COUNTING KNIGHTS AND KNAVES

OSCAR LEVIN AND GERRI M. ROBERTS

Oscar Levin (oscar.levin@unco.edu) received his Ph.D. from the University of Connecticut in 2009. Currently he is an assistant professor at the University of Northern Colorado. In addition to research in mathematical logic and computability theory, he enjoys recreational mathematics, especially for use in the classroom. He is the creator of the Math Puzzle Wiki (mathpuzzlewiki.com), which aims to collect the best mathematical puzzles for this purpose.

Gerri M. Roberts (midnightroxy13@gmail.com) started working on this project at the 2011 Frontiers in Science Institute at the University of Northern Colorado. She recently graduated from Poudre High School in May 2013, where she was an avid member of Math Club, Science Olympiad, and Science Bowl. She hopes to continue doing research through out her undergraduate work and is planning on majoring in mathematics or chemistry.

Walking through a fictional forest, you arrive at a bridge guarded by three trolls. Each troll is either a *knight*, who always tells the truth, or a *knave*, who always lies. They will not let you pass until you determine which belongs to which group. Each gives a clue:

Troll 1: Only one of us is a knave.
Troll 2: Only one of us is a knight.
Troll 3: We are all knaves.

Which trolls are which?

These puzzles appear frequently in books of mathematical recreations. This version is relatively simple (hint: Troll 3 can't be telling the truth). To see just how complex they can get, look at one of Raymond Smullyan's excellent books ([4] or [5], for example). Smullyan connects these puzzles with deep theorems in the foundations of mathematics. The self reference needed for knights and knaves puzzles is present, for example, in Gödel's incompleteness theorems. Additionally these puzzles are excellent teaching tools in any introductory logic class. This paper is not about logic however, it is about counting.

2000 *Mathematics Subject Classification.* 05A19, 97A20.

After solving some of these puzzles (further hint: what if Troll 1 were a knight) you might wish to create your own. How many different ones are there? That sounds like a fun counting problem and is precisely what we do here. We first count the number of knights and knaves puzzles and then count how many solutions they have. The two counting problems are surprisingly connected. We end with a variety of open questions.

Have you solved our challenge puzzle? That's right, Troll 2 is the sole knight. Troll 3 is a knave - if he were a knight, then his statement would be true making him (and everyone) a knave, a contradiction. Troll 1 is also a knave. If he spoke the truth, there would be a single knave, which we now know would be Troll 3, and two knights (Troll 1 and Troll 2), which makes Troll 2 a knave (his statement is false), another contradiction. This leaves Troll 2. If he were a knave then there would be three knaves making Troll 3's statement true. That's no good, so Troll 2 is a knight. Isn't logic fun?

Counting Puzzles

There are many variations of these puzzles. We limit ourselves to those in which each troll makes a single statement about the number of knights or knaves in the group. Additionally, since, for example, the assertion that 2 out of 5 trolls are knaves is equivalent to the assertion that 3 out of the 5 trolls are knights, we assume all statements give the number of knights instead of the number of knaves. Thus an *n-troll puzzle* consists of n trolls each of whom makes a statement of the form "x of us are knights," where x ranges from 0 to n.

Now we can count. Each of the n trolls can make one of $n+1$ different statements, so there are $(n+1)^n$ different n-troll puzzles.

Not very satisfying though, is it? Notice we count the puzzle in which Troll 1 says "0 of us are knights" and Troll 2 says "1 of us is a knight" separately from the one in which Troll 1 says "1 of us is a knight" and Troll 2 says "0 of us are knights," but they are essentially the same. The *name* of the troll making a particular statement shouldn't matter, only how *many* trolls make each statement. So, for the time being, we ignore the order in which the statements are made.

Perhaps the solution is still easy? There are $n+1$ different statements a troll can make, and we need statements for n trolls. Order doesn't matter, so aren't there just $\binom{n+1}{n}$ puzzles?

Not so fast. That is correct only if no two trolls make the same statement. There's no such rule. Trolls may choose statements multiple times. Instead of $n+1$ *choose* n, we need $n+1$ *multichoose* n, written

$\left(\!\!\binom{n+1}{n}\!\!\right)$ (following [2]). In general, $\left(\!\!\binom{n}{k}\!\!\right)$ is the number ways to select k objects from a group of n objects, order ignored, repetition allowed. Given this definition, the number of n-troll puzzles is $\left(\!\!\binom{n+1}{n}\!\!\right)$. What is this number? Let's count.

Counting puzzles by how many trolls make each statement

Since we don't distinguish the order statements are made, a puzzle is uniquely determined by the numbers of trolls making each statement. For example, a 5-troll puzzle may have two trolls who say "0 of us are knights," one troll saying "2 of us are knights," and two trolls making the statement "3 of us are knights." In an n-troll puzzle, there are $n+1$ different statements that can be made. We need only decide how many trolls make each statement.

With this in mind, we represent each puzzle as a string of stars and bars. For example, the 5-troll puzzle described above is represented by

$$* * || * | * * ||.$$

A star represents a troll; a bar represents a switch to the next type of statement in increasing order of the number of purported knights. For another example, the diagram

$$| * | * * * |||*,$$

again represents a 5-troll puzzle (there are 5 stars). Which one? It is the puzzle in which no trolls claim there are 0 knights (there is no star to the left of the first bar), one troll claims there is 1 knight, three trolls claim there are 2 knights (there are three stars between the second and third bar), no trolls claim there are 3 or 4 knights (no stars between those bars) and one troll claims there are 5 knights (there is one star to the right of the fifth bar). Every puzzle is represented uniquely by such a diagram, and every diagram describes exactly one puzzle, so counting puzzles amounts to counting diagrams.

The diagrams for n-troll puzzles have n stars, one for each troll, and n bars, one less than the number of possible statements (since we can put stars to the left of the first bar and to the right of the last bar). Each puzzle is an arrangement of these $2n$ symbols, so the number of n-troll puzzles is

$$\left(\!\!\binom{n+1}{n}\!\!\right) = \binom{2n}{n}.$$

This counts the puzzles, but for later use let's find a general formula for $\left(\!\!\binom{n}{k}\!\!\right)$. Suppose k trolls each make one of n different statements. Now the diagrams contain k stars and $n-1$ bars, and out of the $k+n-1$ symbols, k are stars, so

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{k+n-1}{k}.$$

Counting puzzles by the number of distinct statements

Here's another inventory of puzzles. This time we group them by how many different statements are made.

How many puzzles are there in which all trolls make the same statement? There are $n+1$ possible statements, so the answer is $n+1$. We write this as $\binom{n+1}{1}\binom{n-1}{0}$.

Next, how many puzzles are there in which exactly two statements are made? First we choose the statements. There are $\binom{n+1}{2}$ choices. Then we decide which trolls make which statement. We assume that all trolls making the first statement speak first, followed by all the trolls making the second statement. Of the n trolls, one of the first $n-1$ is the last troll to make the first statement. It cannot be troll n because at least one troll makes the second statement. There are $\binom{n-1}{1}$ ways to select one of the first $n-1$ trolls to play this role, therefore the total number of puzzles containing two distinct statements is $\binom{n+1}{2}\binom{n-1}{1}$.

Similarly the number of puzzles using three distinct statements is $\binom{n+1}{3}\binom{n-1}{2}$. We first select 3 of the $n+1$ statements, and then choose 2 of the first $n-1$ trolls to be the last troll making the first two types of statement.

The pattern is clear: the number of puzzles with k distinct statements is $\binom{n+1}{k}\binom{n-1}{k-1}$. So the total number of n-troll puzzles is

$$\left(\!\!\binom{n+1}{n}\!\!\right) = \sum_{k=1}^{n} \binom{n+1}{k}\binom{n-1}{k-1}.$$

As happens when one counts something in two ways, we have an identity:

$$\sum_{k=1}^{n} \binom{n+1}{k}\binom{n-1}{k-1} = \binom{2n}{n}.$$

What does it say? In Pascal's triangle, take a row and multiply each entry by the one directly below it, skipping a row. The sum of these products gives the middle entry in the row twice as far down as the row you skipped. For example, when $n=3$

$$(1\cdot 5)+(3\cdot 10)+(3\cdot 10)+(1\cdot 5) = 70,$$

as in Figure 1. If you like this identity, [3] has more.

$$\begin{array}{c}
1\\
1 \quad 1\\
1 \quad 2 \quad 1\\
1 \quad 3 \quad 3 \quad 1\\
1 \;\Big|\; 4 \;\Big|\; 6 \;\Big|\; 4 \;\Big|\; 1\\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1\\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1\\
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1\\
1 \quad 8 \quad 28 \quad 56 \quad \boxed{70} \quad 56 \quad 28 \quad 8 \quad 1
\end{array}$$

FIGURE 1

COUNTING SOLUTIONS

We have counted puzzles, but not all of them have solutions! Some have none, some have a unique solution, and some have multiple solutions. We investigate this phenomenon and come up with a way to count the number of solutions arising from all n-troll puzzles.

What is a solution? A solution labels each troll as either a knight or a knave, so that the statements made by knights are true and those made by knaves are false. For example, consider the 6-troll puzzle in which one troll says "1 of us is a knight," two trolls say "2 of us are knights," and three trolls say "3 of us are knights." Note that all the trolls in each group must be of the same type. Perhaps the first troll is a knight and the other five knaves. This is consistent with the statements since then only the first troll's statement is true. Alternatively, the two trolls who say "2 of us are knights" might be knights and the other four knaves. We can also label the three trolls claiming that "3 of us are knights" as knights, and the other three as knaves. A final solution labels all six as knaves. These are the only 4 solutions, since we have exhausted the possible numbers of knights.

It turns out that it is easy to solve our puzzles, partly because each troll makes a claim about the number of *knights*. If exactly k trolls say "k of us are knights" then one solution labels those k as knights and the others as knaves. The solutions to the example above were obtained by this principle.

For the purpose of counting, each solution is a labeling paired with the puzzle it solves. For example, the labeling in which the first three trolls are knaves and the last three are knights, which solves our example puzzle, also solves a different puzzle: three trolls say "1 of us is a knight" and three trolls say "3 of us are knights." We consider these different solutions though the labeling is identical. The key to our counting argument is that we count all the *puzzles* with a particular labeling, and then sum over all distinct labellings.

Counting solutions by the number of knights

First, consider the solution labeling all trolls as knights. There is only one puzzle with this solution; the puzzle in which all trolls say "n of us are knights." This puzzle has another solution in which all trolls are knaves, but we count that later.

Second, consider solutions which label all but one troll a knight. Puzzles with this solution must have $n-1$ trolls claiming there are $n-1$ knights. This leaves a lone troll to make one of the other n statements. There are n solutions of this type.

Next, suppose a solution bestows knight status upon $n-2$ trolls. These $n-2$ trolls say "$n-2$ of us are knights" while the other two trolls make any of n different statements. We have already counted the ways that can happen: $\left(\!\binom{n}{2}\!\right)$.

In general, if a solution labels $n-k$ trolls as knights, then $n-k$ trolls must claim that there are $n-k$ knights while the other k trolls make other claims. The k other trolls can choose from n statements (repeats allowed), so there are $\left(\!\binom{n}{k}\!\right)$ solutions of this kind. This agrees with our first two cases since $\left(\!\binom{n}{0}\!\right) = \binom{n-1}{0} = 1$ and $\left(\!\binom{n}{1}\!\right) = \binom{n}{1} = n$.

Putting this all together gives,

$$\left(\!\!\binom{n}{0}\!\!\right) + \left(\!\!\binom{n}{1}\!\!\right) + \left(\!\!\binom{n}{2}\!\!\right) + \left(\!\!\binom{n}{3}\!\!\right) + \cdots + \left(\!\!\binom{n}{n}\!\!\right),$$

which equals

$$\binom{n-1}{0} + \binom{n}{1} + \binom{n+1}{2} + \binom{n+2}{3} + \cdots + \binom{2n-1}{n},$$

because $\left(\!\binom{n}{k}\!\right) = \binom{k+n-1}{k}$. This looks familiar. For example, if $n = 3$ we get

$$\binom{2}{0} + \binom{3}{1} + \binom{4}{2} + \binom{5}{3} = \binom{6}{3},$$

as illustrated in Figure 2. This is an instance of the famed *hockey stick theorem* (a combinatorial proof can be found in [2]), from which we deduce that the number of solutions to all n-troll puzzles is

$$\binom{n-1}{0} + \binom{n}{1} + \binom{n+1}{2} + \binom{n+2}{3} + \cdots + \binom{2n-1}{n} = \binom{2n}{n}.$$

$$
\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}
$$

FIGURE 2

SURPRISE!

The number of n-troll puzzles and the number of solutions to all n-troll puzzles are the same! This might not be anything to brag about except that some puzzles have no solutions, and some have multiple solutions. The extra solutions of the multiple-solution puzzles exactly balance the number of puzzles without solutions. Surprising, but it gets stranger.

Let's go back to our first enumeration of puzzles, when for a moment order did matter. We found the number of n-troll puzzles to be $(n+1)^n$. How many solutions are there to *these* puzzles?

We count them just as we did when ignoring order. First, there is only one solution with all knights. Second, if the solution has $n-1$ knights, then there is exactly 1 troll making a statement other than "$n-1$ of us are knights." There are n choices for this other statement, but order counts now, so it matters which of the n trolls makes it. Thus there are $n\binom{n}{1}$ solutions of this type.

Considering solutions in which $n-2$ trolls are knights, there are now 2 trolls making other statements. They can do so in n^2 ways (n choices for the first troll, n choices for the second). We must also select which two trolls are knaves. So there are $n^2\binom{n}{2}$ solutions of this type.

In general, there are $n^k\binom{n}{k}$ solutions in which $n-k$ trolls are knights. Summing gives the number of solutions when order matters:

$$\binom{n}{0} + n\binom{n}{1} + n^2\binom{n}{2} + \cdots + n^n\binom{n}{n}.$$

This may not look like $(n+1)^n$, but it is if you expand using the binomial theorem! Again, the number of n-troll puzzles equals the number of solutions to those puzzles.

Stranger still, if you consider the number of n-troll puzzles in which exactly k distinct statements are made (as we did in our second attempt to count n-troll puzzles), again the number of such puzzles is the number of solutions. We leave this as an exercise. It is not very complicated but you will need the hockey stick theorem again.

What is going on?

Since the number of puzzles equals the number of solutions, there must be a bijection between the two sets - many bijections. Perhaps one gives insight into *why* the two sets have the same size.

Here's one possibility. Given an n-troll puzzle P, let x be the number of trolls who say "0 of us are knights." We map P to a solution S that labels those x trolls as knights. We must say which puzzle this labels. Here P does not work. We need x trolls to say "x of us are knights." So we swap the statements of the trolls who originally said "0 of us are knights" with those who said "x of us are knights," leaving all other statements the same.

It is straight forward to check that this is a bijection. Our map also makes sense whether order matters or not. What's more, it takes puzzles with k distinct statements to solutions of puzzles with k distinct statements, so we see that the number of puzzles equals the number of solutions even when so restricted.

Open Questions

Does our bijection really explain what's going on? Perhaps. It is true that any troll who says "0 of us are knights" *must* be lying. Whether trolls making other statements must be lying is contingent on how many trolls agree with them. Our map turns these necessarily false statements to true statements. Is this relevant? Maybe but exactly how is unclear. So our first open questions are: *why* is the number of puzzles equal to the number of solutions to those puzzles? If the given bijection is informative, how so? If not, what is a better bijection?

While we have totaled puzzles and solutions, we do not have a nice formula for the number of puzzles with a unique solution, or, for that matter, the number of puzzles with no solutions or with multiple solutions. Each of these appears to be difficult, if not impossible, depending on your definition of nice. Counting unique solutions appears to be a complicated inclusion/exclusion problem, which often have no nice closed formula. Are we missing something? We ask: how many n-troll puzzles have a unique solution? How many n-troll puzzles have no solutions? How many n-troll puzzles have multiple solutions?

We have really just begun to classify and count knights and knaves puzzles. Trolls might make more than one statement, or make statements of the form "between x and y of us are knights." If trolls can make more than one statement, an amusing variation is to allow a third class of trolls - *alternators* - who switch back and forth between telling the truth and lying (as in [1]).

What other knights and knaves puzzles would be fun to count?

Summary

To better understand some of the classic knights and knaves puzzles, we count them. Doing so reveals a surprising connection between puzzles and solutions, and highlights some beautiful combinatorial identities.

References

[1] B. Averbach and O. Chein, *Problem Solving through Recreational Mathematics*, Dover Publications, Mineola NY, 2000.

[2] A.T. Benjamin and J.J. Quinn, *Proofs that Really Count*, Mathematical Association of America, Washington DC, 2003.

[3] P. Hilton and J. Pedersen, Looking into Pascal's triangle: Combinatorics, arithmetic, and geometry, *Math. Mag.*, **60**, (1987) 305-316

[4] R. Smullyan, *The Lady or the Tiger?: and Other Logic Puzzles*, Dover Publications, Mineola NY, 2009.

[5] ———, *What Is the Name of This Book?: The Riddle of Dracula and Other Puzzles*, Dover Publications, Mineola NY, 2011.