Problem of the Month
Solution to Problem 5: February 2020

Solution

Before answering the specific parts of this question, we will work through a few examples to convey the spirit of the arguments presented.

In parts (a) and (b), the general idea will be to find the smallest integer $n$ with the property that $\frac{1}{n} < \alpha$ and then “reduce” the problem to writing the number $\alpha - \frac{1}{n}$ as the sum of reciprocals of distinct positive integers.

As suggested in the statement of part (a), let us do this for $\alpha = \frac{2}{7}$. You can check that $\frac{1}{2}$ and $\frac{1}{3}$ are both greater than $\frac{2}{7}$, but $\frac{1}{4} < \frac{2}{7}$. This means 4 is the smallest positive integer whose reciprocal is less than $\frac{2}{7}$. Thus, we compute $\frac{2}{7} - \frac{1}{4} = \frac{1}{28}$.

We are left with a rational number having a numerator of 1, so we can rearrange to get

$$\frac{2}{7} = \frac{1}{4} + \frac{1}{28}.$$ 

In the language of the problem statement, we would take $n_1 = 4$ and $n_2 = 28$.

The process stopped after 1 step this time, but it will generally take longer. For example, consider $\alpha = \frac{13}{17}$. This time, $\frac{1}{2} < \alpha$ so 2 is the smallest positive integer whose reciprocal is less than $\frac{13}{17}$. We set $n_1 = 2$ and compute $\frac{13}{17} - \frac{1}{2} = \frac{9}{34}$. Since $\frac{9}{34} - \frac{1}{3} = -\frac{7}{102}$ is negative, $\frac{1}{3} > \frac{9}{34}$. On the other hand, $\frac{9}{34} - \frac{1}{4} = \frac{2}{136} = \frac{1}{68}$ is positive, so $\frac{1}{4} < \frac{9}{34}$. This means 4 is the smallest positive integer $n$ satisfying $\frac{1}{n} < \frac{9}{34}$, so we set $n_2 = 4$. Now compute

$$\frac{13}{17} - \frac{1}{2} - \frac{1}{4} = \frac{9}{34} - \frac{1}{4}$$

$$= \frac{2}{136}$$

$$= \frac{1}{68}$$

which leads to

$$\frac{13}{17} = \frac{1}{2} + \frac{1}{4} + \frac{1}{68}.$$ 

With $n_3 = 68$, the process is complete after two steps. In general, this procedure will take many steps and would be difficult or impossible to implement by hand, but it is a nice exercise to program a computer to carry out this procedure.

In part (b), it was suggested that we write $\frac{2}{7}$ as a sum of reciprocals of distinct positive integers all of which are larger than 5. The smallest positive integer $n$ with the property that $\frac{1}{n} < \frac{2}{7}$ is 4, but we only want integers greater than 5. Note that since $\frac{1}{4} < \frac{2}{7}$, we also have that $\frac{1}{5}$ and $\frac{1}{6}$ are less than $\frac{2}{7}$, so we take $n_1 = 6$ this time since it is the smallest positive integer whose reciprocal is less than $\frac{2}{7}$. 

that we are allowed to use. Compute \(\frac{2}{7} - \frac{1}{6} = \frac{5}{42}\). You can check that the smallest positive integer \(n\) satisfying \(\frac{1}{n} < \frac{5}{42}\) is \(n = 9\) so we take \(n_2 = 9\). It is important to note that \(n_2\) is larger than 5 and different from \(n_1\). Now compute \(\frac{5}{42} - \frac{1}{9} = \frac{3}{378} = \frac{1}{126}\) which has a numerator of 1, so we take \(n_3 = 126\). Thus,

\[
\frac{2}{7} = \frac{1}{6} + \frac{1}{9} + \frac{1}{126}.
\]

This expression is completely different from the one previously found for \(\frac{2}{7}\), which will be important in part (c).

This type of procedure where we always take the “first thing that works” is often called a “greedy algorithm”. The term “greedy” refers to the fact that each choice made is the best based on the current information only; consideration is not given to how a decision may affect later decisions. The procedures outlined above always work (as we will show), but they do not always lead to the expression involving the fewest terms or the smallest denominators. For example, following the procedure above for \(\frac{19}{88}\) (and no restriction on positive integers that can be used) leads to

\[
\frac{19}{88} = \frac{1}{5} + \frac{1}{63} + \frac{1}{27700}
\]

which is true, but a much simpler representation is

\[
\frac{19}{88} = \frac{1}{8} + \frac{1}{11}.
\]

This illustrates a common defect of greedy algorithms. While they can often be used to achieve what is wanted, they may not produce an optimal result. On the other hand, greedy algorithms tend to be simple to understand and implement, which can be advantageous if you are not concerned with optimization.

We will now proceed with part (a).

(a) We first state and prove two facts to be used in solving this problem. You may want to ignore the proofs for now and come back to them later.

**Fact 1:** Suppose \(a\) and \(b\) are positive integers with \(0 < \frac{a}{b} < 1\), \(a > 1\), and \(\frac{a}{b}\) in lowest terms. If \(n\) is the smallest positive integer satisfying \(\frac{1}{n} < \frac{a}{b}\), then the numerator of \(\frac{a}{b} - \frac{1}{n}\) after combining into one fraction and reducing to lowest terms is less than \(a\).

**Proof.** Since \(n\) is the *smallest* positive integer satisfying \(\frac{1}{n} < \frac{a}{b}\), it must be the case that \(\frac{a}{b} \leq \frac{1}{n-1}\). Since \(\frac{a}{b}\) is in lowest terms and \(a > 1\), it is impossible for this to be an equality, which means \(\frac{a}{b} < \frac{1}{n-1}\). This can be rearranged to \(a(n-1) < b\) or \(an - b < a\).

Choose \(a'\) and \(b'\) to be the positive integers satisfying \(\frac{a'}{b'} = \frac{a}{b} - \frac{1}{n}\) with \(\frac{a'}{b'}\) in lowest terms. The claim is that \(a' < a\). By finding a common denominator, we have that

\[
\frac{a'}{b'} = \frac{an-b}{bn}
\]
and since $\frac{a'}{b'}$ is in lowest terms, this means $a' \leq an - b$. We established in the first paragraph that $an - b < a$, so these inequalities can be combined to get $a' < a$. \qed

**Fact 2:** Suppose $\alpha$ is a rational number and $n$ and $m$ are positive integers. If $\alpha < 1$ and $n$ is the smallest positive integer with $\frac{1}{n} < \alpha$, and $\frac{1}{m} \leq \alpha - \frac{1}{n}$, then $n < m$.

*Proof.* Adding $\frac{1}{n}$ to both sides of $\frac{1}{m} \leq \alpha - \frac{1}{n}$ gives

$$\frac{1}{n} + \frac{1}{m} \leq \alpha. \quad (1)$$

Since $\frac{1}{n}$ and $\frac{1}{m}$ are both positive, this means $\frac{1}{m} < \alpha$. Since $n$ is the smallest integer satisfying $\frac{1}{n} < \alpha$, we must have $n \leq m$. We need only rule out the possibility that $n = m$. If this were the case, inequality (1) would become

$$\frac{1}{n} + \frac{1}{n} = \frac{2}{n} \leq \alpha. \quad (2)$$

Since $n$ is the smallest positive integer such that $\frac{1}{n} < \alpha$, we also know that $\alpha \leq \frac{1}{n-1}$, so

$$\frac{2}{n} \leq \frac{1}{n-1}.$$  

This can be rearranged to $2(n - 1) \leq n$ or $n \leq 2$. We cannot have $n = 1$ since our assumptions would imply $1 = \frac{1}{n} < \alpha$. If $n = 2$, we get $1 = \frac{2}{2} = \frac{2}{n} \leq \alpha$ from (2). Either way, this contradicts the assumption that $\alpha < 1$. This means $m \neq n$, so $n < m$ as claimed. \qed

We now describe a procedure to produce $n_1, \ldots, n_k$. The proof that it works is essentially a combination of Facts 1 and 2.

We begin with a rational number $\alpha$ with $0 < \alpha < 1$. For notational convenience, we label $\alpha = \alpha_1$. Now do the following

- Set $n_1$ to be the smallest positive integer satisfying $\frac{1}{n_1} < \alpha_1$.
- Set $\alpha_2 = \alpha_1 - \frac{1}{n_1}$ written as a reduced fraction. If the numerator of $\alpha_2$ is 1, stop.
- Set $n_2$ to be the smallest positive integer satisfying $\frac{1}{n_2} < \alpha_2$.
- Set $\alpha_3 = \alpha_2 - \frac{1}{n_2}$ written as a reduced fraction. If the numerator of $\alpha_3$ is 1, stop.
- ...
- Set $n_i$ to be the smallest positive integer satisfying $\frac{1}{n_i} < \alpha_i$.
- Set $\alpha_{i+1} = \alpha_i - \frac{1}{n_i}$ written as a reduced fraction. If the numerator of $\alpha_{i+1}$ is 1, stop.
• Continue until a numerator of 1 is reached.

By Fact 1, the numerator of $\alpha_{i+1} = \alpha_i - \frac{1}{n_i}$ is smaller than the numerator of $\alpha_i$. We are also choosing the $n_i$ so that the $\alpha_i$ are positive, which means the numerator of $\alpha_k$ must be 1 for some $k$. If we set $n_k$ to be the denominator of $\alpha_k$, we will have

$$\alpha = \alpha_1 = \frac{1}{n_1} + \alpha_2 = \frac{1}{n_1} + \frac{1}{n_2} + \alpha_3$$

$$\vdots$$

$$= \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \cdots + \frac{1}{n_{k-1}} + \alpha_k$$

$$= \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \cdots + \frac{1}{n_{k-1}} + \frac{1}{n_k}.$$  

For each $i$ from 1 to $k - 1$, we have that $\frac{1}{n_{i+1}} \leq \alpha_{i+1} = \alpha_i - \frac{1}{n_i}$. [The inequality will be strict for all $i$ except $i = k - 1$.] By Fact 2, this means $n_1 < n_2, n_2 < n_3$, and so on so that $n_1 < n_2 < n_3 < \cdots < n_{k-1} < n_k$.

(b) First, let us introduce some notation. For positive integers $a$ and $b$ with $a < b$, we define $S(a, b)$ to be the sum

$$S(a, b) = \frac{1}{a} + \frac{1}{a+1} + \frac{1}{a+2} + \cdots + \frac{1}{b-1} + \frac{1}{b}.$$  

For example,

$$S(5, 9) = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9}.$$  

Notice that if $a < b < c$, then $S(a, c) = S(a, b) + S(b + 1, c)$. You might want to check this by considering $a = 5, b = 7,$ and $c = 9$ above. As well, for convenience we define $S(a, a) = \frac{1}{a}$.

We start with a fact that is possibly of independent interest: For every positive integer $\ell > 1$, $S(\ell + 1, 2\ell) > \frac{1}{2}$. This follows from the fact that

$$\frac{1}{\ell+1} > \frac{1}{\ell+2} > \frac{1}{\ell+3} > \cdots > \frac{1}{2\ell},$$

which means

$$S(\ell + 1, 2\ell) = \frac{1}{\ell+1} + \frac{1}{\ell+2} + \frac{1}{\ell+3} + \cdots + \frac{1}{2\ell}$$

$$> \frac{1}{2\ell} + \frac{1}{2\ell} + \frac{1}{2\ell} + \cdots + \frac{1}{2\ell}$$

$$= \ell \left( \frac{1}{2\ell} \right)$$

$$= \frac{1}{2}. \quad (3)$$
where the sums in lines (3) and (4) have \( \ell \) terms, and the inequality holds because \( \frac{1}{2\ell} \) is the smallest quantity in the sum in line (3).

We now let \( m \) be an arbitrary positive integer greater with \( m > 1 \). The above fact with \( \ell = m \) implies \( S(m + 1, 2m) > \frac{1}{2} \), and with \( \ell = 2m \) implies

\[
S(\ell + 1, 2\ell) = S(2m + 1, 4m) > \frac{1}{2}.
\]

Thus,

\[
S(m + 1, 4m) = S(m + 1, 2m) + S(2m + 1, 4m) > \frac{1}{2} + \frac{1}{2} = 1
\]

where the first equality in the line above comes from the fact stated after the definition of \( S(a, b) \).

We now verify the claim in the problem. To do this, first choose \( m \) to be the smallest positive integer satisfying both \( \frac{1}{m+1} < \alpha \) and \( T < m + 1 \). [It may seem like an odd choice to use \( m + 1 \) instead of \( m \), but it will make the notation a bit easier.] By our assumption and what we just showed,

\[
\frac{1}{m+1} < \alpha < 1 < S(m + 1, 4m).
\]

This means there is an integer \( r \geq m + 1 \) with the property that

\[
S(m + 1, r) < \alpha \leq S(m + 1, r + 1).
\]

This is because we know the sum \( S(m + 1, i) \) exceeds \( \alpha \) for \( i = 4m \) and hence for all \( i > 4m \), so we can choose \( r \) to be the largest positive integer satisfying \( S(m + 1, r) < \alpha \). We set \( \alpha' = \alpha - S(m + 1, r) \) which is positive by the choice of \( r \), and less than 1 since \( \alpha \) is less than 1 and \( S(m + 1, r) \) is positive.

We now apply part (a) to \( \alpha' \) to find integers \( p_1 < p_2 < \cdots < p_\ell \) satisfying

\[
\alpha' = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \cdots + \frac{1}{p_\ell}.
\]

Since \( \alpha' = \alpha - S(m + 1, r) \), we have that

\[
\alpha = \left( \frac{1}{m+1} + \frac{1}{m+2} + \cdots + \frac{1}{r} \right) + \left( \frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_\ell} \right).
\]

We have that \( m + 1 < m + 2 < \cdots < r - 1 < r \) since they are consecutive integers and \( p_1 < p_2 < \cdots < p_\ell \) from part (a). If we can show that \( r < p_1 \), we will be done.

By part (a), \( p_1 \) was chosen to be the smallest positive integer satisfying \( \frac{1}{p_1} < \alpha' \). We know that

\[
\alpha \leq S(m + 1, r + 1) = S(m + 1, r) + \frac{1}{r+1},
\]
\[
\alpha - S(m + 1, r) \leq \frac{1}{r+1}
\]
which is the same as
\[
\alpha' \leq \frac{1}{r+1}.
\]
It follows that \(\alpha' \leq \frac{1}{2}\) for any \(d \leq r + 1\), which implies \(p_1 > r + 1\) since \(p_1\) does not satisfy \(\alpha' \leq \frac{1}{p_1}\). Hence, \(p_1 > r\) as claimed. We can now take the integers \(n_1, n_2, \ldots, n_k\) to be the integers 
\(m + 1, m + 2, m + 3, \ldots, r, p_1, p_2, \ldots, p_\ell\) in that order. Here, \(k\) will be 
\(r - m + \ell\).

(c) Suppose \(\alpha\) is any positive rational number. Set \(r = \lfloor \alpha \rfloor\), the floor of \(\alpha\). Additionally, define \(\beta = \alpha - r\). That is, \(r\) is the largest integer that does not exceed \(\alpha\), and \(0 \leq \beta < 1\) is some rational number.

By part (a), if \(\beta > 0\), then there are distinct positive integers whose reciprocals sum to \(\beta\). In this situation, let \(T\) be the largest of these positive integers. If \(\beta = 0\), skip this step and set \(T = 0\).

Now use part (b) to find distinct positive integers all larger than \(T\) whose reciprocals sum to \(\frac{1}{2}\). Replace the value of \(T\) with the largest positive integer whose reciprocal was used in this sum.

Use part (b) again to find another set of distinct positive integers all larger than this new value of \(T\) and whose reciprocals sum to \(\frac{1}{2}\). We now have distinct positive integers whose reciprocals sum to \(\frac{1}{2} + \frac{1}{2} + \beta\). Once again, update \(T\) to be the largest integer used so far.

Repeat this process \(2r - 2\) more times always updating \(T\) to be the largest positive integer used so far then applying part (b) to find new positive integers whose reciprocals sum to \(\frac{1}{2}\). The sum of the reciprocals of the distinct positive integers will now be \(2r\) copies of \(\frac{1}{2}\) plus \(\beta\), or \(2r \left(\frac{1}{2}\right) + \beta = \alpha\).

(d) Suppose \(M\) is a positive real number and take \(r\) to be any integer with \(M < r\). By part (c), there are distinct integers \(n_1, n_2, \ldots, n_k\) so that
\[
r = \frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_k}.
\]
Then we certainly have
\[
r \leq 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots + \frac{1}{n_k},
\]
the sum of the reciprocals of the first \(n_k\) positive integers. This means
\[
M < 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots + \frac{1}{n_k}.
\]
Thus, the sum

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} \cdots$$

eventually exceeds any given real number, so the infinite sum cannot exist.

Additional Notes:

1. A sum of reciprocals of distinct positive integers is known as an *Egyptian Fraction*. You might want to do an internet search. There are lots of interesting questions about them. For example, how do you find “optimal” representations of rational numbers (using the smallest possible denominators or using as few fractions as possible)?

2. A few people sent in solutions that used the following fact. For any positive integer $k$, one has

$$\frac{1}{k} = \frac{1}{k+1} + \frac{1}{k(k+1)}.$$

This seems like a very inviting approach for problems (a) and (b), and I do believe it can be made to work. For example, to express $\frac{2}{3}$ as a sum of reciprocals of distinct positive integers, we could use this fact to get

$$\frac{2}{3} = \frac{1}{3} + \frac{1}{3}$$

$$= \frac{1}{3} + \frac{1}{4} + \frac{1}{34}$$

$$= \frac{1}{3} + \frac{1}{4} + \frac{1}{12}.$$

I did consider writing a solution using this fact, but found a general argument for the uniqueness of the positive integers to be a bit problematic. For example suppose (relating to part (b)) one wanted to use this fact to turn $\frac{1}{2} + \frac{1}{6} + \frac{1}{7} + \frac{1}{12} = \frac{75}{84}$ into an expression for $\frac{75}{84}$ using only integers larger than 5. The $\frac{1}{2}$ could be split into $\frac{1}{3} + \frac{1}{6}$. The $\frac{1}{3}$ is still too small, so we would need to split the $\frac{1}{3}$ into $\frac{1}{4} + \frac{1}{12}$; then the $\frac{1}{4}$ into $\frac{1}{5} + \frac{1}{20}$; and $\frac{1}{5}$ into $\frac{1}{6} + \frac{1}{30}$. In this process, we now have three copies of $\frac{1}{6}$ and two copies of $\frac{1}{12}$, so the fact has to be applied several more times. But how many more times? Splitting $\frac{1}{6}$ into $\frac{1}{7} + \frac{1}{42}$ gets rid of a $\frac{1}{6}$, but there is now an extra $\frac{1}{7}$ and we still have an extra copy of $\frac{1}{6}$.

My belief is that this technique can be used to solve parts (a) and (b), but it seems like an intricate argument is needed in order to guarantee that this process can be used to eventually lead to a sum of reciprocals of distinct positive integers. In other words, there is a systematic way of applying this process to make sure to get rid of all duplications. Of course, there may be a simple argument as well. If anyone has such an argument, I would love to see it!