Problem of the Month

Solution to Problem 4: January 2023

(a) The positive integer 1 cannot be expressed as the sum of more than 1 positive integer, so if A compresses $[1:9]$, then A must contain the integer 1. We want A to be a list of positive integers of length four, and 1 is the smallest positive integer, so we can assume that $A = [1, a, b, c]$ where a, b, and c are integers with $1 \leq a \leq b \leq c$.

The largest sum in $f(A)$ must be the sum of the three largest items in A, which is $a+b+c$, and since $f(A) = [1:9]$, we have $a + b + c = 9$.

Suppose that $a = 1$ and $b = 1$, then $a+b+c = 9$ implies that $c = 7$, but then $A = [1, 1, 1, 7]$, which does not satisfy $f(A) = [1 : 9]$ since, for example, 4 is not in $f(A)$. Therefore, a and b are not both 1.

Suppose that $a = 1$ and $b = 2$. Then $a + b + c = 9$ implies $c = 6$, so $A = [1, 1, 2, 6]$, but then $f(A)$ does not contain 5 so $f(A) \neq [1 : 9]$. Therefore, we cannot have $a = 1$ and $b = 2$.

Suppose that $a = 1$ and $b = 3$. Then $c = 5$, so $A = [1, 1, 3, 5]$. It can be verified that $f([1, 1, 3, 5]) = [1 : 9]$, so this gives one possible list A.

Suppose $a = 1$ and $b \geq 4$. Then $c \geq 4$ as well since $b \leq c$, but if $A = [1, 1, b, c]$ where $b \geq 4$ and $c \geq 4$, then $f(A)$ does not contain 3 which means $f(A) \neq [1 : 9]$.

So far, we have shown that $A = [1, 1, 3, 5]$ is the only list of the form we seek with $a = 1$ and $f(A) = [1 : 9]$.

We will now similarly examine the possibilities when $a = 2$.

If $a = 2$ and $b = 2$, then $c = 5$, so $A = [1, 2, 2, 5]$. It can be checked that $f(A) = [1 : 9]$ in this case.

If $a = 2$ and $b = 3$, then $c = 4$ and $A = [1, 2, 3, 4]$. It can be checked that $f(A) = [1 : 9]$ in this case.

If $a = 2$ and $b \geq 4$, then $a + b + c = 9$ implies that $c < 4$, but we are assuming that $b \leq c$, so this is impossible.

Therefore, the only possibilities with $a = 2$ are $A = [1, 2, 2, 5]$ and $A = [1, 2, 3, 4]$.

If $a \geq 3$, then $A = [1, a, b, c]$ has $b \geq 3$ and $c \geq 3$ as well, but this would prevent 2 from being in $f(A)$, so A cannot compress $[1 : 9]$ in this case.

Therefore, the only lists of length four that compress $[1 : 9]$ are $[1, 1, 3, 5]$, $[1, 2, 2, 5]$, and $[1, 2, 3, 4]$.

To see why no shorter list can compress $[1 : 9]$, we use the general observation that if A is a list of length k, then there are at most $2^k - 2$ integers in $f(A)$. This is because for each sum computed for $f(A)$, each of the k items in A is either included in the sum or it is not. This gives 2^k possible ways of computing a sum of some of the items in A. However, this count includes the sum of none of the items in A (all are excluded from the sum) and the sum of all of the items in A. Both of these sums are excluded from $f(A)$, so there are $2^k - 2$ ways to compute a sum to go in $f(A)$. We note that there could be multiple ways to express the same integer in $f(A)$ as a sum of items in A, which is why we can only say there are *at most* $2^k - 2$ integers in $f(A)$ – in practice, there could be fewer than $2^k - 2$.

If $k \leq 3$, then $2^k - 2 \leq 2^3 - 2 = 6$, and so if A is a list of length at most three, then $f(A)$ has at most 6 integers. Therefore, $[1 : 9]$, which has nine integers, cannot be compressed by a list of length less than four.

(b) At the end of the solution to part (a), we argued that if A is a list of length k, then there are at most $2^k - 2$ integers in $f(A)$. Since $2^6 - 2 = 62$ and $2^7 - 2 = 126$, we need $k \geq 7$ to achieve $2^k - 2 \geq 100$. Therefore, a list that compresses $[1 : 100]$ must have length at least seven. Now, we will provide a list of minimal length (seven) that compresses $[1 : 100]$.

Consider $A = [1, 2, 4, 8, 16, 32, 38]$. Note that 1 is in $f(A)$ and that the sum of all items in A is $1 + 2 + 4 + 8 + 16 + 32 + 38 = 101$. Since $f(A)$ has sums of some but not all of the items in A, the integers in $f(A)$ are at most 100. Rather than computing all of the sums, we will explain why $f(A) = [1 : 100]$ in a way that will give some insight for part (c).

We first consider the sums that are achievable without using the integer 38. The other integers in A are $1 = 2^0$, $2 = 2^1$, $4 = 2^2$, $8 = 2^3$, $16 = 2^4$, and $32 = 2^5$. The sums that can be obtained by adding some or all of these integers are exactly the integers from 1 through $63 = 2^6 - 1$. To get an idea of how this works, read about "binary expansions" or "binary representations" of integers. As an example, to represent 53 as a sum of powers of 2, first, find the largest power of 2 that is no larger than 53, which is 32. Then compute $53 - 32 = 21$ to get $53 = 32 + 21$. Now find the largest power of 2 that is no larger than 21, which is 16. Subtract to get $21 - 16 = 5$ or $21 = 16 + 5$. Now substitute to get $53 = 32 + 16 + 5$. Repeating this process, find the largest power of 2 that is no larger than 5, which is 4. Subtracting, $5 - 4 = 1$ so $5 = 4 + 1$, but what remains, 1, is a power of 2, so we get $53 = 32 + 16 + 4 + 1 = 2^5 + 2^4 + 2^2 + 2^0$.

The integers from 1 through 63 are all in $f(A)$. To write the integers from 64 through 100 as a sum of integers in A, notice that $100 - 38 = 62$, and so if $64 \leq m \leq 100$, then $m - 38 \leq 62$. To write such m as a sum of integers from A, compute $r = m - 38 \leq 62 < 63$, write r as a sum of the integers from A other than 38, then include 38 in the sum. For example, to see that 91 is in $f(A)$, compute $r = 91 - 38 = 53$, then use that $53 = 32 + 16 + 4 + 1$ to get that $91 = 1 + 4 + 16 + 32 + 38$.

The only thing left to check is that none of the sums described above require using all seven items in A. To see why this is not a concern, recall that the sum of all items in A is 101, so if we express an integer that that is no larger than 100 as a sum of items from A, then it cannot possibly use every item in A.

(c) The result of part (b) generalizes as follows. For every positive integer n, the minimum length of a list A that compresses $[1 : n]$ is $\lceil \log_2(n + 2) \rceil$. Notice that when $n = 100$, since $100 + 2 = 102$ is strictly between $2^6 = 64$ and $2^7 = 128$, we have $\lceil \log_2(100 + 2) \rceil = 7$, which agrees with the result from part (b).

We will prove that $\lceil \log_2(n + 2) \rceil$ is the minimum length of a list that compresses $[1 : n]$ and, in the process, prove that $[1 : n]$ is always compressible.

To see that $\lceil \log_2(n + 2) \rceil$ is the minimum length of a list that compresses $[1 : n]$ in general, we will show that a list of length less than $\lceil \log_2(n + 2) \rceil$ cannot possibly compress $[1 : n]$, and then we will construct a list of length exactly $\lceil \log_2(n + 2) \rceil$ that does compress $[1 : n]$.

Suppose A has length $k < \lceil \log_2(n+2) \rceil$. Since k and $\lceil \log_2(n+2) \rceil$ are both integers, this implies $k \leq \lceil \log_2(n+2) \rceil - 1$. For every integer x, it is true that $\lceil x \rceil - 1 < x \leq \lceil x \rceil$, so we conclude that $k \leq \lceil \log_2(n+2) \rceil - 1 < \log_2(n+2)$.

From $k < \log_2(n+2)$, we get $2^k < n+2$ or $2^k - 2 < n$. As argued earlier, a list A of length k has the property that there are at most $2^k - 2$ distinct integers in $f(A)$. Since $2^k - 2 < n$, we cannot have $f(A) = [1:n]$ when A is a list of length $k < \lceil \log_2(n+2) \rceil$ since $[1:n]$ contains n integers.

We have shown that if A compresses $[1:n]$, then it must have length at least $\lceil \log_2(n+2) \rceil$. We will now produce a list of length $\lceil \log_2(n+2) \rceil$ that compresses $[1:n]$. This requires explaining how to produce the list, then showing that the list has the correct length.

Suppose n is a positive integer. Define k to be the largest non-negative integer with the property that $2^k \leq n+1$ and define $m = n + 2 - 2^k$. The list A consisting of the powers of 2 from 1 through 2^{k-1} together with m will compress $[1:n]$ and have length exactly $\lceil \log_2(n+2) \rceil$. Notice that it is possible for m to be equal to one of the powers of 2 from 1 through 2^{k-1}. In this situation, the list A will include two copies of that power of 2.

Before verifying that the list described above does what is required, we will work through a couple of examples.

- When $n = 1$, we observe that $2^0 = 1$ and $2^1 = 2$ are the powers of 2 that are no larger than $n+1 = 2$, and so $k = 1$. Thus, $2^{k-1} = 1$ and $m = n + 2 - 2^k = 1 + 2 - 2 = 1$, so the list is $A = [1,1]$. Indeed $f([1,1]) = [1]$.
- When $n = 100$, we get $k = 6$, so $2^{k-1} = 32$ and $m = n + 2 - 2^k = 100 + 2 - 64 = 38$, so the list is $A = [1, 2, 4, 8, 16, 32, 38]$, which is exactly the list from part (b).

We will now show that list A has length $\lceil \log_2(n+2) \rceil$. The integer k is the largest non-negative integer with the property that $2^k \leq n+1$, and A contains 2^0, 2^1, and so on up to 2^{k-1}, along with the integer m. This gives a total of $k+1$ items in A. Therefore, it suffices to show that with k chosen as described above, we have $\lceil \log_2(n+2) \rceil = k+1$.

The function $\log_2$ is increasing, meaning that if x and y are positive real numbers with $x < y$, then $\log_2(x) < \log_2(y)$. Using this fact along with the fact that $2^k \leq n+1$, we get that $k \leq \log_2(n+1) < \log_2(n+2)$. As well, k is the largest non-negative integer with the property that $2^k \leq n+1$, which means $n+1 < 2^{k+1}$.

Suppose $k+1 < \log_2(n+2)$. Then $2^{k+1} < n+2$. From above, we also have $n+1 < 2^{k+1}$, so we conclude that $n+1 < 2^{k+1} < n+2$. The quantities $n+1$ and $n+2$ are consecutive integers, so the integer 2^{k+1} cannot lie strictly between them. Therefore, it is impossible for $k+1 < \log_2(n+2)$, which means we must have $\log_2(n+2) \leq k+1$. Combining this inequality with $k < \log_2(n+2)$, we have $k < \log_2(n+2) \leq k+1$, and so we conclude that $\lceil \log_2(n+2) \rceil = k+1$.

It remains to show that A compresses $[1:n]$. As discussed earlier, since list A contains the items 2^0, 2^1, and so on up to 2^{k-1}, as well as at least one other item, m, $f(A)$ contains all of the integers from 1 through $2^{k-1+1} - 1 = 2^k - 1$. This is because these integers can be expressed using the sum of some or all of the powers of 2 from 1 through 2^{k-1}. These are exactly the integers that can be expressed without using m.

If we do use m, then we can express every integer from m through $m + 2^k - 1$ as a sum of items in A. Since $m + 2^k - 1$ is the sum of all items in A, it is excluded from $f(A)$ and so we have that $f(A)$ contains all of the integers from m to $m + 2^k - 2$, and nothing larger. By definition, $m = n + 2 - 2^k$, so $m + 2^k - 2 = (n + 2 - 2^k) + 2^k - 2 = n$.

We have shown that $f(A)$ is the list consisting of the integers that are in $[1 : 2^k - 1]$ or $[m : n]$. To see that $f(A)$ is exactly $[1 : n]$, we need to show that $m \leq 2^k$. There might be overlap corresponding to multiple ways to express some integers in $[1 : n]$ in $f(A)$, but this is allowed.

Suppose $m > 2^k$. Since these quantities are integers, we must have $m \geq 2^k + 1$. By definition, $m = n+2-2^k$, and so we get that $n+2-2^k \geq 2^k+1$, which can be rearranged to get $n + 1 \geq 2^k + 2^k = 2^{k+1}$. Therefore, we have $2^{k+1} \leq n + 1$, but k was chosen to be the largest integer with $2^k \leq n + 1$, so it is not possible that $2^{k+1} \leq n + 1$. Therefore, our assumption that $m > 2^k$ must be wrong, so we conclude that $m \leq 2^k$, as desired.

(d) Fix a positive integer $k \geq 3$. We will show that for every integer $m \geq k$, $[m : m + k - 1]$ is not compressible.

To see this, suppose A is a list that compresses $[m : m + k - 1]$. Since $k \geq 3$, there are at least three items in $[m : m + k - 1]$. If A has only two items, then $f(A)$ has at most two integers since the allowable sums are just the two "singleton sums". Therefore, A also contains at least three items. Note that since m is the smallest integer in $f(A)$, m must also be the smallest integer in A. (If r is in A, then the "singleton sum" r must also be in $f(A)$, so all integers in A must be at least m. Also, since there is nothing smaller than m in A, the only way to produce m in $f(A)$ is by the "singleton sum" m.)

Since $k \geq 3$, $m + 1$ is also in $f(A)$. If $m + 1$ is the sum of at least two items in A, then they must all be smaller than $m + 1$, but the only integer in A that is smaller than $m + 1$ is m. This would mean 1 must be in A, but this is impossible since $1 < k \leq m$ and m is the smallest element in A. Therefore, we must also have $m + 1$ in A, and so both m and $m + 1$ are in A.

Now recall that A has at least three items, so there is at least one element other than m and $m + 1$, and so $m + m + 1 = 2m + 1$ is in $f(A)$. Since $f(A) = [m : m + k - 1]$, we must then have $2m + 1 \leq m + k - 1$, which can be rearranged to get $m \leq k - 2$, which is impossible since $m \geq k$.

Therefore, it is not possible to compress $[m : m + k - 1]$ when $m \geq k$. This means that there are only finitely many m for which $[m : m + k - 1]$ is compressible since all such m must be at most k.

(e) As mentioned in the hint, the answer is 39. This means $[5 : 39]$ is not compressible, but $[5 : k]$ is compressible for all $k \geq 40$. We will include a sketch of the proof here.

One can check that the lists $A = [5, 6, 7, 8, 9, 10]$, $B = [5, 5, 6, 6, 7, 8, 9]$, $C = [5, 5, 6, 7, 7, 8, 9]$, $D = [5, 5, 6, 7, 8, 8, 9]$, and $E = [5, 5, 6, 7, 8, 9, 9]$ compress the lists $[5 : 40]$, $[5 : 41]$, $[5 : 42]$, $[5 : 43]$, and $[5 : 44]$, respectively.

Now suppose a list A compresses $[5 : k]$ and suppose B is the list A with a 5 added to it. It can be shown that B compresses the list $[5 : k + 5]$. Since $[5 : 40]$ is compressible, this shows that $[5 : 45]$ is compressible. Since $[5 : 41]$ is compressible, $[5 : 46]$ is compressible. Since we have five consecutive values of k for which $[5 : k]$ is compressible ($k = 40$ through

$k = 44$), this reasoning can be used to show that $[5 : k]$ is compressible for all $k \geq 40$. Note that the lists to compress $[5 : k]$ given by this inductive process will not, in general, be as short as possible.

Now suppose a list A compresses $[5 : 39]$. It can be argued using reasoning similar to that from earlier parts that 5 must be in A, 5 is the smallest integer in A, and that the integers 6, 7, 8, and 9 also appear in A. Since the smallest integer in A is 5, the largest sum in $f(A)$ is 5 less than the sum of all items in A. Therefore, the sum of all items in A must be $39 + 5 = 44$. We already have 5, 6, 7, 8, and 9 in A which have a sum of $5 + 6 + 7 + 8 + 9 = 35$, and so the remaining items in A have a sum of $44 - 35 = 9$.

Next, note that using only the five items 5, 6, 7, 8, and 9, a sum of 10 is impossible. The list A cannot contain the integer 10 itself since we already determined that the remaining items in A have a sum of 9. It also cannot include the integers 1, 2, 3, or 4 since 5 is the smallest integer in A. Therefore, the 10 in $f(A)$ must come from the sum of two 5s, which means A must include a second 5.

We now have that A contains (at least) the items 5, 5, 6, 7, 8, and 9, which have a total of 40. Since the sum of all items in A is 44, there must be an additional item in A that is no larger than $44 - 40 = 4$. This is impossible, so we conclude that $[5 : 39]$ is not compressible.