

MATH 6

LOGIC I: KNIGHTS AND KNAVES — FULL 2026–27

A Little Story

Long ago, travelers told tales of a strange island with only two kinds of people:

- **Knights**, who always tell the truth, and
- **Knaves**, who always lie.

When you meet someone on this island, you cannot tell which one they are just by looking. You must listen carefully to what they say and use logic to figure out who is telling the truth.

A simple conversation like:

A says: “B is a Knight.”

can turn into a fun deduction puzzle. Are they telling the truth? Are they lying? Could they even be talking about themselves?

These puzzles turn you into a detective: your job is to examine each possibility, spot contradictions, and uncover the truth. They look like games, but they teach real mathematical thinking: working by cases, eliminating impossible scenarios, and understanding how statements can be true or false.

Welcome to the Island of Knights and Knaves!

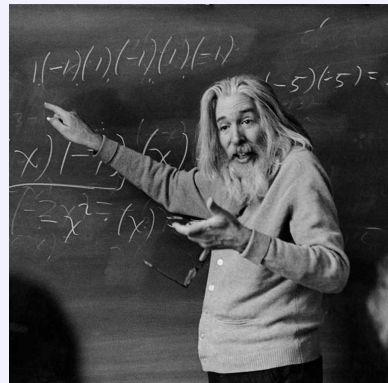
Historical Note

Stories about truth-tellers and liars go all the way back to ancient Greece. A saying traditionally attributed to the Cretan thinker *Epi-menides* is “All Cretans are liars.” It sounds self-defeating, although it is not quite the same as the sharper paradox “This statement is false.”

The American mathematician and logician **Raymond Smullyan** (1919–2017) made modern truth-teller-and-liar puzzles especially famous. He explored the fictional *Island of Knights and Knaves* in books such as *What Is the Name of This Book?* (1978) and *The Lady or the Tiger?* (1982).

He designed these puzzles to make logic playful and accessible. The same habits of testing cases and finding contradictions also appear in Boolean logic, programming, and artificial intelligence.

Today these puzzles remain popular because they are simple to start with, fun to solve, and excellent training for careful mathematical reasoning.



Raymond Smullyan (1919-2017)

The Core Rule

A **Knight's** statements are always true.

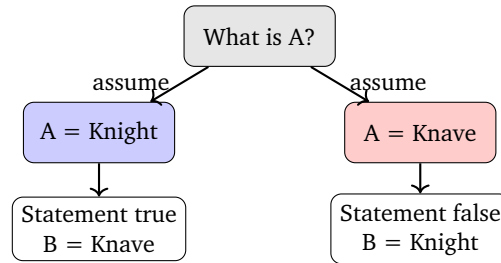
A **Knave's** statements are always false.

When solving problems:

1. Assume a person is a Knight. Their statement must be true.
2. Assume they are a Knave. Their statement must be false.

3. Keep every assignment that works. There may be no solution, exactly one solution, or several solutions.

Decision Tree: “A says: B is a Knave”



Conclusion: A and B are always opposite types

Quick Check

1. Can a Knight say “I am a Knave”?
2. Can a Knave say “I am a Knave”?
3. A says “B is a Knight.” If A is a Knave, what is B?
4. A says “We are the same type.” If A is a Knight, what is B?

Examples

Single Person

A says: “I am a Knight.”

What can we say about A?

We try both possibilities.

Case 1: A is a Knight. Then the statement “I am a Knight” is true. This is consistent with A being a Knight.

Case 2: A is a Knave. Then the statement “I am a Knight” is false. This is also consistent, because Knaves are allowed to make false statements.

So both assignments

A is a Knight and A is a Knave

are possible. From this statement alone, we *cannot* determine whether A is a Knight or a Knave.

Impossible Statement

A says: “I am a Knave.”

What can we say about A?

If A is a Knight, the statement is false — contradiction. If A is a Knave, the statement “I am a Knave” would be true, but Knaves cannot tell the truth.

Thus this statement is impossible. No inhabitant can say “I am a Knave.”

Two People

A says: "B is a Knave."

What can we say about A and B?

Case 1: If A is a Knight, then B is a Knave.

Case 2: If A is a Knave, then the statement is false, so B is a Knight.

Both assignments are possible, and A and B are necessarily different.

Two People with More Information

A says: "We are both Knaves."

What can we say about A and B?

A cannot be a Knight (a Knight cannot say they are a Knave). Thus A is a Knave. A's statement is false, so they are *not* both Knaves. Therefore B must be a Knight.

Same Kind or Different Kinds?

Alice says: "We are the same kind." Bob says: "We are of different kinds."

What are Alice and Bob?

Solution: These two statements are exact opposites: one asserts sameness, the other asserts difference. Therefore, **exactly one of them must be true**. This immediately means:

one is a Knight, and the other is a Knave.

But one Knight and one Knave are of *different* kinds. So Bob's statement is the true one, and Alice's is the false one: Bob is the Knight and Alice is the Knave.

Thus the only consistent assignment is:

Alice = Knave, Bob = Knight.

Two Cities Puzzle

There are two neighboring cities: one is a city of **Knights** (everyone always tells the truth), and the other is a city of **Knaves** (everyone always lies). The residents of these cities often travel between them. A traveler arrives at one of the cities, but does not know which one. He asks the first person he meets:

"Which city am I in?"

The person replies: "You are in the Knaves' city."

The traveler thinks for a moment and realizes that this answer does not help. Indeed:

- He might actually be in the **Knaves' city**, and the person he met is a **Knight** visiting there (so the answer would be true).
- Or he might be in the **Knights' city**, and the person he met is a **Knave** visiting there (so the answer would be false).

Both stories fit the answer, so it tells him nothing.

Can you find a better question before reading on? Here is one that works:

"Do you live here?"

Now the logic becomes clear:

- If he is in the **Knights' city**, a Knight living there says YES, and a Knave visiting from the other city lies and also says YES.
- If he is in the **Knaves' city**, a Knave living there must say NO (lying), and a Knight visiting from the other city truthfully says NO.

So whoever he meets, a YES means he is in the Knights' city, and a NO means he is in the Knaves' city.

Three People

A says: "C is a Knave."
B says: "A is a Knight."
C says: "I am a Knight."
What are A, B, and C?

Solution:

Case 1: B is a Knight.

If B is a Knight, then B's statement "A is a Knight" is true, so A is a Knight.

If A is a Knight, then A's statement "C is a Knave" is true, so C is a Knave.

Now check C: C is a Knave saying "I am a Knight." This is a false statement, which is allowed for a Knave.

So one consistent assignment is:

$$A = \text{Knight}, \quad B = \text{Knight}, \quad C = \text{Knave}.$$

Case 2: B is a Knave.

If B is a Knave, then B's statement "A is a Knight" is false, so A is a Knave.

If A is a Knave, then A's statement "C is a Knave" is false, so C is a Knight.

Now check C: C is a Knight saying "I am a Knight." This is a true statement, which is allowed for a Knight.

So a second consistent assignment is:

$$A = \text{Knave}, \quad B = \text{Knave}, \quad C = \text{Knight}.$$

Conclusion. This puzzle has *two* possible solutions. From these statements alone we cannot uniquely determine who is who.

Leonard and Martin on the Beach

Leonard and Martin are sitting on the beach. I ask Leonard:

"Is either of you a Knight?"

Leonard answers *yes* or *no*, and from his single answer I can immediately determine whether each of the two men is a Knight or a Knave. What must they be?

Solution: Let L = Leonard and M = Martin. The question "Is either of you a Knight?" is the statement:

$$(L \text{ is a Knight}) \quad \text{or} \quad (M \text{ is a Knight}).$$

Consider all possibilities and what Leonard would answer.

- **If L is a Knight, M is a Knight:** The statement is true, so Leonard answers **Yes**.

- **If L is a Knight, M is a Knave:** The statement is still true, so Leonard answers **Yes**.
- **If L is a Knave, M is a Knight:** The statement is true, but Leonard must lie, so he answers **No**.
- **If L is a Knave, M is a Knave:** The statement is false, so Leonard (lying) answers **Yes**.

If Leonard said **Yes**, there would still be three possibilities left, and I could *not* tell who is who. If Leonard said **No**, only one possibility is left.

But I *did* figure it out from his answer. So he must have said No, and therefore:

Leonard is a Knave and Martin is a Knight.

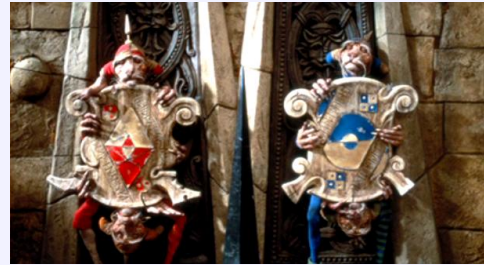
Fork in the Road

John and Bill are standing at a fork in the road. John stands by the left road, Bill stands by the right road. One is a Knight (always tells the truth) and the other is a Knave (always lies). One road leads to **Death**, the other to **Freedom**. You may ask *one yes-no question* to *one* of them. How can you find the road to Freedom?

Idea. We want a question where *both* a Knight and a Knave give an answer that tells us what to do. A classic trick is to ask what the *other* person would say.

Ask John the following question:

“If I asked Bill ‘Does your road lead to Freedom?’, would he say yes?”



Scene from the *Labyrinth* (1986)

Let’s check all four possibilities. “Bill would say” is Bill’s answer to “Does your road lead to Freedom?”

John is a . .	Bill’s road leads to . .	Bill would say	John answers
Knight	Freedom	No (Bill lies)	No (John is honest)
Knight	Death	Yes (Bill lies)	Yes (John is honest)
Knave	Freedom	Yes (Bill is honest)	No (John lies)
Knave	Death	No (Bill is honest)	Yes (John lies)

Look at the last column: John says **Yes** exactly when Bill’s road leads to Death, and **No** exactly when it leads to Freedom — no matter who is who.

Why does this work? The answer passes through *both* men, and exactly one of them is a liar. So it gets flipped exactly once: John’s answer is always the *opposite* of the truth about Bill’s road.

Conclusion. If John answers **Yes**, take **John’s** road. If he answers **No**, take **Bill’s** road.

What questions are we allowed to ask?

In these puzzles we ask questions *about the island* — about who is a Knight or a Knave, or about what someone would say. We do not lean on outside facts that the islanders happen to know. (You *could* ask “Is $2 + 2 = 4$?” and a Knight would say “yes” while a Knave says “no” — but that dodges the real challenge, which is reasoning about the truth-tellers and liars themselves.)

Quick Check

5. A says “B is a Knight.” B says “We are different types.” Is it possible for both to be Knights?
6. Why do the direct questions “Are you a Knight?” and “Are you a Knave?” fail to reveal a person’s type?

Normals

Now let’s add a third kind of islander. Besides **Knights**, who always tell the truth, and **Knaves**, who always lie, there are **Normals**, who sometimes lie and sometimes tell the truth.

A Normal can say anything at all — even “I am a Knave” (a lie, but Normals are allowed to lie). So once Normals are around, this “impossible” sentence is possible again, and it tells us the speaker is Normal!

Normals

We meet three people, A, B, and C. One is a Knight, one is a Knave, and one is a Normal, but we don’t know who is who. They make the following statements:

A: “I am normal.”

B: “That is true.”

C: “I am not normal.”

What are A, B, and C?

Step 1. A is not a Knight: a Knight would never say “I am normal.” So A is a Knave or a Normal.

Step 2. Suppose A is the Normal. Then A’s sentence is true, so B (“That is true”) told the truth. B is not the Normal (A is), so B is the Knight, and C is the Knave. But then C’s sentence “I am not normal” is *true* — and a Knave can’t say something true. Contradiction!

Step 3. So A is the Knave. A’s sentence is false, so B’s “That is true” is false too. B is not the Knave (A is), so B is the Normal, and C is the Knight. Check: C truthfully says “I am not normal.” ✓

A = Knave, B = Normal, C = Knight.

The Island of Bahava

Bahava is another imaginary island, home to Knights, Knaves, and Normals. It has one unusual marriage rule: a Knight may marry only a Knave, a Knave may marry only a Knight, and a Normal may marry only another Normal. Thus every married couple consists either of two Normals or of one Knight and one Knave.

Married Couple

We consider a married couple from the Island of Bahava, 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?

Could Mr. A be a Knave? Then his wife is a Knight (Knaves marry Knights), so she really is not normal, and his sentence is true. A Knave can’t say something true — so no.

Could Mr. A be a Knight? Then his wife is a Knave. But by the same argument with the roles swapped, Mrs. A can’t be a Knave. So no.

So Mr. A is a Normal, and therefore so is his wife: **both are Normals**, and both are lying.

Quick Check

7. Can a Normal say “I am a Knight”?
8. On the Island of Bahava, Mr. X says “My wife is a Knave,” and Mr. X is a Knave. What is Mrs. X?

Patterns Worth Remembering

A says. . .	If A is a Knight	If A is a Knave
“I am a Knight”	true — fine	false — fine
“I am a Knave”	false — impossible!	true — impossible!
“B is a Knight”	B is a Knight	B is a Knave
“B is a Knave”	B is a Knave	B is a Knight
“We are the same type”	same type, so B is a Knight	really different, so B is a Knight
“We are different types”	really different, so B is a Knave	really the same, so B is a Knave

Look at the last two rows: whatever A is, “We are the same type” tells us that B is a Knight, and “We are different types” tells us that B is a Knave!

Key Takeaways

- **Knights** always tell the truth; **Knaves** always lie
- To solve: assume each possibility and check for contradictions
- When everyone is either a Knight or a Knave, no one can say “I am a Knave” (but a Normal can!)
- “We are the same type” and “We are different types” are opposite statements
- When asking questions, consider what *both* a Knight and Knave would answer
- The “ask about the other person” trick works for fork-in-the-road problems

Common Mistakes

- **Forgetting to check both cases:** Always verify your answer works for all assumptions
- **Assuming unique solutions:** Some puzzles have multiple valid answers
- **Confusing “lying” with “wrong”:** A Knave’s statement is false, not just mistaken
- **Forgetting Normals:** When Normals are involved, they can say anything

Practice Problems

1. A says: "I am a Knight." What can we conclude about A?
2. A says: "I am not a Knave." What can we conclude?
3. A says: "I always lie." What can we conclude?
4. A says: "B is a Knight." B says: "We are different types." What are A and B?
5. A says: "At least one of us is a Knave, and B is a Knight." What are A and B?
6. A says: "B is a Knave." B says: "A is a Knave." What can we conclude?
7. A says: "At least one of us is a Knight," pointing at himself and B. What can we conclude?
8. A says: "B is a Knight." B says: "A is a Knight." What can we conclude about A and B?
9. A says: "We are of different types." What can we conclude about A and B?
10. A says: "B is a Knave." B says: "We are the same type." What are A and B?
11. A says: "B is a Knave." B says: "C is a Knave." C says: "A is a Knave." What can we conclude?
12. A says: "All three of us are Knaves," pointing at A, B, and C. What can we conclude?
13. A says: "I am a Knave or I am a Knight." What can we conclude about A?
14. A says: "At least one of us is a Knave," pointing at himself and B. What are A and B?
15. A says: "I am the same type as B." B says: "I am a Knave." What can we conclude?
16. A says: "B and I are both Knights." B says: "That is false." What are A and B?
17. A says: "C is a Knave." B says: "A is a Knight." C says: "B is a Knave." What are A, B, and C?
18. You want to know whether it will rain tomorrow. You may ask *one* islander (a Knight or a Knave — you don't know which) a single yes/no question. What should you ask?
19. A says: "We are the same type." B says: "A is a Knave." What are A and B?
20. A says: "All of us are the same type." B says: "Exactly one of us is a Knight." (There are three: A, B, C.) What are A, B, and C?

Practice Problems: Solutions

1. **We cannot determine A's type.** If A is a Knight, the statement is true (consistent). If A is a Knave, the statement is false (also consistent).
2. "I am not a Knave" means the same as "I am a Knight." Same as problem 1: **Cannot determine A's type.**
3. "I always lie" is equivalent to "I am a Knave." A Knight can't say this (it would be false). A Knave can't say this (it would be true). **Impossible statement—no inhabitant can say this!**
4. If A is a Knight, then B is a Knight. But then B's statement "We are different types" would be false, contradiction. So A is a Knave. Then B is not a Knight, so B is a Knave. Check: B says "We are different types" which is false. ✓
Answer: A = Knave, B = Knave
5. If A were a Knight, his statement would be true, so B would be a Knight — but then there would be no Knave at all, and "at least one of us is a Knave" would be false. Contradiction. So A is a Knave, and his statement is false. Its first part, "at least one of us is a Knave," is true (A himself is one), so the part that fails must be "B is a Knight."
Answer: A = Knave, B = Knave
6. If A is a Knight, then B is a Knave, so B's statement is false, meaning A is not a Knave. Consistent!
If A is a Knave, then B is not a Knave, so B is a Knight. B's statement "A is a Knave" is true. Consistent!
Both scenarios work: either (A=Knight, B=Knave) or (A=Knave, B=Knight). We only know they are different types.
7. If A is a Knight, "At least one of us is a Knight" is true (A himself is). Consistent.
If A is a Knave, the statement is false, so neither is a Knight. That means B is also a Knave. Consistent.
Two solutions: A=Knight (B=either) or A=Knave and B=Knave.
8. If A is a Knight, then B is a Knight (A's statement), and B (a Knight) truly says "A is a Knight." ✓ If A is a Knave, then B is a Knave (A's statement is false), and B (a Knave) falsely says "A is a Knight." ✓ **A and B are the same type: both Knights or both Knaves (two solutions).**
9. If A is a Knight, "different types" is true, so B is a Knave. If A is a Knave, "different types" is false, so they are the same type, making B a Knave too. Either way, **B is a Knave**; A cannot be determined.
10. If A is a Knight, B is a Knave; then B (Knave) says "same type," which is false (they differ) ✓. If A is a Knave, B is a Knight; then B (Knight) says "same type," but they differ — a Knight can't lie, contradiction. **A = Knight, B = Knave.**
11. Suppose A is a Knight: then B is a Knave, so "C is a Knave" is false, so C is a Knight, so "A is a Knave" is true — contradiction. Suppose A is a Knave: then B is a Knight, so C is a Knave, so "A is a Knave" is false — contradiction. **No consistent assignment exists** (a three-person cycle of "the next one is a Knave" has no solution).
12. A can't be a Knight (a Knight would never truthfully call all three, himself included, Knaves). So A is a Knave, and "all three are Knaves" is false. **A = Knave, and at least one of B, C is a Knight** (that is all we can determine).
13. "I am a Knave or I am a Knight" is always true (everyone is one or the other). A Knave cannot make a true statement, so **A must be a Knight.**
14. A can't be a Knave (then "at least one of us is a Knave" would be true, but Knaves lie). So A is a Knight; his true statement means at least one is a Knave, so B is the Knave. **A = Knight, B = Knave.**
15. B's statement "I am a Knave" is one that *no* islander can ever make. So the situation is **impossible** — there is no such B, whatever A is.

16. If A is a Knight, both are Knights, so B is a Knight; but then B's "that is false" (they are not both Knights) is false — a Knight lying, contradiction. So A is a Knave, "both Knights" is false, and B's "that is false" is true. **A = Knave, B = Knight.**
17. If A is a Knight: "C is a Knave" is true, so C is a Knave; B says "A is a Knight," true, so B is a Knight; check C (Knave) "B is a Knave" — false ✓. Gives (Knight, Knight, Knave).
If A is a Knave: C is a Knight; B says "A is a Knight," false, so B is a Knave; check C (Knight) "B is a Knave" — true ✓. Gives (Knave, Knave, Knight). **Two solutions: (Knight, Knight, Knave) or (Knave, Knave, Knight).**
18. Ask: "If I asked you whether it will rain tomorrow, would you say 'yes'?" A Knight answers this honestly; a Knave, lying twice (about the weather and about what he would say), also lands on the truth. So the islander says "yes" exactly when it will rain — whichever type he is.
19. If A is a Knight, "same type" is true, so B is a Knight; but then B's "A is a Knave" is false — a Knight lying, contradiction. So A is a Knave, "same type" is false, so B is a Knight. Check B (Knight) "A is a Knave" — true ✓. **A = Knave, B = Knight.**
20. If A is a Knight, "all the same type" is true, so all three are Knights — but then "exactly one Knight" is false, a contradiction for a Knight B. So A is a Knave and "all the same type" is false. If B were a Knave, then (A, B Knaves) the third, C, would have to be a Knight so the types aren't all the same — making "exactly one Knight" true, which a Knave can't say. So B is a Knight, "exactly one Knight" is true, and A, C are Knaves. **A = Knave, B = Knight, C = Knave.**

Homework

1. Alice and Bob are standing next to each other. Alice says, “At least one of us is a Knight.” Bob says, “Alice is a Knave.” What are Alice and Bob?
2. You meet two inhabitants: Claire and Desmond. Claire suddenly says, “We are both Knights.” After this, Desmond says “Either Claire is a Knight or I am a Knight, but we are not both Knights.” What are Claire and Desmond?
3. **H** There are three native islanders, named Elena, Fernando, and Gary, standing together. You ask Elena, “How many Knights are among you?” Elena answers, but you cannot quite hear what she says. You then ask Fernando, “What did Elena say?” Fernando replies, “Elena said there is one Knight among us.” At this point Gary says, “Don’t believe Fernando; he is lying.”
Is it possible to determine what Elena is? Is it possible to determine what Fernando is? Is it possible to determine what Gary is?
4. You meet two inhabitants: Alice and Bob. Alice tells you that Bob is a knave. Bob says, “Neither Alice nor I are knaves.” Who is a knight and who is a knave?
5. Three inhabitants—A, B, and C—are standing together in a garden. A traveler asks A, “Are you a Knight or a Knave?” A answers too quietly to be heard. The traveler then asks B, “What did A say?” B replies, “A said that he is a Knave.” C says, “Don’t believe B; he is lying!” What are B and C?
6. Suppose Bridget says, “Either I am a knave or Connor is a knight.” What are Bridget and Connor?
7. Harry says that all three people—Harry, Irma, and Janet—are Knaves. Irma then says that there is exactly one Knave among them. What can you conclude?
8. Suppose Deb says, “I am a knave, but Ed isn’t.” What are Deb and Ed?
9. We again have three inhabitants, Alice, Bob, and Charlie, 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. Alice and Bob make the following statements:
Alice: “Bob is a knave.”
Bob: “Alice and Charlie are of the same type.”
What is Charlie?
10. **M** Harry, Irma, and Jane are standing together. Harry says, “Irma and Jane are of the same type.” Someone then asks Jane, “Are Harry and Irma of the same type?” What does Jane answer?
11. Consider a married couple on the Island of Bahava, Mr. and Mrs. A. They make the following statements:
Mr. A: My wife is normal.
Mrs. A: My husband is normal.
What are Mr. and Mrs. A?
12. **M** **Who are you?** You meet a single islander and want to find out whether he is a Knight or a Knave. The direct questions “Are you a Knight?” and “Are you a Knave?” don’t work — why not? Find one yes/no question that does.
13. A says: “B is a Knight.” B says: “A is a Knave.” What are A and B?

Quick Check Answers

1. No — a Knight cannot make a false statement.
2. No — a Knave cannot make a true statement.
3. B is a Knave (the statement is false, so B is not a Knight).
4. B is a Knight (the statement is true, so they are the same type).
5. No. If A is a Knight, B is a Knight. Then B's statement "different types" is false—but Knights don't lie!
6. Both types answer "Yes" to "Are you a Knight?": the Knight answers truthfully, while the Knave lies. Both answer "No" to "Are you a Knave?": again, the Knight tells the truth and the Knave lies. So neither *direct* question separates them.
7. Yes — Normals can say anything (sometimes true, sometimes false).
8. Mrs. X is a Knight. On Bahava a Knave marries a Knight, and this fits: Mr. X (a Knave) falsely says "my wife is a Knave," and indeed she isn't one.