Enumerating Rationals using Calkin-Wilf trees

Sunaina Pati

Chennai Mathematical Institute

February 17,2024

Outline

Introduction

What do we mean by countability?

A bijective function $f : X \to Y$ is a one-to-one (injective) and onto (surjective) mapping of a set X to a set Y.

Countable

We say a set X is countable if there is a bijection between X and a subset of $\mathbb{N}$

Same cardinality

Let A, B be sets. We say A and B are equipotent (or have the same cardinality) if there exists a bijection $f : A \to B$.

Now, let us consider some examples!

Examples

Is $\mathbb{N}$ countable?

Yes! Simply consider the bijective function as $\phi(n) = n$. It is easy to see why it is both injective and subjective.

Do you think $\mathbb{Z}$ is countable?

Try finding constructions (or counter-constructions)!

Is $\mathbb{Z}$ countable?

There can be several constructions! However, here are two:

- Consider $\phi : \mathbb{N} \to \mathbb{Z}$ such that

$$\phi(n) = \begin{cases} 0 & \text{if } n = 1 \\ \frac{n}{2} & \text{if } n = 2k, k \in \mathbb{N} \\ -\frac{n-1}{2} & \text{if } n = 2k, k \in \mathbb{N} \end{cases} \tag{1}$$

- Another construction but $\sigma : \mathbb{Z} \to \mathbb{N}$ such that

$$\sigma(n) = \begin{cases} 1 & \text{if } n = 0 \\ 2^n & \text{if } |n| = n \neq 0 \\ 3^{-n} & \text{if } |n| \neq n \end{cases} \tag{2}$$

Injection to $\mathbb{N}$ is sufficient

Note

If an injection exists such that $f : A \to \mathbb{N}$, the set A is countable

Since there is an injection into the naturals, then clearly there's a bijection into a subset of the naturals, call it B, then since $f : A \to B$ is a bijection we have $|A| = |B|$, and since $B \subset \mathbb{N}$ we have $|A| = |B|$, thus A is countable.

In general, Subsets of countable sets are countable.

$A \times A$ is countable if A is countable

Since A is countable, $\exists \phi$ such that $\phi : A \to \mathbb{N}$ is an injection. So let $(a, b) \in A \times A$. Then map $\sigma : (a, b) \to 2^{\phi(a)} \cdot 3^{\phi(b)}$. Note that this map σ is injective and lies inside $\mathbb{N}$.

Essentially, we can show A^n is countable if A is countable, as we know that number of primes are infinite. (Why?)

Also, if A and B are countable then $A \times B$ is countable.

Okay so Cartesian products are nice! What about unions of countable sets?

Union of finite number of countable sets is countable

Let the sets be $A_1, \ldots, A_n$ with each nonempty. For each n fix a function $f_n : A_n \to \mathbb{N}$ a one-one function.
Define a function $f : \cup A_n \to \mathbb{N}$ as follows.

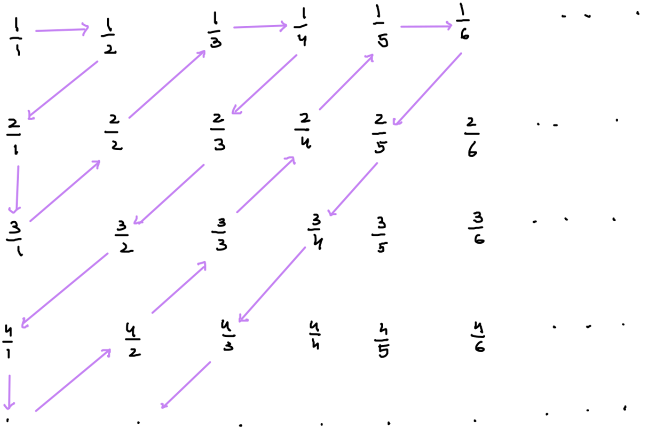
Take $x \in \cup A_n$. Let i be the first integer such that $x \in A_i$. Put $f(x) = 2^i 3^{f_i(x)}$. This is injective and is in $\mathbb{N}$. Hence countable.

Positive rationals are countable

Claim: The set of pairs $S = \{(m, n) : m, n \in \mathbb{N}\}$ is countable.
Proof: Put $f(m, n) = 2^m 3^n$ on S into $\mathbb{N}$. This is injective. Hence, countable.

Set of positive rational numbers is an infinte subset of the above by identifying a positive rational m/n with m/n where m, n have no common factors.

Geometric way to show countability of rationals



Rationals are countable!

Let $f(1), f(2), \dots$ be the mapping from natural numbers from $\mathbb{N} \to +\mathbb{Q}$.
Then, note that there is also a bijection from $\mathbb{N} \to -\mathbb{Q}$ by simply mapping $i \in \mathbb{N} \to -f(i)$.
And to create the bijection from $g : \mathbb{N} \to \mathbb{Q}$, consider

$$g(n) = \begin{cases} f(k) & \text{if } n = 2k, k \in \mathbb{N} \\ -f(k) & \text{if } n = 2k+1, k \in \mathbb{N}, k > 1 \\ 0 & \text{if } n = 1 \end{cases} \tag{3}$$

Algebraic Numbers are countable

Algebraic number

Algebraic number is a complex number which is root of a polynomial with rational coefficients.

Fix an integer $n \geq 1$. If we show set A_n of algebraic numbers which are roots of n-th degree polynomials is a countable set then the set of algebraic numbers $\cup A_n$ would also be countable by above. Let us fix $n \geq 1$. Consider one Polynomial $r_n x^n + r_{n-1} x^{n-1} + \cdots + r_1 x + r_0$. This polynomial has at most n distinct roots. The set of such rationals $(r_0, r_1, \ldots, r_n)$ is countable set and so A_n is countable union of finite sets and hence countable.

Reals are uncountables

The first solution uses nested intervals!
Clearly R is infinite. Let $f : \mathbb{N} \to \mathbb{R}$. We show that there is a real number x which is not $f(n)$ for any $n \in \mathbb{N}$.
start with the interval $[0, 1] = [a_1, b_1]$. Note there exists $[a_2, b_2] \subset [a_1, b_1]$ such that $f(1) \notin [a_2, b_2]$. Then get $[a_3, b_3] \subset [a_2, b_2]$ such that $f(2) \notin [a_3, b_3]$. In general, we have $[a_n, b_n] \subset \cdots \subset [a_1, b_1]$. Note that

$$a_1 \leq a_2 \leq \cdots \leq b_3 \leq b_2 \leq b_1.$$

Let now $A = \{a_n : n \geq 1\}$. This set is bounded above by b. Let $a = \sup A$. Note that $a \in [a_n, b_n]$. This a is our required x.

Power set of $\mathbb{N}$ is uncountable

This is essentially Cantor's Theorem.

Cantor's Theorem

Let f be a map from A to $P(A)$. Then f is not subjective. Hence not bijective.

Proof: Suppose not. Then f is surjective. Let $B = \{x \in A | x \notin f(x)\}$. Since f is surjective, there is a pre-image of B. Say $f(y) = B$. If $y \notin B$ then $y \in B$. If $y \in B$ then $y \notin B$. Contradiction.

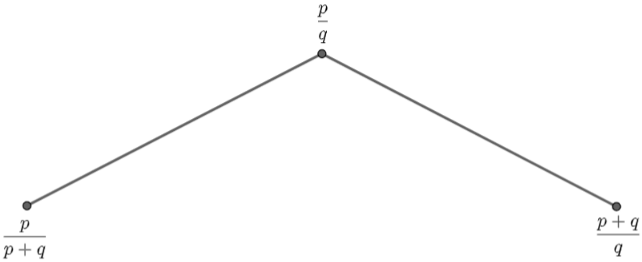
Diagonalisation

Cantor considered the set T of all infinite sequences of binary digits (i.e. each digit is zero or one). We shall show T is uncountable and that, there is an injection from T to $\mathbb{R}$ thus proving $\mathbb{R}$ to be uncountable.
For the injection, simply map the binary number, say $x_1x_2\dots$ to $0.x_1x_2\dots$
If $s_1, s_2, \dots, s_n, \dots$ is any enumeration of elements from T, then an element s of T can be constructed that doesn't correspond to any s_n in the enumeration which is constructed by: for every n, the nth digit as complementary to the nth digit of s_n.

Calkin-Wilf tree

What is it?

Every node of this binary tree is assigned a positive fraction.And the starting node is $\frac{1}{1}$. If the positive fraction is $\frac{p}{q}$, then we define it's left child as $\frac{p}{p+q}$ and the right child as $\frac{p+q}{q}$. We call the fraction $\frac{p}{q}$ as the parent of $\frac{p}{p+q}$ and $\frac{p+q}{q}$.



Introduction

In general, if we have the fraction as x, then note that it's left child is $\frac{x}{x+1}$ and the right child is $x+1$. However, before proceeding, we shall prove this.

Claim: If the parent is x, then it's left child is $\frac{x}{x+1}$ and the right child is $x+1$.

Proof:Let $x = \frac{p}{q}$. Then by definition,

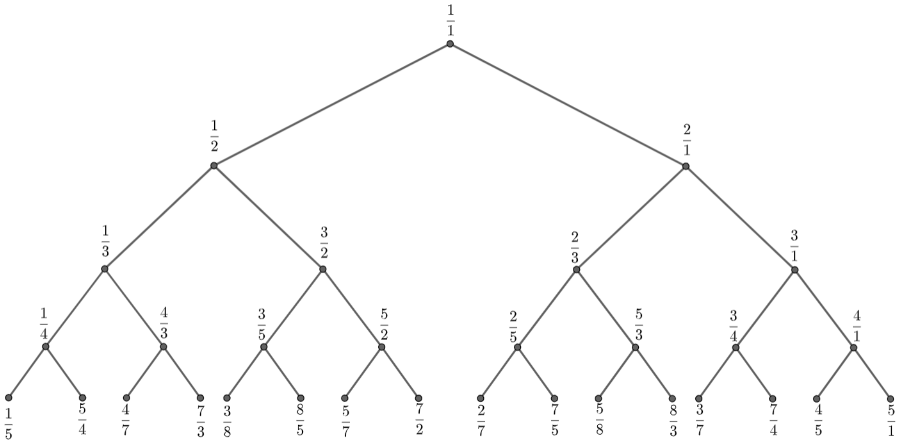
$$\text{the right child is } \frac{p+q}{q} = \frac{p}{q} + 1 = x + 1.$$

And by definition,

$$\text{the left child is } \frac{p}{p+q} = \frac{1}{\left(\frac{p+q}{q}\right)} \cdot \frac{p}{q} = \frac{x}{x+1}.$$

Infinite Calkin-Wilf Tree

We start with $\frac{1}{1}$ and proceed to make the left child and right child. Note that this process is infinite and hence the name.



Some properties

Here are a few properties of the infinite cute tree which we will proceed to prove in the talk.

Proposition

Every positive rational number appears in the tree and appears uniquely on the tree.

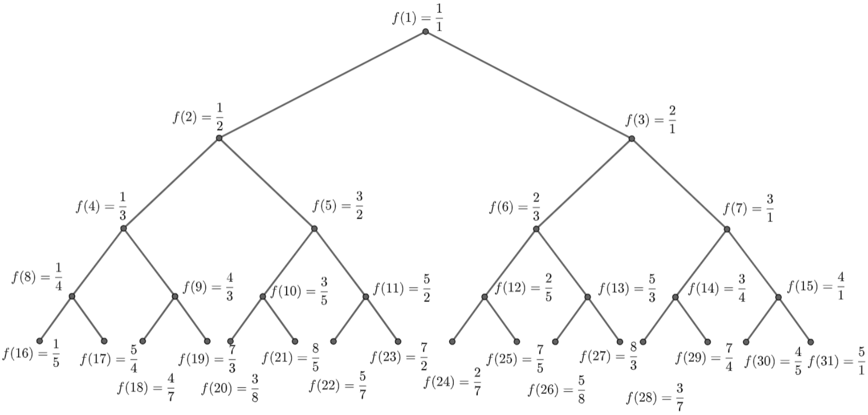
Proposition: Assuming this proposition is true, we can simply number every node in every row systematically. But how are we numbering? Start numbering from the first row and then number starting from the right side of the next row and continue. We get

$$f(1) \to \frac{1}{1}$$

$$f(2) \to \frac{1}{2}$$

$$f(3) \to \frac{2}{1}$$

Some Properties



Properties continues

Proposition

All the fractions in the infinite cute tree are in reduced form.

Proof: We can prove this using induction.
Claim: If $\frac{p}{q}$ is reduced then so is $\frac{p}{p+q}, \frac{p+q}{q}$.
Note that

$$(p,q)=1 \implies (p+q,p)=1,(q,p+q)=1 \implies \frac{p}{p+q},\frac{p+q}{q}$$

are in reduced form if $\frac{p}{q}$ is in reduced form . Clearly elements of row 1 are reduced. Say elements are reduced for kth row. Then by our above claim, we get that $k+1$th row is reduced, and by induction, we are done.

Properties Continues!

Proposition

The left child of any vertex is always strictly less than 1.

Proof: Let $\frac{p}{q}$ is the parent, then the left child is $\frac{p}{p+q}$ which is less than $\frac{p}{p} = 1$.

Proposition

The right child is always strictly greater than 1.

Proof: Let $\frac{p}{q}$ is the parent, then the right child is $\frac{p+q}{q}$ which is greater than $\frac{q}{q} = 1$.

Example

- For example, the left child of $\frac{20}{23}$ is $\frac{20}{43}$ which is less than 1.
- For example, the right child of $\frac{20}{23}$ is $\frac{43}{23}$ which is greater than 1.

And it continues!

Proposition

Every vertex is the product of its childs.

Proof For any node, the childs are $\frac{p}{q}$ are $\frac{p}{p+q}$ and $\frac{p+q}{q}$. And the clearly, the product of the children are $\frac{p}{p+q} \times \frac{p+q}{q} = \frac{p}{q}$.

Relations with Euclidean-Algorithm

Cool-Sequence Algorithm

This algorithm can be used to find the unique path from any given reduced fraction to $1/1$. Moreover, this algorithm also proves the first proposition we had!

$$\text{For any fraction } a/b, \text{ the parent is} \begin{cases} a/b-a & \text{if } b>a \\ a-b/b & \text{if } a>b \end{cases} \tag{4}$$

Example

For example, take any fraction, $\frac{13}{47}$ then the parent of $\frac{13}{47}$ is $\frac{13}{34}$. We will now trace it back to $1/1$ as all fractions originate from $1/1$. So we have

$$13/47 \to 13/34 \to 13/21 \to 13/8 \to 5/8 \to 5/3 \to 2/3 \to 2/1 \to 1/1.$$

Continues

Cool-sequence Path

Let it denote the path which take any fraction p/q to $1/1$.

Note that for any fraction which appears in the tree will have a unique parent and hence a unique cool sequence path. Hence, any fraction appearing in the cute tree appears only once.
Now, we claim that any reduced fraction appears on this tree. However, note that our algorithm works for any reduced fraction.

More Properties

Proposition

Number of elements in nth row is 2^{n-1}.

Proof: We can prove this by induction. The base case is true. Suppose it's true for kth row, then by definition of $k+1$th row, the $k+1$th row will contain childs of kth row. Note that these childs will be unique and distinct by the first proposition, we get that the number of elements in the $k+1$th row is 2 times the number of elements in the kth row, which by induction is 2^{k-1}. And we are done.

Proposition

The ith node from the left in any given row is the reciprocal of ith vertex from the right of the tree of that row.

Proof: We proceed with induction. True for the first row. Say it is true for the kth row. Then say the ith node from the left is p/q and the ith node from the left is q/p. Then we consider the children. The right child of p/q is the $2i - 1$th from left in row $k + 1$ and the left child of q/p is the $2i - 1$th from right in row $k + 1$. However, The right child of p/q is $p/p + q$ and the left child of q/p is $q + p/p$, and both are reciprocal of each other. Similarly, the left child of p/q and the right child of q/p are the $2i$th element from left and right, respectively, in row $k + 1$. And the left child of p/q is $p/p + q$ and the right child of q/p is $p + q/p$, which are clearly reciprocal of each other. And our hypothesis is true for $k + 1$th row. So, we are done by induction!

Proposition

The product of all the elements in a given row is 1.

Proof: True for the first row. For any $k > 1$th row, we get that there are even number of elements. Using the above proposition, we get that every ith element from the left can be paired with the ith element from the right. Note that both these elements are reciprocal of each other, and hence the product of them is 1. And hence the product is 1.

Proposition

The 1st node from the left in nth row is $1/n$ and the $1st$ node from right in nth row is $n/1$.

Proof: True for the first row. Say it is true for the kth row. Then the leftest node must be $1/k$ and rightest node is $k/1$. Note that the leftest node in the $k+1$th row is the left child of the leftest node in kth row. Hence it must be $1/(k+1)$. Similarly, we get that the rightest node in the $k+1$th row is the right child of the rightest node in kth row. Hence it must be $(k+1)/1$.

Proposition

Sum of all elements in nth row is $3 \cdot 2^{n-2} - \frac{1}{2}$.

Proof: We proceed with induction. It is true for $n = 1$ row. Say it's true for kth row. Now consider the $k+1$ row. For any fraction, we know it's reciprocal is there. So for a/b in the kth row, we know that b/a is there too. We consider the children which are in the $k+1$th row and the sum. So we get

$$a/(a+b) + (a+b)/b + b/(a+b) + (a+b)/a = 3 + a/b + b/a.$$

Hence, we get the sum of elements in $k+1$th row is

$$3 \times 2^{k-2} + 3 \cdot 2^{k-2} - \frac{1}{2} = 3 \times 2^{k-1} - \frac{1}{2}.$$

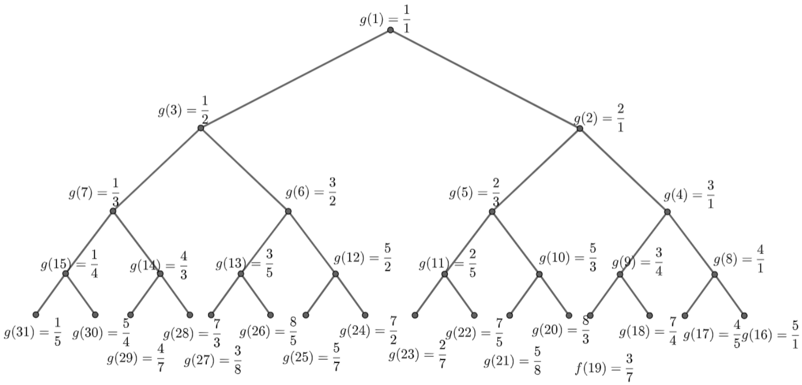
And we are done by induction.

Binary Preimage

Using our cool sequence algorithm, we showed that our function f covers all positive rational numbers uniquely. Now, we will show that given any fraction, how we can find its preimage (which is obviously unique).
Remember our Cool sequence algorithm? We will do the same thing, except now we will also care about the left and right child.
For example, we know that 13/47 is leftchild of 13/34 and so on. So we get

$$\frac{13}{47} \xrightarrow{L} \frac{13}{34} \xrightarrow{L} \frac{13}{21} \xrightarrow{L} \frac{13}{8} \xrightarrow{R} \frac{5}{8} \xrightarrow{L} \frac{5}{3} \xrightarrow{R} \frac{2}{3} \xrightarrow{L} \frac{2}{1} \xrightarrow{R} \frac{1}{1}.$$

We define a new function, $g(n)$ which is basically the same as $f(n)$ but we define $g(n)$ as the following.



Note that

$$f(n) = g(2^{[}\log_2(n)] + 2^{[}\log_2(n) + 1] - 1 - n).$$

Note that if the number $g(n)$ is the parent, then its right child is $g(2n)$ and the left child is $g(2n+1)$. Now consider Now, consider the binary representation of $n, 2n$ and $2n+1$. If the $n = (a_1 a_2 \dots a_k)_2$ then note that $2n = (a_1 \dots a_k 0)_2$ and $2n + 1 = (a_1 \dots a_k 1)_2$.
So we add 1 at the end of the binary reprsentation whenever we go to the left child and 0 if right.
Hence, for the above example, we start with $1/1$. Since $2/1$ is right child, we add 1 to binary expression. Hence $(11)_2$. And we proceed like this.
We get that

$$13/47 = g((110101000)_2) = g(424) = f(512 + 256 - 1 - 424) = f(343).$$

So did we prove?

Note that, using our **proposition** and **binary preimage**, we get that there is a bijection between naturals and positive rationals. We get that there is a bijection between naturals and rational numbers. Hence, rationals are countable.

Connection with Euclidean ALgorithm

We consider 13/47 as the example. Recall, we had

$$\frac{13}{47} \xrightarrow{L} \frac{13}{34} \xrightarrow{L} \frac{13}{21} \xrightarrow{L} \frac{13}{8} \xrightarrow{R} \frac{5}{8} \xrightarrow{L} \frac{5}{3} \xrightarrow{R} \frac{2}{3} \xrightarrow{L} \frac{2}{1} \xrightarrow{R} \frac{1}{1}.$$

Note that

$$47 = 3 \times 13 + 8 \text{ hence three times L}$$

$$13 = 1 \times 8 + 5 \text{ hence one time R}$$

$$8 = 1 \times 5 + 3 \text{ hence one time L}$$

$$5 = 1 \times 3 + 2 \text{ hence one time R}$$

$$3 = 1 \times 2 + 1 \text{ hence one time L}$$

$$2 = 2 \times 1 + 0$$

Note that at every step, our L, R are alternating. This is because if $a = qb + r$ was say L, then that implies the fraction was b/a and then after qL's we get to b/r, but we have $b > r$ by euclidean algorithm. So we get that b/r must be a right child. Hence they are alternating.
Moreover, whenever we get the equation of the form $n = n \times 1 + 0$ at the end of the euclidean algorithm (note that we are supposed to get this as the fractions are in reduced form), we write $n - 1$ times the symbol.
Hence, for our current example, we have $n - 1 = 2 - 1 = 1$ times R.
Since we can connect euclidean algorithm and finite continued fractions, we can also connect the fractions with finite continued fractions.

Some Formulas

Address Formula

Using the binary preimage, we can now figure out the actuall address for any fraction.
For any fraction a/b, we figure out the n such that $f(n) = a/b$. The the row $f(n)$ belongs to is $[\log_2(n)] + 1$ as in the kth row, $f(2^{k-1}), \ldots, f(2^k - 1)$ are there. And then the fraction would be $n - 2^{[\log_2(n)]}$th element from left.

Algberaic Formula

We claim the following.

Theorem

We have the following recursive formula

$$f(n+1) = \frac{1}{[f(n)] + 1 - \{f(n)\}}$$

with $f(1) = 1/1 = 1$.

We firstly simplify the formula.

Lemma

Note that

$$f(n+1)=\frac{1}{[f(n)]+1-\{f(n)\}}=\frac{1}{2[f(n)]+1-f(n)}.$$

Proof: As we have $f(n)=[f(n)]+\{f(n)\}$, we get that

$$[f(n)]+1-\{f(n)\}=[f(n)]+1-(f(n)-[f(n)])=2[f(n)]+1-f(n).$$

Claim

: For any fraction $f(n+1)$ where $f(n+1)$ is the left child, we can say that $f(n+1) = \dfrac{1}{2[f(n)] + 1 - f(n)}$

Proof: We get that the left child is $f(n) = \frac{x}{x+1}$ and right child is $f(n+1) = x+1$ for any vertex x. We have to show that

$$\frac{1}{2[f(n)] + 1 - f(n)} = f(n+1)$$

or show that

$$2f(n+1)[f(n)] + f(n+1) - f(n)f(n+1) = 1.$$

But note that $f(n)$ is a left child, so we get that $2f(n+1)[f(n)] = 0$. Moreover, we get that

$$f(n+1) - f(n)f(n+1) = x + 1 - x = 1.$$

And we are done!

So our theorem holds true whenever $f(n)$ is a left child. Now, we see the case when $f(n)$ is the right child. Since $f(n)$ is the right child, we have $f(n+1)$ is the left child of some other parent. Since all fractions are getting generated by $1/1$, the fraction $f(n), f(n+1)$ must have same parent some rows ago. Let a/b be that common parent fraction $k+1$ rows. Then note that $f(n)$ is generated by taking k consecutive right childs after one left child from a/b and $f(n+1)$ is generated by taking k consecutive left child after one right child from a/b.

Now, before proceeding further, we present a lemma.

Lemma

For any number $x = p/q$, the rightmost child after n rows is $(p + nq)/q = x + n$, and the leftmost child is $p/(np + q) = x/(nx + 1)$.

Proof: The proof is just induction. It is true for $n = 1$ case. Say it is true for $n = k$. Then note that the right child of rightmost child in kth row is rightmost child in $k + 1$th row. And right child is $(p + (k + 1)q)/q = x + k + 1$. And similarly for left child. And we are done by induction.

Final Slide!

The left child of a/b is $a/(a+b)$ and right child is $(a+b)/b$. As $f(n)$ is is generated by taking k consecutive right childs after one left child from a/b, by above lemma, we have

$$f(n) = a/(a+b) + k$$

and we have

$$\begin{aligned} f(n+1) &= [(a+b)/b]/[k(a+b)/b+1] \\ &= (a+b)/(k(a+b)+b) = 1/(k+b/(a+b)). \end{aligned}$$

Now, note that $[f(n)] = k$. And hence we have

$$f(n) - k = a/(a+b) \implies f(n) - [f(n)] = a/(a+b).$$

So note that

$$f(n+1) = \frac{1}{1 - a/(a+b) + k} = \frac{1}{1 - f(n) + 2k} = \frac{1}{1 - f(n) + 2[f(n)]}.$$

Hence, we prove for the right child too.

Fun things to check out!

- Now, you can finally check out Hilbert's Infinite hotel paradox (which isn't a paradox but rather counter intuitive)
- If you know haskell, perhaps check Functional pearl's Enumerating the Rationals
- Do check the original paper! Calkin Neil, Wilf Herbert, Recounting the rationals, American Mathematical Monthly, Mathematical Association of America, 107 (4): (2000) 360-363, doi:10.2307/2589182, JSTOR 2589182