An Interesting Way to Combine Numbers

Joshua Zucker and Tom Davis
October 12, 2016

Abstract

This exercise can be used for middle school students and older. The original problem seems almost impossibly difficult, but there are obviously many ways to approach it by considering simpler problems. As they investigate it, students are actually drilling their multiplication and addition facts. In addition, depending on how detailed a study you want to make, the students will be forced to do a lot of simple algebra to obtain the results they need. They may also learn something about commutativity, associativity, and symmetric functions.

1 Introduction

This document is meant for the teacher. It describes an interesting problem and then lists a number of ways the problem can be used in a classroom. Depending on the age and sophistication of the students, the classroom discussion can be taken in different directions.

As usual, if you are the teacher, you will probably find this document more useful if, before reading our discussion about classroom presentation, you try to make some headway on the problem yourself. It will help you to "think like a student" and you may come up with additional ideas and strategies that did not occur to the authors.

2 The Problem

Imagine that all the numbers from 1 to 100 inclusive are written on the blackboard. At every stage, you are allowed to erase two numbers that appear on the board (let's call the numbers you erased x and y) and in place of the two erased numbers, write the number $x + y + xy$. Repeat this operation until only a single number remains.

What are the possible values for that remaining number?

3 Classroom Presentation Strategies

Obviously you do not need to list all the numbers from 1 to 100. But write down a few of them, possibly something like this:
$$1, 2, 3, 4, 5, 6, 7, 8, \ldots, 98, 99, 100$$

Also, depending on the sophistication and level of the class, you may not want to give the algebraic form $x + y + xy$; just say that the new number is obtained by adding three numbers: the two original numbers and their product. Do a couple of examples, as follows:

Suppose the two numbers you choose are 2 and 4, The new number is $2 + 4 + 2 \times 4 = 14$. On the blackboard, physically erase the 2 and the 4 and write down the number 14. Point out that the board now has two copies of the number 14, and that there's nothing wrong with that. The two 14's could even be combined at that point, yielding $14 + 14 + 14 \times 14 = 224$. At every stage throughout your presentation, make sure the *students* do the calculations.

Get the students to explain to you why the process must come to an end. (You want them to say something like "At every stage two numbers are erased and only one number is added, so after each step there is one fewer number. Since we started with only 100 numbers, after 99 steps there will remain only a single number.") This is a "monotonicity argument": we can assign a positive number to a situation (in this case, simply the number of numbers remaining on the board) and show that at each stage this number is reduced (in this case, by one), and thus the process must eventually come to an end since once we are down to a single number the process must halt.

Depending on the level of the class, you may need to work out one or two more examples of how the process works in general. You could show, for example, that if you begin with the numbers $1, 2, 3$ and 4, that the process ends after three steps. Not only will this drive home the fact that all such processes have to end, but it will remind the students a few more times exactly how numbers are combined.

If you work with $\{1, 2, 3, 4\}$ as your initial set, don't choose pairs to eliminate in any straightforward order. For example, first combine 4 and 2 to yield 14. Then combine 1 and 3 to yield 7. Finally combine the 7 and 14 to yield the final result of 119.

Ask the students if they have any ideas for how to proceed. Ask them if they can guess what strategies might lead to a large final number and which strategies might lead to a small final outcome.

4   How to Investigate the Problem

It's pretty clear that it would be crazy to start work on the full problem that specifies that we begin with 100 different values on the blackboard. There are a *tremendous* number of ways that pairs of numbers can be selected and every one of them might yield a different final outcome.

4.1   A Smaller, Simpler Problem

A good strategy in situations like this is to invent a similar, smaller problem that we can look at. We can look at problems that are both smaller (fewer than 100 initial numbers) or simpler. The word "simpler" can be a little misleading. Obviously we'd like to try some examples with numbers that are easy to calculate with. This might mean small numbers, like $1, 2, 3$, or perhaps look at the situation where all the initial numbers are the same (say $1, 1$ and 1). What would happen if we allow the number 0 to be included?

We can gain some insight into the problem with examinations like this, but later on we will also find that looking at the general algebraic form of the combining operation will actually be, in some senses, simpler, since it will show us how exactly how the original numbers fit into combinations, especially when we combine more than two of them.

Ask the students to suggest simpler related problems, and you may get suggestions like, "Just list the numbers from 1 to 10." That's a good suggestion, but try to convince them that even 10 is a very large

number, perhaps by starting a calculation and seeing how many possibilities there are.

Ask them, "What is the *easiest* problem?" Starting with one number is probably the easiest, since at that point, the game is over, and the score is simply whatever number you started with.

What happens if there are two numbers, say 1 and 2? You have no choice but to pick those two numbers, and the result will be $1 + 2 + 1 \times 2 = 5$. Eventually you want to write down the algebraic form for the combination, but depending on the class, you might want to do more examples with concrete numbers first: "What's the result of combining 4 and 5? Of combining 1 and 100? Of 100 and 100? Of x and y? Be sure that you do eventually get them to state the algebraic notation for the combination operation: $x + y + xy$.

Ask the students to use the formula $x + y + xy$ to figure out what number would replace, say, 3 and 5. When somebody gives the correct answer $3 + 5 + 3 \times 5 = 23$, ask them how they did it. They'll probably say something like, "$x = 3$ and $y = 5$, so I just plugged those numbers into the formula." Ask them how they knew that it was the x that was 3. What if x were the 5 and y were the 3?

Obviously, you get the same answer, but ask the students why? What is it about the formula $x + y + xy$ that makes it give the same answer if you swap the values of x and y? The answer, of course, is that if you put an "x" everywhere there's a "y" and a "y" everywhere there's an "x" (in other words, swap x and y) then the algebraic form is the same. In our simple example, swapping x and y would yield $y + x + yx$ which is equal to $x + y + xy$. If our rule for combining two numbers x and y by another had been to replace them with, say, $2x + y + xy$, then the order *would* matter, so choosing 5 for x and 3 for y would yield a different result than choosing 3 for x and 5 for y.

Symmetric functions: If there's interest and the class has a little algebraic sophistication, you can examine this idea of a "symmetric" function – a function whose inputs can be permuted and yet yield the same value. Here are some symmetric functions. Note that in functions with more than two variables, the inputs can be rearranged in any way you want with no change to the output. Mathematically, this can be stated as $f(x, y) = f(y, x)$, or

$$f(x, y, z) = f(x, z, y) = f(y, x, z) = f(y, z, x) = f(z, x, y) = f(z, y, x).$$

There would be 24 equations to illustrate all the equalities required for a symmetric function of 4 variables, and in general, $n!$ equalitites for a function of n variables. Check that these are symmetric and try to find others.

$$
\begin{aligned}
f(x, y) &= xy \\
f(x, y) &= x^2 - 3xy + y^2 \\
f(x, y) &= x^2 y + y^2 x + x + y - 17x^2 y^2 \\
f(x, y, z) &= xyz + xy + yz + zx \\
f(x, y, z) &= (x + y)(y + z)(z + x) \\
f(x, y, z) &= (x + 1)(y + 1)(z + 1) - 1 \\
f(x, y, z) &= \sqrt{\frac{(x + y + z)(x + y - z)(x - y + z)(-x + y + z)}{16}} \\
f(x, y, z, w) &= 1
\end{aligned}
$$

Because the replacement function is symmetric, the first interesting case to consider is when we begin with three numbers, since we can choose any pair to begin with. Suppose the three numbers are 1, 2 and

3. Then we can begin with 1 and 2, with 1 and 3, or with 2 and 3. After that choice is made the next one is fixed since only two numbers will remain. Have the students work out the answers for each of the three initial choices, and they *should* discover that all yield the same final result; namely, 23. It's easy to find the errors; anyone who gets an answer for any of the examples that is not 23 has made one!

At this point there is a little evidence that the final answer is independent of the starting values, but could that be simply because we chose the simple sequence 1, 2, 3? Test that by starting with a different set; say, 3, 7 and 8. (Doing this will cause the students to drill their arithmetic.) In this case, the final results will again be independent of the order that the numbers are combined, and the final number will always be 287. You can ask the class if there is any way to show that the answer is independent of the order for any three starting numbers. You're looking for an answer that involves algebra and variable names for the three numbers, where the initial numbers are combined in different orders.

Now you can start looking at what happens with four initial numbers, 1, 2, 3 and 4, but you can impress upon the students that there are a *lot* of possibilities to be considered. The first choice can be made in 6 ways and then there are 3 ways to make the second choice, for a total of 18 different ways. Your students should work out a few examples (all of which should yield 119 as the final number), but even now it should be clear that checking all of them would require a lot of work, and for the case of 100 starting numbers, checking all possibilities would be impossible. Some other approach must be found.

Commutativity and Associativity of Operations

When students are taught about commutativity and associativity, it always seems to be in the context of the operations of addition and multiplication that they know so well that it may seem to them almost stupid to have names for these "obvious" properties. (Of course they could be told about non-commutative and non-associative operations like subtraction and division, but somehow that's hardly ever done.) The great thing about this problem is that we have effectively defined a new operation for putting two numbers together – adding both numbers to their product – that is just another binary operation that takes two numbers as input and yields a single number as output. What follows shows a way to lead the students to look at this new, unusual operation and to try to find its properties. (It will turn out to be both commutative and associative, and knowing that, we can show symbolically that the final result is completely independent of the order in which the numbers on the blackboard are combined.)

When specific calculations (like the ones that use specific numbers) do not shed enough light on the problem, sometimes it's easier to look at the general case. For example, suppose there are three numbers on the board: a, b and c. Can we work out the result of doing them in different orders? For example, first combine a and b, then combine that result with c. Then do the same thing, but choose b and c first, then combine the result with a. Here are the results of both of those calculations.

The first step is easy: combining a and b yields $a + b + ab$. To combine that with c we obtain:

$$(a + b + ab) + c + (a + b + ab)c \tag{1}$$
$$= \quad a + b + c + ab + ac + bc + abc \tag{2}$$

It might be clearer to the students if you do this in two steps as follows. First, give a name, like x, to the quantity $a + b + ab$, so when you combine x with c, you first get $x + c + xc$, and then if you substitute $a + b + ab$ for x, you obtain the first line in the pair of equations above.

If b and c are combined first, we obtain $b + c + bc$ and then we combine that with a to obtain:

$$(b + c + bc) + a + (b + c + bc)a \tag{3}$$
$$= \quad a + b + c + ab + ac + bc + abc \tag{4}$$

which is exactly the same result, and a symmetric expression in a, b and c as well!

This means that if you have three numbers, the order in which you combine them makes no difference. The same calculation could be done for four numbers, et cetera, but again, the calculations start to become unmanageable. With four numbers it is not totally intractable, and if you wish to have the students do it, the result should be the following 15-term symmetric expression (where the initial list contains a, b, c and d):

$$a + b + c + d + ab + ac + ad + bc + bd + cd + abc + abd + acd + bcd + abcd.$$

A slightly sophisticated symmetry argument. In the calculations above, we calculated the value resulting from combining three numbers in two different orders. In fact, a slightly sophisticated symmetry argument shows us that this is true after only a single calculation. Equation 2 was calculated by combining first a and b, and then the result was combined with c. In the next line we recalcuated in a different order and obtained the same result, but that second calculation was unnecessary, in a sence, since the original result in Equation 2 is symmetric and there is no difference in the way that a, b and c enter into that calculation. Since the resulting equation is symmetric, that means that the inputs can be permuted any way you want, and in this case, the "input" basically specified the order of combination of initial numbers. Because of the symmetry, we can do the input in different orders and obtain the same result, so the second calculation to obtain Equation 4 was unnecessary. This argument may be too complicated for many students, but it is an important idea.

At this point, we can take one more step toward generality that will simplify notation and perhaps improve comprehension.

In the preceeding, we have repeatedly used expressions like "combine a and b". In reality, it is just an arithmetic operation that doesn't have a common name or symbol. We will simply invent a new symbol, "$\oplus$" that represents our new operation. In the same way that $a + b$ uses the symbol "$+$" to indicate the addition of a and b, we will use our new symbol, "$\oplus$" to indicate our combining operation. Symbolically, $a \oplus b = a + b + ab$. If you consider the calculations we did in equations 2 and 4, since a, b and c were arbitrary numbers, we have:

$$(a \oplus b) \oplus c = (b \oplus c) \oplus a. \tag{5}$$

We also noted earlier that:

$$a \oplus b = b \oplus a. \tag{6}$$

In other words, the $\oplus$ operator is commutative. Because of this commutativity, we could rearrange the terms in equation 5 as follows (which puts it in the usual form used to express the result that the operation is associative):

$$(a \oplus b) \oplus c = a \oplus (b \oplus c). \tag{7}$$

Equations 6 and 7 say that our new operation $\oplus$ is both commutative and associative. Since it's commutative and associative, it doesn't matter in what order the numbers are selected to be combined; the result is always the same.

At this point, there are a couple of ways to go, depending on the sophistication of the class. The easy way is just to note that "Associativity means that you don't need parentheses" and "Commutativity means that you can swap the order of the operands any way you want". With these two ideas, it's clear that any sequence that combines any number of initial numbers leads to the same result, and that may be good enough and you should skip the following section.

4.2   Commutativity and Associativity: Detailed Drill

One problem students have with "general" formulas like:

$$A \oplus (B \oplus C) = (A \oplus B) \oplus C$$

or even simple formulas like the Pythagorean theorem:

$$A^2 + B^2 = C^2$$

is that it is hard for them to get the idea that the A, B, and C can represent *general* expressions. For Pythagorean triples, it may be obvious to students that if you know that legs of a right triangle have lengths $A = 3$ and $B = 4$ that they can plug those numbers in to obtain a formula for C, but they don't know how to find C in a triangle where the lengths of the legs are given as $A = 1 + u^2$ and $B = 1 - u^2$ like this:

$$C^2 = (1 + u^2)^2 + (1 - u^2)^2 = 2 + 2u^4,$$

so $C = \sqrt{2 + 2u^4}$.

To step through the details for rearranging formulas involving the operation $\oplus$ using the primitive formulas for commutativity and associativity, students will have to understand and apply this idea over and over.

If you are interested in drilling the ideas of commutativity and associativity, you can go through an example with combining, say, five numbers to show that commutativity and associativity implies that two different combining orders yield the same result, as follows:

Suppose the five initial numbers are a, b, c, d and e. Your claim is that no matter how the numbers are combined, the result will be the same as combining d with e, then that number with c, then that number with b, and finally, that number with a. Suppose the students say this:

Combine b with d (call that x). Next combine c with e (call that y). Now combine y with a (making z) and then z with x (producing the final number N). You probably want to give names to each step as the students make them up, like this:

$$
\begin{aligned}
x &= b \oplus d \\
y &= c \oplus e \\
z &= y \oplus a \\
N &= z \oplus x
\end{aligned}
$$

You can then work backwards to find a mathematical expression for N. Begin with the final equation above, and do one substitution at each step. In the list below we substitute first for x, then for z, then for

y to obtain each successive line:

$$\begin{aligned} N &= z \oplus x \\ N &= z \oplus (b \oplus d) \\ N &= (y \oplus a) \oplus (b \oplus d) \\ N &= ((c \oplus e) \oplus a) \oplus (b \oplus d) \end{aligned}$$

Make sure the students see that this final equation expresses in one line exactly the operations that they suggested.

Next, we can take the final equation that involves only our original variables and convert it, step by step, to our "target" form:

$$N = a \oplus (b \oplus (c \oplus (d \oplus e))).$$

Here are the steps that will do it (see below for the detailed explanation for each step):

$$\begin{aligned} N &= ((c \oplus e) \oplus a) \oplus (b \oplus d) & (8) \\ N &= (a \oplus (c \oplus e)) \oplus (b \oplus d) & (9) \\ N &= a \oplus ((c \oplus e) \oplus (b \oplus d)) & (10) \\ N &= a \oplus ((b \oplus d) \oplus (c \oplus e)) & (11) \\ N &= a \oplus (b \oplus (d \oplus (c \oplus e))) & (12) \\ N &= a \oplus (b \oplus ((c \oplus e) \oplus d)) & (13) \\ N &= a \oplus (b \oplus (c \oplus (e \oplus d))) & (14) \\ N &= a \oplus (b \oplus (c \oplus (d \oplus e))) & (15) \end{aligned}$$

We will see that each equation is derived from the previous using only associativity and commutativity of the operation $\oplus$. We'll use the following shorthand. If we apply commutativity, we effectively convert $X \oplus Y$ to $Y \oplus X$. This will be represented as in the second line below, where X stands for a and Y stands for $(c \oplus e)$. If associativity is involved, we will convert $A \oplus (B \oplus C)$ to $(A \oplus B) \oplus C$ (or the reverse). The easiest one to see is in the seventh line below. Notice the strategy: we use commutativity to bring the letters into the order a, b, c, d, e, and associativity to group them the way we want.

Equation 8: The original equation.
Equation 9: Commutativity: $X = a$ and $Y = (c \oplus e)$
Equation 10: Associativity: $A = a$, $B = (c \oplus e)$, $C = (b \oplus d)$
Equation 11: Commutativity: $X = (c \oplus e)$, $Y = (b \oplus d)$
Equation 12: Associativity: $A = b$, $B = d$, $C = (c \oplus e)$
Equation 13: Commutativity: $X = d$, $Y = (c \oplus e)$
Equation 14: Associativity: $A = c$, $B = e$, $C = d$
Equation 15: Commutativity: $X = d$, $y = e$

5 The Final Result

Now that we know that the order in which we combine the original numbers on the board makes no difference, we can compute a few simple examples by using simple arithmetic. It's not too hard to do this

(and the class should) work out values for initial sets containing up to six natural numbers, $\{1\}$, $\{1, 2\}$, $\{1, 2, 3\}$, ..., $\{1, 2, 3, 4, 5, 6\}$. The following table lists the results they should obtain:

Initial List	Final Value
1	1
$1, 2$	5
$1, 2, 3$	23
$1, 2, 3, 4$	119
$1, 2, 3, 4, 5$	719
$1, 2, 3, 4, 5, 6$	5039

As you help with the calculations, you can point out a couple of things if the students themselves don't notice them. First, as you're working on the first two or three, point out that since the order doesn't matter, students can use whatever numbers they want that seem like they'll make for easier calculations. Then, say after the value 23 is computed for the third example, point out that rather than starting from scratch for the fourth example, note that they've already combined $1, 2$ and 3 and to obtain the next number, they just need to combine the 23 already obtained with 4 to obtain: $23 + 4 + 23 \times 4 = 119$. Given this idea, the students may want to calculate a couple of more steps. (If they do, the answers for initial lists containing the first $7, 8$ and 9 whole numbers are given by $40319, 362879$ and 3628799, respectively.) Sophisticated students may from this information alone be able to guess the general formula, which, for the initial list $\{1, 2, 3, \ldots, n\}$ is given by $(n + 1)! - 1$.

Even if you've guessed the formula, it still may be difficult to write down a convincing argument that it is correct. One approach is this:

5.1 Wishful Thinking

When we see the formula $x \oplus y = x + y + xy$, we may notice that the form on the right can *almost* be factored. If only there were an additional 1 on the right, we'd be in business, since

$$(1 + x)(1 + y) = 1 + x + y + xy.$$

Well, just add a 1! But you can't do that without changing the value of the expression, so you'd better subtract a 1 to make up for it:

$$x \oplus y = 1 + x + y + xy - 1 = (1 + x)(1 + y) - 1.$$

As we shall see, this is a *much* better form of the algebraic expression that represents $x \oplus y$.

If we use this form and want to calculate $x \oplus y \oplus z$, let's see what happens. (Note that because of associativity, we don't need to worry about parentheses in our "$\oplus$" expressions.)

$$
\begin{aligned}
x \oplus y \oplus z &= ((1 + x)(1 + y) - 1) \oplus z \\
&= (1 + (1 + x)(1 + y) - 1)(1 + z) - 1 \\
&= (1 + x)(1 + y)(1 + z) - 1
\end{aligned}
$$

Point out where the pair "1" and "−1" inside the initial set of parentheses in the second line above comes from – that when we apply $\oplus$ to two numbers, we add one to each, multiply them, and subtract 1 from the

8

result. One of the numbers is $(1+x)(1+y)-1$ and the other is z. Adding 1 to both gives $(1+x)(1+y)$ and $(1+z)$, multiplying them gives $(1+x)(1+y)(1+z)$, and subtracting 1 gives the result.

Lead the students through another couple of steps exactly as above to show that:

$$
\begin{aligned}
x \oplus y \oplus z &= (1+x)(1+y)(1+z)-1 \\
x \oplus y \oplus z \oplus w &= ((1+x)(1+y)(1+z)(1+w)-1 \\
x \oplus y \oplus z \oplus w \oplus v &= ((1+x)(1+y)(1+z)(1+w)(1+v)-1
\end{aligned}
$$

It would be possible, of course, to write down a formal proof by mathematical induction that the general formula holds, but just doing the operation a few times should make the general pattern clear. The general pattern is that you add 1 to every number in the list, multiply all of those numbers together, and subtract one from the final result.

Thus the result obtained from an initial list containing $\{1, 2, 3, 4, 5, 6, 7\}$ is given by:

$$(1+1) \times (2+1) \times (3+1) \times (4+1) \times (5+1) \times (6+1) \times (7+1) - 1$$

which is:

$$2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 - 1 = 8! - 1.$$

Clearly, the result for an the initial list from 1 to 100 always yields $101! - 1$, which, in case anyone wants to multiply it out, is the following 160-digit number:

$$94259477598383594208516231244829367495623127947025437683278893534169775993162214765030878615918083469116234900035495995833697063026032639999999999999999999999999999$$