a}{2^i b} \lessdot \frac{p}{q}$. Thus we obtain both

$$\frac{a}{b} = [a_0/a_1/\cdots/a_{n-1}],$$
$$\frac{r}{s} = \frac{2^i a + p}{2^i b + q} = [a_0/a_1/\cdots/a_{n-1}/a_n + 2^i],$$

which completes the proof of (iii) $\Rightarrow$ (i).

Finally we have to prove that (ii) implies (i), hence assume for some $j \geq 0$ that $\frac{r}{s} = \sigma \circ 0 \circ 1^j$ and $\frac{p}{q} = \sigma \circ 1 \circ 0^j$. Two automatons decoding the bitstrings $\sigma 0 1^j$ and $\sigma 1 0^j$ after reading the string σ, will both be in the same state, having parsed and decoded an initial sequence of partial quotients $\{a_0, a_1, \cdots, a_{n-1}\}$, being in the state of decoding the n'th partial quotient a_n, when encountering the 0 or 1 following σ. The automatons will be in one of four possible states, depending on n being even or odd (reading in true or complemented form), and either reading the initial (unary including switch bit) part or the trailing (binary) part of $\ell(a_n)$.

We will start with n being even, and let us first assume the automatons are reading the unary part of a_n, having already seen k ones of the unary part. The 10 in the LCF expansion of $\frac{p}{q}$ thus implies that the unary part is completed, and

$$\frac{p}{q} = [a_0/a_1/\cdots/a_{n-1}/2^{k+1}],$$

since a suitable amount of zeroes just completes $a_n = 2^{k+1}$ as the final partial quotient (n is even). The interpretation of $\mathrm{LCF}(\frac{r}{s})$, however, depends on the relation between k and j, since the zero brings the automaton into the state of decoding the binary part.

If $j \leq k$ we find

$$\frac{r}{s} = [a_0/a_1/\cdots/a_{n-1}/(2^{j+1} - 1)2^{k-j}],$$

whereas when $j > k$ there are more ones than needed to complete the binary part, and the remaining ones then are interpreted as two additional partial quotients:

$$\frac{r}{s} = [a_0/a_1/\cdots/a_{n-1}/2^{k+1} - 1/1/2^{j-k}].$$

so that in both cases, $\frac{p}{q}$ and $\frac{r}{s}$ are found to be bijacent of degree $k - j$.

Now assume the automatons are reading the binary part, still having to read $k \geq 1$ bits to complete a_n, when σ has been read. If $k \geq j + 1$, then for some a we obtain

$$\frac{r}{s} = [a_0/a_1/\cdots/a_{n-1}/a \cdot 2^{k-j-1}] \text{ and } \frac{p}{q} = [a_0/a_1/\cdots/a_{n-1}/(a+1)2^{k-j-1}].$$

If $1 \leq k \leq j$, for some a_n we obtain

$$\frac{p}{q} = [a_0/a_1/\cdots/a_n + 1] \text{ and } \frac{r}{s} = [a_0/a_1/\cdots/a_n/1/2^{j-k}],$$

thus in both cases $\frac{p}{q}$ and $\frac{r}{s}$ are bijacent.

For n odd, the bits have to be inverted when reading a_n, and the continued fractions have to be written in even order expansion form, i.e. an $(n+1)$' partial quotient has to be considered. The results then follow similarly. □

By Theorem 8, the bijacency relation is precisely the relation that holds among neighbors in the inorder traversal of the LCF tree given to any depth k, specifically between consecutive fractions of Q_k. For the purpose of determining the members of Q_{k+1} given Q_k, again intuitively what we want is to characterize, in fraction or continued fraction form, is the rational number from Q_{k+1} whose LCF representation is obtained by one-bit extensions of the LCF representations of members of Q_k. Specifically, given two bijacent neighbors from Q_k with LCF representations $\sigma \circ 0 \circ 1^j$ and $\sigma \circ 1 \circ 0^j$, we want to characterize the unique rational $\sigma \circ 0 \circ 1^{j+1}$ from Q_{k+1}, falling between these.

Let us then for a moment digress to the classical concept of Farey fractions. Recall (e.g. from [Hardy and Wright 79]) that given the *Farey-set*

$$F_n = \left\{ \tfrac{p}{q} \;\middle|\; 0 \leq p \leq n, 1 \leq q \leq n, qcd(p,q) = 1 \right\}$$

a member of F_{2n} can be constructed as the *mediant* $\frac{p+r}{q+s}$ of two successive rationals $\frac{p}{q} < \frac{r}{s}$ already in F_n, which can always be shown to be adjacent as defined previously. We further obtain $\frac{p}{q} < \frac{p+r}{q+s} < \frac{r}{s}$. However, when say $\frac{p}{q}$ is a "very simple fraction", whereas $\frac{r}{s}$ is not (e.g. $p \ll r$ and $q \ll s$), their mediant will be of numeric value very close to $\frac{r}{s}$, and rather distant from $\frac{p}{q}$. In general, the spacing between consecutive members of F_n is quite erratic, varying between n^{-1} and n^{-2} [Matula and Kornerup 80]. Now if p and q above both were multiplied by some common factor c, chosen such that cp and cq were both of the same order of magnitude as r and s respectively, then the rational value of the expression $\frac{cp+r}{cq+s}$ would split the interval between p/q and r/s in two intervals of more nearly equal widths.

This observation leads us to the introduction of an alternative form of mediant, more nearly bisecting each gap between consecutive members of Q_k, which turns out to be precisely the way we can generate members of Q_{k+1} from Q_k.

The *binary mediant* of the bijacent fractions $\frac{p}{q} \lessdot \frac{r}{s}$ is the irreducible fraction $\frac{u}{v}$ with value

$$u/v = (2^j p + r)/(2^j q + s),$$

where j is the largest integer such that $\frac{2^j p}{2^j q} \lessdot \frac{r}{s}$. The binary mediant $\frac{u}{v}$ is defined to be in reduced form, however $gcd(2^j p + r, 2^j q + s)$ can only have value one or two. The following alternative characterizations and properties of the binary mediant are readily obtained by extending the proof of Theorem 8.

Theorem 9. *The binary mediant of two bijacent fractions $\frac{p}{q}, \frac{r}{s}$ is the irreducible fraction $\frac{u}{v}$ equivalently determined by either of the following three conditions:*

i) *Continued fraction form: If*

$$\begin{aligned} \frac{p}{q} &= [a_0/a_1/\cdots/a_{n-1}/a_n] \\ \frac{r}{s} &= [a_0/a_1/\cdots/a_{n-1}/a_n + 2^i] \\ \textit{with } a_n &= k2^i \geq 2 \textit{ when } i \geq 1, \textit{ then} \\ \frac{u}{v} &= [a_0/a_1/\cdots/an - 1/a_n + 2^{i-1}]. \end{aligned}$$

ii) *LCF expansion form: If $\frac{p}{q} < \frac{r}{s}$ and for some bit string σ and integer $j \geq 0$,*

$$\begin{aligned} \frac{p}{q} &= \sigma \circ 0 \circ 1^j, \\ \frac{r}{s} &= \sigma \circ 1 \circ 0^j, \\ \textit{then} & \\ \frac{u}{v} &= \sigma \circ 0 \circ 1^{j+1}. \end{aligned}$$

iii) *Fraction form: If $\frac{p}{q} \lessdot \frac{r}{s}$ are bijacent of degree i, then*

$$\frac{u}{v} = \begin{cases} \dfrac{2^{-i}p + r}{2^{-i}q + s} & \textit{for } i \leq 0, \\ \dfrac{(p+r)/2}{(q+s)/2} & \textit{for } i > 0. \end{cases}$$

Lemma 10. *The binary mediant $\frac{u}{v}$ of the bijacent fractions $\frac{p}{q}, \frac{r}{s}$ is bijacent to both $\frac{p}{q}$ and $\frac{r}{s}$.*

Theorems 8, 9 and Lemma 10 provide us the computational means of generating sets of bijacent fractions partitioning any interval specified by two bijacent fractions given in either the fraction, continued fraction or lexicographic continued fraction form of representation.

Regarding the LCF tree we then immediately obtain the following "LCF Tree Labelling Lemma", here employing in an obvious way the "infinite" fraction $\frac{1}{0}$.

Lemma 11. *In the LCF tree, the fraction $\frac{1}{1}$ is assigned to the root, and the fraction assigned to any other node is the binary mediant of the fraction assigned to the nearest left ancestor node (or $\frac{0}{1}$ if none exists) and the fraction assigned to the nearest right ancestor node (or $\frac{1}{0}$ if none exists) in the LCF tree. Furthermore, with edges labelled 0 for left branch and 1 for right branch, the bitstring composed from the labels on the edges of the path from the root to a particular node, with a terminal unit appended, provides the minimal LCF expansion of the fraction assigned to that node.*

4 Extremes and Distribution of Gap Sizes over Q_k

Traditional fixed point binary representation with k bits to the right of the radix point $0.b_1b_2\cdots b_k$ allows for the representation of 2^k+1 values over the unit interval $[0,1]$, with all gaps of uniform size 2^{-k}. The LCF expansion $0b_1b_2\cdots b_k$ with k bits beyond the reciprocal zero bit, determine an equal sized 2^k+1 membered set Q_k of representable values over the unit interval. The gap sizes over Q_k necessarily vary about the mean 2^{-k} to accommodate exact representations of the "simple" rational fractions. In this section we shall discuss bounds on the variations on the size of gaps over Q_k for large k. We first confirm the existence of certain relatively large and small gaps in Q_k, that we contend are indicative asymptotically of the maximum and minimum gap sizes in Q_k.

Lemma 12. *Given $\epsilon > 0$, then for sufficiently large k, the maximum gap size in Q_k will be at least $2^{-(a+\epsilon)k}$ for*

$$a = \frac{1}{4}\log_2(5+2\sqrt{6}) = 0.82682\cdots,$$

and the minimum gap size will be no bigger that $2^{-(b-\epsilon)k}$ for

$$b = \log_2\left(\frac{3+\sqrt{5}}{2}\right) = 1.38848\cdots.$$

Proof. For the minimum gap result, consider that $z = \frac{\sqrt{5}-1}{2} = [0/1/1/\cdots] = 010101\cdots = 0.618033\cdots$ has the sequence of convergents $\frac{0}{1}, \frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \cdots$ which (deleting $\frac{0}{1}$) are also seen to be the sequence of biconvergents to z. The numerators (and denominators) are the well known Fibonacci numbers and grow asymptotically at the rate $(1+\sqrt{5})/2$. The bounding intervals on z determined by the sequence of biconvergents are then $\frac{1}{1\times2}, \frac{1}{2\times3}, \frac{1}{3\times5}, \frac{1}{5\times8}, \cdots$, and decrease at a rate approaching $4/(1+\sqrt{5})^2 = 2/(3+\sqrt{5})$, verifying the minimum gap size bound.

For the maximum gap size result note that $y = \sqrt{6}-2 = [0/2/4/2/4/\cdots] = 001111000011110000\cdots = 0.449489\cdots$ has a sequence of convergents, where the rate of increase of numerators (and denominators) in two steps, $\frac{p_{i+2}}{p_i}$, approaches $5+2\sqrt{6} = 9.898979\cdots$. Note then that after eight bits in the LCF expansion corresponding to encoding another pair of partial quotients $2,4$, the bounding interval only determined by the biconvergents will decrease at a rate approaching $1/(5+2\sqrt{6})^2$, from which the lower bound on the maximum bound follows. □

For the purpose of bounding the approximation error in rounding to a bi-convergent in Q_k, we are concerned with determining an upper bound on the maximal gap size in Q_k. By direct computation we first obtain the maximum gap size in Q_k for moderate orders, given in terms of

$$a_k = -\frac{1}{k}\log_2(\text{max gap in } Q_k),$$

the values of which are listed for $k = 1, 2, \cdots, 20$ in the following table:

k	a_k	k	a_k
1	1.000	11	0.812
2	0.792	12	0.816
3	0.774	13	0.819
4	0.792	14	0.820
5	0.817	15	0.816
6	0.812	16	0.819
7	0.804	17	0.821
8	0.810	18	0.822
9	0.815	19	0.818
10	0.818	20	0.821

In these computations we found for the larger k that the maximum size gaps in Q_k always had a boundary point $\frac{p}{q} \in Q_k$ whose LCF representation contains replications of the bit pattern 00001111, consistent with the example of Lemma 12. This lemma further gives us an upper bound on any limiting value for a_k, and we suggest in the following that this is indeed the correct value.

Conjecture 13. $\lim_{k\to\infty} a_k = \frac{1}{4}\log_2(5+2\sqrt{6}) = 0.82682\cdots$, *where furthermore the gaps in Q_k containing the real number $\sqrt{6}-2 = [0/2/4/2/4/\cdots]$ decrease in size asymptotically in k as fast as the maximum size gaps in Q_k.*

The proof of the conjecture at this point appears quite tedious. We quote here without proof from [Kornerup and Matula 85] a somewhat weaker result, which in conjunction with Lemma 12 provides reasonable tight bounds on a_k.

Theorem 14. *Given $\epsilon > 0$, then for sufficiently large k, the maximum gap size in Q_k is no greater than $2^{-(a-\epsilon)k}$ for $a = \frac{1}{4}\log_2(16/153) = 0.814347\cdots$.*

Also from [Kornerup and Matula 85] we quote findings on some computations and simulations on the gap size distribution. For values of k up through 24 the distributions were computed exhaustively. In each case it was also found that the minimum gap in Q_k fell between two consecutive rationals of the form $\frac{f_{n-2}}{f_{n-1}}$ and $\frac{f_{n-1}}{f_n}$, where f_n denotes the nth Fibonacci number, in correspondence with the observation about the LCF expansion of $[0/1/1/\cdots/1]$.

The main purpose of the computations was, however, to obtain graphs of the distribution of gap sizes. Exhaustive computations up to $k = 24$ and simulations for $k = 32, 64, 128$ showed the distribution of the negative base 2 logarithm of gaps in Q_k to be bell-shaped between approximately 0.8 and 1.3, and centered around 1.0, the bell-shape getting narrower and higher peaked for increasing values of k.

Overall our study shows that the gaps between LCF representable values are subject to a variation in size corresponding to a worst case 19% precision loss (or equivalently a 19% storage capacity loss), and a best case 38% precision gain, in comparison with an equivalent fixed point binary system with uniform gap size over the unit interval. This appears to be a small price to pay for achieving exact representation of all simple rationals.

5 Conclusions

A binary representation of the rationals has been described and analyzed. It is capable of representing in finite precision a set of rationals fairly regularly spaced on the unit interval. It supports an online arithmetic unit for rational arithmetic, which can alternatively be considered an approximative real arithmetic with embedded exact computations on simple rationals.

The LCF representation is non-redundant, which for an on-line (digit serial, most significant digit first) arithmetic has the implication that the delay between input and output can vary unboundedly. To be able to bound and reduce such delays, it is necessary to introduce redundancy in the representation. It is straightforward to introduce redundancy in the continued fraction representation of rationals (allowing partial quotients suitably restricted to become negative), and furthermore to introduce redundancy in the binary encoding of the individual partial quotients. Arithmetic units supporting such redundant continued fraction representations have also been investigated and reported [see Kornerup and Matula 90], however an analysis of these representations has never been conducted and deserves a similar study.

Acknowledgements

This paper reports some previously unpublished results presented as part of an invited talk entitled "Rational Number Systems and Arithmetic", given by the first author at the symposium "Real Numbers and Computers", in Saint Etienne, April 4-6, 1995. The work has been supported by The Danish Research Councils, grants no. 11-8243 and 8.21.08.02, and by The National Science Fondation under grant DCR-8315289.

References

[Blachman 84] N. M. Blachman. The Continued Fraction as an Information Source. *IEEE Transactions on Information Theory*, IT-30(4):671–674, July 1984.

[Hardy and Wright 79] C. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, London, fifth edition, 1979.

[Khinchin 35] A. Y. Khinchin. *Continued Fractions*. The University of Chicago Press, 1935. Translated from Russian by P. Wynn, P. Noordhoff Ltd., Grooningen, 1963. Also by H. Eagle, The University of Chicago Press, 1964.

[Kornerup and Matula 83] P. Kornerup and D. W. Matula. Finite Precision Rational Arithmetic: An Arithmetic Unit. *IEEE Transactions on Computers*, C-32(4):378–387, April 1983.

[Kornerup and Matula 85] P. Kornerup and D. W. Matula. Finite Precision Lexicographic Continued Fraction Number Systems. In *Proc. 7th IEEE Symposium on Computer Arithmetic*, pages 207–214, 1985. Also reprinted in the collection *Computer Arithmetic, Vol II, E.E. Swartzlander, ed.*, IEEE Computer Society Press, Washington, 1990, pages 341-348.

[Kornerup and Matula 88] P. Kornerup and D. W. Matula. An On-line Arithmetic Unit for Bit-Pipelined Rational Arithmetic. *Journal of Parallel and Distributed Computing*, 5(3):310–330, May 1988.

[Kornerup and Matula 90] P. Kornerup and D. W. Matula. An Algorithm for Redundant Binary Bit-Pipelined Rational Arithmetic. *IEEE Transactions on Computers*, C-39(8):1106–1115, August 1990.

[Knuth 81] D. E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison Wesley, 2 edition, 1981.

[Knuth 82] D. E. Knuth. Supernatural Numbers. *in Mathematical Gardner, D.A. Klarner, ed.*, 1982.

[Matula and Kornerup 80] D. W. Matula and P. Kornerup. Foundations of Finite Precision Rational Arithmetic. *Computing, Suppl. 2*, pages 88–111, February 1980.

[Matula and Kornerup 83] D. W. Matula and P. Kornerup. An Order Preserving Finite Binary Encoding of the Rationals. *Proc. 6th IEEE Symposium on Computer Arithmetic*, pages 201–209, 1983.