

The Hardest Logic Puzzle Ever¹

By *George Boolos*

SOME YEARS AGO, THE LOGICIAN AND PUZZLE-MASTER Raymond Smullyan devised a logical puzzle that has no challengers I know of for the title of Hardest Logical Puzzle Ever. I'll set out the puzzle here, give the solution, and then briefly discuss one of its more interesting aspects.

The puzzle: Three gods A, B, and C are called, in some order, True, False, and Random. True always speaks truly, False always speaks falsely, but whether Random speaks truly or falsely is a completely *random* matter. Your task is to determine the identities of A, B, and C by asking three yes-no questions; each question must be put to exactly one god. The gods understand English, but will answer all questions in their own language, in which the words for "yes" and "no" are "da" and "ja," in some order. *You do not know which word means which.*²

Before I present the somewhat lengthy solution, let me give answers to certain questions about the puzzle that occasionally arise:

- It could be that some god gets asked more than one question (and hence that some god is not asked any question at all).

GEORGE BOOLOS IS Professor of Philosophy at MIT and one of the founders of the field of "Provability Logic." He is the author of *The Logic of Provability* (Cambridge, 1993) and, with Richard Jeffrey, *Computability and Logic* (Cambridge, 1974).

- What the second question is, and to which god it is put, may depend on the answer to the first question. (And of course similarly for the third question.)
- Whether Random speaks truly or not should be thought of as depending on the flip of a coin hidden in his brain: if the coin comes down heads, he speaks truly; if tails, falsely.
- Random will answer da or ja when asked any yes-no question.

The Solution: Before solving The Hardest Logic Puzzle Ever, we will set out and solve three related, but much easier, puzzles. We shall then combine the ideas of their solutions to solve the Hardest Puzzle. The last two puzzles are of a type that may be quite familiar to the reader, but the first one is not well known (in fact the author made it up while thinking about the Hardest Puzzle).

Puzzle 1: Noting their locations, I place two aces and a jack face down on a table, in a row; you do not see which card is placed where. Your problem is to point to one of the three cards and then ask me a single yes-no question, from the answer to which you can, with certainty, identify one of the three cards as an ace. If you have pointed to one of the aces, I will answer your question truthfully. However, if you have pointed to the jack, I will answer your question yes or no, completely at random.

Puzzle 2: Suppose that, somehow, you have learned that you are speaking not to Random but to True or False — you don't know which — and that whichever god you're talking to has condescended to answer you in English. For some reason, you need to know whether Dushanbe is in Kirghizia or not. What one yes-no question can you ask the god from the answer to which you can determine whether or not Dushanbe is in Kirghizia?

Puzzle 3: You are now quite definitely talking to True, but he refuses to answer you in English and will only say da or ja. What one yes-no question can you ask True to determine whether or not Dushanbe is in Kirghizia?

HERE'S ONE SOLUTION TO PUZZLE 1: POINT TO THE middle card and ask, "Is the left card an ace?" If I answer yes, choose the left card; if I answer no, choose the right card. *Whether the middle card is an ace or not*, you are certain to find an ace by choosing the left card if you hear me say yes and choosing the right card if you hear no. The reason is that if the middle card is an ace, my answer is truthful, and so the left card is an ace if I say yes, and the right card is an ace if I say no. But if the middle card is the Jack, then *both* of the other cards are aces, and so again the left card is an ace if I say yes (so is the right card but that is now irrelevant), and the right card is an ace if I say no (as is the left card, again irrelevantly).

To solve puzzles 2 and 3, we shall use *iff*.

Logicians have introduced the useful abbreviation "iff," short for "if, and only if." The way "iff" works in logic is this: when you insert "iff" between two statements that are either both true or both false, you get a statement that is true; but if you insert it between one true and one false statement, you get a false statement. Thus, for example, "The moon is made of Gorgonzola iff Rome is in Russia" is *true*, because "The moon is made of Gorgonzola" and "Rome is in Russia" are both false. But, "The moon is made of Gorgonzola iff Rome is in Italy" and "The moon lacks air iff Rome is in Russia" are false. However, "The moon lacks air iff Rome is in Italy" is true. ("Iff" has nothing to do with causes, explanations, or laws of nature.)

To solve puzzle 2, ask the god not the simple question, "Is Dushanbe in Kirghizia?" but the more complex question, "Are you True iff Dushanbe is in Kirghizia?" Then (in the absence of any geographical information) there are four possibilities:

- 1) The god is True and D. is in K.: then you get the answer yes.
- 2) The god is True and D. is not in K.: this time you get no.
- 3) The god is False and D. is in K.: you get the answer yes, because only

one statement is true, so the correct answer is no, and the god, who is False, falsely says yes.

4) The god is False and D. is not in K.: in this final case you get the answer no, because both statements are false, the correct answer is yes, and the god False falsely says no.

So you get a yes answer to that complex question if D. is in K. and a no answer if it is not, *no matter to which of True and False you are speaking*. By noting the answer to the complex question, you can find out whether D. is in K. or not.

The point to notice is that if you ask either True or False, "Are you True iff X?" and receive your answer in English, then you get the answer yes if X is true and no if X is false, regardless of which of the two you are speaking to.

The solution to puzzle 3 is quite similar: Ask True not, "Is Dushanbe in Kirghizia?" but, "Does da mean yes iff D. is in K.?" There are again four possibilities:

- 1) Da means yes and D. is in K.: then True says da.
- 2) Da means yes and D. is not in K.: then True says ja (meaning no).
- 3) Da means no and D. is in K.: then True says da (meaning no).
- 4) Da means no and D. is not in K.: then both statements are false, the statement "Da means yes iff D. is in K." is true, the correct answer (in English) to our question is yes, and therefore True says ja.

Thus you get the answer da if D. is in K. and the answer ja if not, regardless of which of da and ja means yes and which means no.

The point this time is that if you ask True, "Does da mean yes iff Y?" then you get the answer da if Y is true and you get the answer ja if Y is false, regardless of which means which.

Combining the two points, we see that if you ask one of True and False (who we again suppose only answer da and ja), the very complex question, "Does da mean yes iff, you are True iff X?" then *you will get the answer da if X is true and get the answer ja if X is false*, regardless of whether you are addressing the god True or the god False, and regardless of the meanings of da and ja.

We can now solve The Hardest Logic Puzzle Ever.

Your first move is to find a god who you can be certain is not Random, and hence is either True or False.

To do so, turn to A and ask Question 1: *Does da mean yes iff, you are True iff B is Random?* If A is True or False and you get the answer da, then as we have seen, B is Random, and therefore C is either True or False; but if A is True or False and you get the answer ja, then B is not Random, therefore B is either True or False.

But what if A is Random?

If A is Random, then neither B nor C is Random!

So if A is Random and you get the answer da, C is not Random (neither is B, but that's irrelevant), and therefore C is either True or False; and if A is Random

and you get the answer ja, B is not random (neither is C, irrelevantly), and therefore B is either True or False.

Thus, *no matter whether A is True, False, or Random*, if you get the answer da to Question 1, C is either True or False, and if you get the answer ja, B is either True or False!

Now turn to whichever of B and C you have just discovered is either True or False — let us suppose that it is B (if it is C, just interchange the names B and C in what follows) — and ask Question 2: *Does da mean yes iff Rome is in Italy?* True will answer da, and False will answer ja. Thus, with two questions, you have either identified B as True or identified B as False.

For our third and last question, turn again to B, whom you have now either identified as True or identified as False, and ask Question 3: *Does da mean yes iff A is Random?*

Suppose B is True. Then if you get the answer da, then A is Random, and therefore A is Random, B is True, C is False, and you are done; but if you get the answer ja, then A is not Random, so A is False, B is true, C is Random, and you are again done.

Suppose B is False. Then if you get the answer da, then since B speaks falsely, A is not Random, and therefore A is True, B is False, C is Random, and you are done; but if we get ja, then A is Random, and thus B is False, and C is True, and you are again done. FINIS.

Well, I wasn't speaking falsely or at random when I said that the puzzle was hard, was I?

A brief remark about the significance of the Hardest Logic Puzzle Ever:

There is a law of logic called "the law of excluded middle," according to which either X is true or not-X is true, for any statement X at all. ("The law of non-contradiction" asserts that statements X and not-X aren't both true.) Mathematicians and philosophers have occasionally attacked the idea that excluded middle is a logically valid law. We can't hope to settle the debate here, but can observe that our solution to puzzle 1 made essential use of excluded middle, exactly when we said "Whether the middle card is an ace or not...." It is clear from The Hardest Logic Puzzle Ever, and even more plainly from puzzle 1, that our ability to reason about alternative possibilities, even in everyday life, would be almost completely paralyzed were we to be denied the use of the law of excluded middle.

By the way, Dushanbe is in Tajikistan, not Kirghizia. ϕ

Endnotes

¹A version of this article, translated by Massimo Piattelli-Palmarini, appeared in *La Repubblica* on 16 April 1992 under the title "L'indovinello più difficile del mondo."

²The extra twist of not knowing which are the gods' words for "yes" and "no" is due to the computer scientist John McCarthy.