# 3 Knights and Knaves

## A. THE ISLAND OF KNIGHTS AND KNAVES

There is a wide variety of puzzles about an island in which certain inhabitants called "knights" always tell the truth, and others called "knaves" always lie. It is assumed that every inhabitant of the island is either a knight or a knave. I shall start with a well-known puzzle of this type and then follow it with a variety of puzzles of my own.

### 26.

According to this old problem, three of the inhabitants—A, B, and C—were standing together in a garden. A stranger passed by and asked A, "Are you a knight or a knave?" A answered, but rather indistinctly, so the stranger could not make out what he said. The stranger than asked B, "What did A say?" B replied, "A said that he is a knave." At this point the third man, C, said, "Don't believe B; he is lying!"

The question is, what are B and C?

### 27.

When I came upon the above problem, it immediately struck me that C did not really function in any essential way; he was sort of an appendage. That is to say, the moment B spoke, one could tell without C's testimony that B was lying (see solution). The following variant of the problem eliminates that feature.

Suppose the stranger, instead of asking A what he is, asked A, "How many knights are among you?" Again A answers indistinctly. So the stranger asks B, "What did A say? B replies, "A said that there is one knight among us." Then C says, "Don't believe B; he is lying!"

Now what are B and C?

## 28.

In this problem, there are only two people, A and B, each of whom is either a knight or a knave. A makes the following statement: "At least one of us is a knave."

What are A and B?

## 29.

Suppose A says, "Either I am a knave or B is a knight." What are A and B?

## 30.

Suppose A says, "Either I am a knave or else two plus two equals five." What would you conclude?

## 31.

Again we have three people, A, B, C, each of whom is either a knight or a knave. A and B make the following statements:

*A*: All of us are knaves.
*B*: Exactly one of us is a knight.

What are A, B, C?

## 32.

Suppose instead, A and B say the following:

*A*: All of us are knaves.
*B*: Exactly one of us is a knave.

Can it be determined what B is? Can it be determined what C is?

## 33.

Suppose A says, "I am a knave, but B isn't."
What are A and B?

## 34.

We again have three inhabitants, A, B, and C, each of whom is a knight or a knave. Two people are said to be of the *same type* if they are both knights or both knaves. A and B make the following statements:

*A*: B is a knave.
*B*: A and C are of the same type.

What is C?

## 35.

Again three people A, B, and C. A says "B and C are of the same type." Someone then asks C, "Are A and B of the same type?"
What does C answer?

## 36. An Adventure of Mine.

This is an unusual puzzle; moreover it is taken from real life. Once when I visited the island of knights and knaves, I came across two of the inhabitants resting under a tree. I asked one of them, "Is either of you a knight?" He responded, and I knew the answer to my question.

What is the person to whom I addressed the question—is he a knight or a knave; And what is the other one? I can assure you, I have given you enough information to solve this problem.

## 37.

Suppose *you* visit the island of knights and knaves. You come across two of the inhabitants lazily lying in the sun. You ask one of them whether the other one is a knight, and you get a (yes-or-no) answer. Then you ask the second one whether the first one is a knight. You get a (yes-or-no) answer.

Are the two answers necessarily the same?

## 38. Edward or Edwin?

This time you come across just one inhabitant lazily lying in the sun. You remember that his first name is either Edwin or Edward, but you cannot remember which. So you ask him his first name and he answers "Edward."

What is his first name?

## B. KNIGHTS, KNAVES, AND NORMALS

An equally fascinating type of problem deals with three types of people: knights, who always tell the truth; knaves, who always lie; and normal people, who sometimes lie and sometimes tell the truth. Here are some puzzles of mine about knights, knaves, and normals.

## 39.

We are given three people, A,B,C, one of whom is a knight, one a knave, and one normal (but not necessarily in that order). They make the following statements:

*A*: I am normal.
*B*: That is true.
*C*: I am not normal.

What are A, B, and C?

## 40.

Here is an unusual one: Two people, A and B, each of whom is either a knight, or knave, or a normal, make the following statements:

*A*: B is a knight.
*B*: A is not a knight.

Prove that at least one of them is telling the truth, but is not a knight.

## 41

This time A and B say the following:

*A*: B is a knight.
*B*: A is a knave.

Prove that either one of them is telling the truth but is not a knight, or one of them is lying but is not a knave.

## 42. A Matter of Rank.

On this island of knights, knaves, and normals, knaves are said to be of the *lowest* rank, normals of *middle* rank, and knights of *highest* rank.

I am particularly partial to the following problem: Given two people A,B, each of whom is a knight, a knave, or a normal, they make the following statements:

*A*: I am of lower rank than B.
*B*: That's not true!

Can the ranks of either A or B be determined? Can it be determined of either of these statements whether it is true or false?

## 43.

Given three people A,B,C, one of whom is a knight, one a knave, and one normal. A,B, make the following statements:

*A*: B is of higher rank than C.
*B*: C is of higher rank than A.

Then C is asked: "Who has higher rank, A or B?" What does C answer?

## C. THE ISLAND OF BAHAVA

The island of Bahava is a female liberationist island; hence the women are also called *knights, knaves,* or *normals.* An ancient empress of Bahava once, in a whimsical moment, passed a curious decree that a knight could marry only a knave and a knave could marry only a knight. (Hence a normal can marry only a normal.) Thus, given any married couple, either they are both normal, or one of them is a knight and the other a knave.

The next three stories all take place on the island of Bahava.

## 44.

We first consider a married couple, Mr. and Mrs. A. They make the following statements:

*Mr. A* / My wife is not normal.
*Mrs. A* / My husband is not normal.

What are Mr. and Mrs. A?

## 45.

Suppose, instead, they had said:

> *Mr. A* / My wife is normal.
> *Mrs. A* / My husband is normal.

Would the answer have been different?

## 46.

This problem concerns two married couples on the island of Bahava, Mr. and Mrs. A, and Mr. and Mrs. B. They are being interviewed, and three of the four people give the following testimony:

> *Mr. A* / Mr. B is a knight.
> *Mrs. A* / My husband is right; Mr. B is a knight.
> *Mrs. B* / That's right. My husband is indeed a knight.

What are each of the four people, and which of the three statements are true?

# SOLUTIONS

## 26.

It is impossible for either a knight or a knave to say, "I'm a knave," because a knight wouldn't make the false statement that he is a knave, and a knave wouldn't make the true statement that he is a knave. Therefore A never did say that he was a knave. So B lied when he said that A said that he was a knave. Hence B is a knave. Since C said that B was lying and B was indeed lying, then C spoke the truth, hence is a knight. Thus B is a knave and C is a knight. (It is impossible to know what A is.)

## 27.

The answer is the same as that of the preceding problem, though the reasoning is a bit different.

The first thing to observe is that B and C must be of opposite types, since B contradicts C. So of these two, one is a knight and the other a knave. Now, if A were a knight, then there would be two knights present, hence A would not have lied and said there was only one. On the other hand, if A were a knave, then it would be true that there was exactly one knight present; but then A, being a knave, couldn't have made that true statement. Therefore A could not have said that there was one knight among them. So B falsely reported A's statement, and thus B is a knave and C is a knight.

## 28.

Suppose A were a knave. Then the statement "At least one of us is a knave" would be false (since knaves make false statements); hence they would both be knights. Thus, if A were a knave he would also have to be a knight, which is impossible. Therefore A is not a knave; he is a knight. Therefore his statement must be true, so at least one of them really is a knave. Since A is a knight, then B must be the knave. So A is a knight and B is a knave.

## 29.

This problem is a good introduction to the logic of disjunction. Given any two statements $p, q$, the statement "either $p$ or $q$" means that at least one (and possibly both) of the statements $p, q$ are true. If the statement "either $p$ or $q$" should be false, then *both* the statements $p, q$ are false. For example, if I should say, "Either it is raining or it is snowing," then if my statement is incorrect, it is both false that it is raining and false that it is snowing.

This is the way "either/or" is used in logic, and is the way it will be used throughout this book. In daily life, it is sometimes used this way (allowing the possibility that both alternatives hold) and sometimes in the so-called "exclusive" sense—that one and only one of the conditions holds. As an example of the exclusive use, if I say; "I will marry Betty or I will marry Jane," it is understood that the two possibilities are mutually exclusive—that is, that I will not marry both girls. On the other hand, if a college catalogue states that an entering student is required to have had *either* a year of mathematics *or* a year of a foreign language, the college is certainly not going to exclude you if you had both! This is the "*inclusive*" use of "either/or" and is the one we will constantly employ.

Another important property of the disjunction relation "either this or that" is this. Consider the statement "$p$ or $q$" (which is short for "either $p$ or $q$"). Suppose the statement happens to be true. Then if $p$ is false, $q$ must be true (because at least one of them is true, so if $p$ is false, $q$ must be the true one). For example, suppose it is true that it is either raining or snowing, but it is false that it is raining. Then it must be true that it is snowing.

We apply these two principles as follows. A made a statement of the disjunctive type: "Either I am a knave or B is a knight." Suppose A is a knave. Then the above statement must be false. This means that it is neither true that A is a knave nor that B is a knight. So if A were a knave, then it would follow that he is not a knave—which would be a contradiction. Therefore A must be a knight.

We have thus established that A is a knight. Therefore his statement is true that at least one of the possibilities holds: (1) A is a knave; (2) B is a knight. Since possibility (1) is false (since A is a knight) then possibility (2) must be the correct one, i.e., B is a knight. Hence A,B, are both knights.

## 30.

The only valid conclusion is that the author of this problem is not a knight. The fact is that neither a knight nor a knave could possibly make such a statement. If A were a knight, then the statement that either A is a knave or that two plus two equals five would be false, since it is neither the case that A is a knave nor that two plus two equals five. Thus A, a knight, would have made a false statement, which is impossible. On the other hand, if A were a knave, then the statement that either A is a knave or that two plus two equals five would be true, since the first clause that A is a knave is true. Thus A, a knave, would have made a true statement, which is equally impossible.

Therefore the conditions of the problem are contradictory (just like the problem of the irresistible cannonball and the immovable post). Therefore, I, the author of the problem, was either mistaken or lying. I can assure you I wasn't mistaken. Hence it follows that I am not a knight.

For the sake of the records, I would like to testify that I have told the truth at least once in my life, hence I am not a knave either.

## 31.

To begin with, A must be a knave, for if he were a knight, then it would be true that all three are knaves and hence that A too is a knave. If A were a knight he would have to be a knave, which is impossible. So A is a knave. Hence his statement was false, so in fact there is at least one knight among them.

Now, suppose B were a knave. Then A and B would both be knaves, so C would be a knight (since there is at least one knight among them). This would mean that there was exactly one knight among them, hence B's statement would be true. We would thus have the impossibility of a knave making a true statement. Therefore B must be a knight.

We now know that A is a knave and that B is a knight. Since B is a knight, his statement is true, so there is exactly one knight among them. This knight must be B, hence C must be a knave. Thus the answer is that A is a knave, B is a knight, and C is a knave.

## 32.

It cannot be determined what B is, but it can be proved that C is a knight.

To begin with, A must be a knave for the same reasons as in the preceding problem; hence also there is at least one knight among them. Now, either B is a knight or a knave. Suppose he is a knight. Then it is true that exactly one of them is a knave. This only knave must be A, so C would be a knight. So if B is a knight, so is C. On the other hand, if B is a knave, then C must be a knight, since all three can't be knaves (as we have seen). So in either case, C must be a knight.

## 33.

To begin with, A can't be a knight or his statement would be true, in which case he would have to be a knave. Therefore A is a knave. Hence also his statement is false. If B were a knight, then A's statement would be true. Hence B is also a knave. So A,B are both knaves.

## 34.

Suppose A is a knight. Then his statement that B is a knave must be true, so B is then a knave. Hence B's statement that A and C are of the same type is false, so A and C are of different types. Hence C must be a knave (since A is a knight). Thus if A is a knight, then C is a knave.

On the other hand, suppose A is a knave. Then his statement that B is a knave is false, hence B is a knight.

Hence B's statement is true that A and C are of the same type. This means that C must be a knave (since A is).

We have shown that regardless of whether A is a knight or a knave, C must be a knave. Hence C is a knave.

## 35.

I'm afraid we can solve this problem only by analysis into cases.

*Case One:* A is a knight. Then B,C really are of the same type. If C is a knight, then B is also a knight, hence is of the same type as A, so C being truthful must answer "Yes." If C is a knave, then B is also a knave (since he is the same type as C), hence is of a different type than A. So C, being a knave, must lie and say "Yes."

*Case Two:* A is a knave. Then B,C are of different types. If C is a knight, then B is a knave, hence he is of the same type as A. So C, being a knight, must answer "Yes." If C is a knave, then B, being of a different type than C, is a knight, hence is of a different type than A. Then C, being a knave, must lie about A and C being of different types, so he will answer "Yes."

Thus in both cases, C answers "Yes."

## 36.

To solve this problem, you must use the information I gave you that after the speaker's response, I knew the true answer to my question.

Suppose the speaker—call him A—had answered "Yes." Could I have then known whether at least one of them was a knight? Certainly not. For it could be that A was a knight and truthfully answered "Yes" (which would be truthful, since at least one—namely A—was a knight), or it could be that both of them were knaves, in which case A would have falsely answered "Yes" (which would indeed be false since neither was a knight). So if A had answered "Yes" I would have had no way of knowing. But I told you that I *did* know after A's answer. Therefore A must have answered "No."

The reader can now easily see what A and the other—call him B—must be: If A were a knight, he couldn't have truthfully answered "No," so A is a knave. Since his answer "No" is false, then there is at least one knight present. Hence A is a knave and B is a knight.

## 37.

Yes, they are. If they are both knights, then they will both answer "Yes." If they are both knaves, then again they will both answer "Yes." If one is a knight and the other a knave, then the knight will answer "No," and the knave will also answer "No."

## 38.

I feel entitled, occasionally, to a little horseplay. The vital clue I gave you was that the man was lazily lying in the sun. From this it follows that he was lying in the sun. From this it follows that he was *lying,* hence he is a knave. So his name is Edwin.

## 39.

To begin with, A cannot be a knight, because a knight would never say that he is normal. So A is a knave or is normal. Suppose A were normal. Then B's statement would be true, hence B is a knight or a normal, but B can't be normal (since A is), so B is a knight. This leaves C a knave. But a knave cannot say that he is not normal (because a knave really isn't normal), so we have a contradiction. Therefore A cannot be normal. Hence A is a knave. Then B's statement is false, so B must be normal (he can't be a knave since A is). Thus A is the knave, B is the normal one, hence C is the knight.

## 40.

The interesting thing about this problem is that it is impossible to know whether it is A who is telling the truth but isn't a knight or whether it is B who is telling the truth but isn't a knight; all we can prove is that at least one of them has that property.

Either A is telling the truth or he isn't. We shall prove: (1) If he is, then A is telling the truth but isn't a knight; (2) If he isn't, then B is telling the truth but isn't a knight.

(1) Suppose A is telling the truth. Then B really is a knight. Hence B is telling the truth, so A isn't a knight. Thus if A is telling the truth then A is a person who is telling the truth but isn't a knight.

(2) Suppose A is not telling the truth. Then B isn't a knight. But B must be telling the truth, since A can't be a knight (because A is not telling the truth). So in this case B is telling the truth but isn't a knight.

## 41.

We shall show that if B is telling the truth then he isn't a knight, and if he isn't telling the truth then A is lying but isn't a knave.

(1) Suppose B is telling the truth. Then A is a knave, hence A is certainly not telling the truth, hence B is not a knight. So in this case B is telling the truth but isn't a knight.

(2) Suppose B is not telling the truth. Then A is not really a knave. But A is certainly lying about B, because B can't be a knight if he isn't telling the truth. So in this case, A is lying but isn't a knave.

## 42.

To begin with, A can't be a knight, because it can't be true that a knight is of lower rank than anyone else. Now, suppose A is a knave. Then his statement is false, hence he is not of lower rank than B. Then B must also be a knave (for if he weren't, then A *would* be of lower rank than B). So if A is a knave, so is B. But this is impossible because B is contradicting A, and two contradictory claims can't both be false. Therefore the assumption that A is a knave leads to a contradiction. Therefore A is not a knave. Hence A must be normal.

Now, what about B? Well, if he were a knight, then A (being normal) actually would be of lower rank than B, hence A's statement would be true, hence B's statement false, and we would have the impossibility of a knight making a false statement. Thus B is not a knight. Suppose B were a knave. Then A's statement would be false, hence B's would be true, and we would have a knave making a true statement. Therefore B can't be a knave either. Hence B is normal.

Thus A and B are both normal. So also, A's statement is false and B's statement is true. So the problem admits of a complete solution.

## 43.

*Step 1*: We first show that from A's statement if follows that C cannot be normal. Well, if A is a knight then B really is of higher rank than C, hence B must be normal and C must be a knave. So in this case, C is not normal. Suppose A is a knave. Then B is not really of higher rank than C, hence B is of lower rank, so B must be normal and C must be a knight. So in this case, C again is not normal. The third possible case is that A is normal, in which case C certainly isn't (since only one of A, B, C is normal). Thus C is not normal.

*Step 2:* By similar reasoning, it follows from B's statement that A is not normal. Thus neither A nor C is normal. Therefore B is normal.

*Step 3*: Since C is not normal, then he is a knight or a knave. Suppose he is a knight. Then A is a knave (since B is normal) hence B is of higher rank than A. So C, being a knight, would truthfully answer, "B is of higher rank." On the other hand, suppose C is a knave. Then A must be a knight, so B is not of higher rank than A. Then C, being a knave, would lie and say, "B is of higher rank than A." So regardless of whether C is a knight or a knave, he answers that B is of higher rank than A.

## 44.

Mr. A cannot be a knave, because then his wife would be a knight and hence not normal, so Mr. A's statement would have been true. Similarly Mrs. A cannot be a knave. Therefore neither is a knight either (or the spouse would then be a knave), so they are both normal (and both lying).

## 45.

For the second problem, the answer is the same. Why?

## 46.

It turns out that all four are normal, and all three statements are lies.

First of all, Mrs. B must be normal, for if she were a knight her husband would be a knave, hence she wouldn't have lied and said he was a knight. If she were a knave, her husband would be a knight, but then she wouldn't have told the truth about this. Therefore Mrs. B is normal. Hence also Mr. B is normal. This means that Mr. and Mrs. A were both lying. Therefore neither one is a knight, and they can't both be knaves, so they are both normal.