2. Recursion :: Five Key Concepts :: Key Concept 1—Reduce and Rephrase

Recursion involves several key concepts.

Notice that in finding the solution to P(5):

$P(5) = 5! = 5 \times P(4)$

we **rephrased** our original problem as a **new, but basically equivalent problem**—what is P(4)?—by **reducing** the "size" of our problem. In regards to the factorial problem, what we mean by "size" is this:

P(4) is a smaller problem than P(5) because computing the factorial of 4 takes less work (or is easier) than to calculate the factorial of 5 ($1 \times 2 \times 3 \times 4$ requires one fewer multiplication than $1 \times 2 \times 3 \times 4 \times 5$).

When a problem of some "size" can be rephrased as a new, but equivalent "smaller" problem then recursion may be useful in solving the larger problem (not always, other problem-specific characteristics have to be present as well).

2. Recursion :: Key Concept 2—Repetition

How did we find P(4)? We defined P(4) as:

$$P(4) = 4! = 4 \times 3! = 4 \times P(3)$$

which was exactly how we defined P(5)

$$P(5) = 5! = 5 \times 4! = 5 \times P(4)$$

albeit with different problem parameters (4 and 3 versus 5 and 4). In other words, the **operations** we performed to compute P(4) were **identical** to the operations that we performed to compute P(5)—only the parameters varied.

Recursion involves repetition. During a recursive procedure we repeatedly reduce the problem parameters to smaller and smaller values (which effectively reduces the "size" of the problem), while performing the same operations in each new problem.

2. Recursion :: Key Concept 3—Recursion Must Stop

When we reached the problem P(1) we stopped because the size of our problem had been reduced to the smallest possible value for which it possible to determine the factorial (ignoring $0! = 1$ again).

In a recursive process, when we determine that it makes no sense to continue reducing our problem to smaller basically-equivalent problems, this situation is referred to as the **base case**. For the factorial problem, the base case is:

$$\mathrm{P}(1) = 1! = 1$$

All recursive procedures must have a base case (there *could* be more than one, but there has to be at least one). Why? Because without a base case you would continue reducing your problem to smaller and smaller sizes ad infinitum.

2. Recursion :: Key Concept 4: Return From the Base Case

What did we do when we determined that P(1) = 1, i.e., when we reached the **base case**? Did we stop and say the answer is 1? No, our original problem was not to find 1! but rather to compute 5!. Upon reaching the base case our original problem was not solved.

Rather, once we determined the solution to P(1) we **returned** to our problem P(2) and computed 2!. But after computing 2! did we stop? Once again, no, because our original problem had not yet been solved.

Rather, once we determined the solution to P(2) we **returned** to our problem P(3) and computed 3!. But after computing 3! did we stop? Once again, no, because our original problem had not yet been solved.

Rather, we repeated this procedure of **returning to the previous, larger, problem** and determining the solution to the larger problem, until we had finally reached and solved our original problem.

This process of returning, returning, ... returning, is one of the hallmarks of recursion and is also one of the most difficult for students to understand.

2. Recursion :: Key Concept 5: Subproblem Solutions Lead to the Solution

What did we do when we returned to problem P(2) after solving P(1)? Remember, P(2) was:

$$\mathrm{P}(2) = 2! = 2 \times 1! = 2 \times \underline{\mathrm{P}(1)} = 2 \times \underline{1} = 2$$

What we did was: **we used the solution from subproblem P(1) to obtain the solution to problem P(2)**. And after finding that P(2) = 2 did we stop? No, we returned to problem P(3):

$$\mathrm{P}(3) = 3! = 3 \times 2! = 3 \times \underline{\mathrm{P}(2)} = 3 \times \underline{2} = 6$$

and we used the solution from subproblem P(2) to obtain the solution to problem P(3).

And what did we do after solving P(4)? We returned to P(5):

$$\mathrm{P}(5) = 5! = 5 \times 4! = 5 \times \underline{\mathrm{P}(4)} = 5 \times \underline{24} = 120$$

and we used the solution from subproblem P(4) to obtain the solution to problem P(5), **which solved our original problem!** (I am excited again. Recursion is such a beautiful and elegant concept.)

Recursion reduces large, very-difficult-to-solve-problems into progressively smaller, more-easily-solvable-subproblems. The solutions to those smaller subproblems are used as we return back up the recursion chain to solve each of the larger problems. The end result is that when we return to the original problem, we now have all the pieces of the puzzle, so to speak, to solve the original problem.