Identities and Inverses, Part 2

The identity function and inverses of functions as examples of the concepts of identity and inverse.

In a previous activity, we explored identities and inverses through a careful process that involved four steps:

(a) Specifying the objects,

(b) Specifying the operation,

(c) Defining the identity with respect to that operation on those objects, and

(d) Defining the meaning of inverse with respect to that operation on those objects.

In this activity, we explore identities and inverses for functions.

Remark 1. *In the Common Core State Standards (CCSS), the expectations regarding inverses of functions are quite modest for the audience of all students, requiring only that they "Solve an equation of the form $f(x) = c$ for a simple function f that has an inverse and write an expression for the inverse (F-BF.4.a)." Students in Precalculus and Calculus need more, of course.*

These notes and questions are intended to help teachers understand these subtle ideas well enough to make wise instructional decisions for both populations of students.

Functions as Objects

To begin our exploration, we first decide that *functions are the objects* under consideration. Thinking of functions as objects can be something of a challenge, however, because much classroom experience emphasizes formulas and computing particular function values.

The CCSS includes the following standards about functions:

- 8.F.1. Understand that a function is a rule that assigns to each input exactly one output. The graph of a function is the set of ordered pairs consisting of an input and the corresponding output.

Learning outcomes:
Author(s):

- F-IF.1. Understand that a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range. If f is a function and x is an element of its domain, then $f(x)$ denotes the output of f corresponding to the input x. The graph of f is the graph of the equation $y = f(x)$.

Question 1 *Write a definition of* function *that you used in an upper-level undergraduate mathematics class. Explain how your definition is or is not consistent with the above definitions.*

Free Response: ***Hint:*** *Possible answer: For any sets A and B,* a function f from A to B *is a subset of the Cartesian product $A \times B$ such that every $a \in A$ appears exactly once as the first element of an ordered pair (a, b) in f. When $(a, b) \in f$, the notation $f(a)$ means the corresponding b.*

In the less formal definitions from the CCSS, A is the set of 'inputs' or the elements of the domain. The 'outputs' are the elements of the range, which is a subset of B.

Question 2 *Many students think of function as synonymous with formula. Describe some advantages to thinking of functions as broader than formulas.*

Free Response: ***Hint:*** *Possible answer: In many modeling situations, formulas are not suitable. For example, suppose $T(n)$ denotes the high temperature at an Arps Hall weather station on the n^{th} day of the year.*

Question 3 *Suppose g is a function, $g(3) = 5$ and $1 = g(3) + g(a)$. What can you say about $g(a)$? What can you say about a? About g?*

Free Response: ***Hint:*** *From the two equations, we can conclude that $g(a) = -4$, but we know nothing else about a and nothing else about g.*

To think of a function as an object, it helps to imagine the function as a whole. Saying "f is the squaring function" is a statement about the whole function, whereas $f(3) = 9$ is a statement about a specific function value. In general, a stand-alone letter, such as f, is used to indicate the function as a whole, whereas $f(a)$ indicates the function value for the particular input value a.

To think about the whole function, imagine varying through all possible input values. The whole function is not just the output values but rather all the correspondences between the input values and output values. Thus, if the domain of f is D, we may write $f = \{(x, y) | x \in D, y = f(x)\}$, which is to say the function f is this set of ordered pairs.

A Note on Notation

Mathematicians and teachers are sometimes sloppy regarding the notational distinction between a function value and the function as a whole, allowing $f(x) = x^2$, for example, to be taken as a statement about the whole function.

Consider the following expressions: $f(a)$, $f(x_0)$, and $f(x)$. Without any additional context, many mathematicians and teachers interpret the first two as particular output values, because it is customary to use the letter a and the subscripted x_0 to denote particular values, considered one at a time and conceived as "fixed" while reasoning through a problem. The expression $f(x)$, on the other hand, is ambiguous, for it sometimes denotes a particular output value, yet other times denotes the function as a whole.

Some mathematicians occasionally rail at the use of $f(x)$ for the function as a whole, while others are content that the meaning is usually clear from the context. When specifying a function, some authors and computer algebra systems avoid the ambiguous "specification formula" $f(x) = x^2$ and instead use the notation $f : x \mapsto x^2$, which can be read, "f maps x to x^2."

This distinction and the "maps to" notation are likely too subtle when high school students are first learning function notation, because students already have plenty of difficulty with simple uses of the notation. The distinction can be useful in calculus, however, and it becomes necessary in upper-level undergraduate mathematics courses. And it is important that teachers understand the distinction because some of their students' difficulties will involve this issue.

Specifying the Operation

Once we consider function to be objects, we can specify an operation for combining such objects. To reach the goal of discussing inverses of functions, we must agree that *the operation is function composition.* The open circle symbol, $\circ$, is often used to indicate function composition, so that $f \circ g$ indicates the composition of functions f and g, taken as whole objects. The expression $f(g(x_0))$, in contrast, is about particular function values. The statement $(f \circ g)(x) = f(g(x))$ indicates how the two ideas are related.

Question 4 *Suppose $f : x \mapsto x^2$ and $g : x \mapsto x + 3$. Compute $f \circ g$ and $g \circ f$. What do you notice?*

Free Response: ***Hint:*** *First, $f(g(x)) = f(x + 3) = (x + 3)^2$. So $f \circ g : x \mapsto (x + 3)^2$.*

Second, $g(f(x)) = g(x^2) = x^2 + 3$. So $g \circ f : x \mapsto x^2 + 3$.

Note that $f \circ g \neq g \circ f$, which is to say that function composition is not commutative.

Defining Identity and Inverse

At last, we can define identity function. To keep the discussion within the realm of school mathematics, let's consider only real-valued functions of a real variable. In other words, both the input and output values are assumed to be real numbers, so that both the domain and the range are subsets of the real numbers.

Definition 1. *A function I is said to be an* identity function *on a domain D if $f \circ I = f$ and $I \circ f = f$ for any function f with domain D. Note that these are statements about whole functions.*

Question 5 *Describe the similarities and differences between this definition of identity function and your definition of multiplicative identity.*

Free Response: ***Hint:*** *It is essentially the same idea, with appropriate substitutions. Replace function composition with multiplication, I with 1, and f as any function with x as any number.*

Question 6 *Restate the definition of identity function so that it involves statements about function values.*

Free Response: ***Hint:*** *If function I is called the identity function on domain D if $f(I(x)) = f(x)$ and $I(f(x)) = f(x)$ for all $x \in D$ and for any function f with domain D.*

Question 7 *What is the identity function on the real numbers? Call it I.*

Free Response: ***Hint:*** $I(x) = x$.

Before we define the inverse of a function, we must first acknowledge that not all functions have inverses. We will address this issue in more detail later in the course. For now, let's restrict our attention to a particular collection of invertible functions: those that are both one-to-one and onto a domain D that is a subset of $\mathbb{R}$.

Remark 2. *If the function $f : A \to B$ is not one-to-one, it can be made one-to-one by restricting its domain to a subset X of A. To ensure the function is onto, let $R = f(X)$, the actual range of the restricted function. Then the function $g : X \to R$ given by $x \mapsto f(x)$ for all $x \in X$ is both one-to-one and onto, and hence it is invertible.*

Definition 2. *Suppose $f : D \to D$ is one-to-one and onto. A function $g : D \to D$ is the* inverse *of f if $g \circ f = I$ and $f \circ g = I$. Note that f is then also the inverse of g, and hence f and g are* inverse functions *in the sense that they are inverses of each other.*

Question 8 *Describe the similarities and differences between this definition of the inverse of a function and your definition of multiplicative inverse.*

Free Response: ***Hint:*** *It is essentially the same idea, with appropriate substitutions. Replace function composition with multiplication, I with 1, f with x, and g (the inverse of f) with y (the multiplicative inverse of x).*

Remark 3. *This analogy between multiplicative inverse and inverse of a function is what is behind the confusing notation f^{-1} for the inverse of a function. If a is a nonzero real number, then a^{-1} denotes the inverse of a with respect to multiplication. Similarly, if f is an invertible function, then f^{-1} means the inverse of f with respect to function composition.*

We should apologize to students for the fact that in $\sin^2 x$, the exponent is about multiplication, whereas in $\sin^{-1} x$, the exponent is about function composition. Because these notations are incompatible, some mathematicians and teachers use the notation $\arcsin x$ instead.

Question 9 *Give a definition of* inverse function *that involves statements about function values.*

Free Response: ***Hint:*** *Given a function $f : D \to D$, a function $g : D \to D$ is the inverse of f if $f(g(x)) = x$ and $g(f(x)) = x$ for all $x \in D$.*

A joke. After writing two matrices on the board, a professor asks a student, "Are these matrices inverses?" The student answers, "The first one is, and the second one isn't."

Question 10 *Rewrite the joke as a joke about functions.*

Free Response: ***Hint:*** *After writing two functions on the board, a professor asks a student, "Are these functions inverses?" The student answers, "The first one is, and the second one isn't."*

Question 11 *Suppose a function composed with itself is the identity function. What can you say about the inverse of the function? Can you think of such a function?*

Hint: *Call the function f. Compare the following: (1) that f composed with itself is the identity function, and (2) the definition of the inverse of f.*

Free Response: ***Hint:*** *This is subtle and worth some thought.*

Examples of Inverses of Functions

Addition is a "binary operation," which is to say it takes two numbers (as inputs) and returns their sum (as output). In other words, addition of real numbers is a function with the domain $\mathbb{R} \times \mathbb{R}$, pairs of real numbers, and the range $\mathbb{R}$. How, then, should we think about the oft-heard claim that "Subtraction is the inverse of addition"?

When considering the expression $5+3$, we can "undo" the addition by subtracting 3, and we get back to where we started, which happens to be at 5 but which could have been any number. In other words, we are thinking of the "add 3" function $f : x \mapsto x + 3$.

Question 12 *Use these ideas to determine the inverse of the function $f : x \mapsto x+3$? Use the definition of inverse function to demonstrate that you are correct.*

Free Response: ***Hint:*** *The inverse of f is the the "subtract 3" function $g : x \mapsto x - 3$.*

Check that $g \circ f = I$ as follows: $g(f(x)) = g(x + 3) = (x + 3) - 3 = x$.

Check that $f \circ g = I$ as follows: $f(g(x)) = f(x - 3) = (x - 3) + 3 = x$.

Note the way in which g "undoes" or "reverses" the effect of f. If we know the output of f was 12, we can compute $g(12) = 9$ to know that the input must have been 9.

Question 13 *Suppose $f : x \mapsto 2x - 5$. Use verbal descriptions of f and the idea of "undoing" to determine the inverse of f. Explain your reasoning.*

Free Response: ***Hint:*** *Think of f as "multiply by 2 then subtract 5." To undo that, we would "add 5 then divide by 2." So $g : x \mapsto \dfrac{x+5}{2}$ is the inverse of f.*

Remark 4. *Notice that figuring out the inverse of f does not require that misunderstood procedure that begins by swapping x and y.*

When someone says that "Subtraction is the inverse of addition," the point is more general than what we have just seen. We are really talking about a family of addition functions, $f_a : x \mapsto x + a$, where a is a parameter that indicates a particular addition function, each of which has its own inverse, $g_a : x \mapsto x - a$. We need the following definition:

Definition 3. *A* family of functions *is a parametrized set of related functions.*

Much of high school mathematics is organized around particular families of functions, such as linear functions, quadratic functions, exponential functions, or trigonometric functions. The quadratic functions, for example, are the family $f(x) = ax^2 + bx + c$, where a, b, and c are real numbers and $a \neq 0$. The parameters a, b, and c, when conceived as "fixed," indicate which particular function we are talking about. Note that "trigonometric functions" are not a family in the sense of the above definition.

Question 14 *In the family of quadratic functions described above, why is it important to state that $a \neq 0$?*

Free Response: ***Hint:*** *If $a = 0$ the function is not quadratic but linear.*

The concept of a family of functions is much more varied than the standard families that populate high school mathematics. For example, in problem-solving situations, it might be useful to consider the family of cubic functions with zeros at 0 and 1, which could be parametrized as $f : x \mapsto ax(x-1)(x-r)$.

Question 15 *Use families of functions to explain the statement "division is the inverse of multiplication."*

Free Response: ***Hint:*** *Consider the family of "multiply by a" functions, $f_a : x \mapsto ax$, where $a \neq 0$. The inverse of f_a is the "divide by a" function $g : x \mapsto \dfrac{x}{a}$.*