Why the Hardest Logic Puzzle Ever Cannot Be Solved in Less than Three Questions

Gregory Wheeler · Pedro Barahona

Abstract Rabern and Rabern (Analysis 68:105–112 2) and Uzquiano (Analysis 70:39–44 4) have each presented increasingly harder versions of 'the hardest logic puzzle ever' (Boolos The Harvard Review of Philosophy 6:62–65 1), and each has provided a two-question solution to his predecessor's puzzle. But Uzquiano's puzzle is different from the original and different from Rabern and Rabern's in at least one important respect: it cannot be solved in less than three questions. In this paper we solve Uzquiano's puzzle in three questions and show why there is no solution in two. Finally, to cement a tradition, we introduce a puzzle of our own.

Keywords Boolos · Logic puzzle · Information theory

Rabern and Rabern [2] and Uzquiano [4] have each presented increasingly harder versions of 'the hardest logic puzzle ever' [1], and each has provided a two-question solution to his predecessor's puzzle. But Uzquiano's puzzle is different from the original and different from Rabern and Rabern's in at least one important respect: it cannot be solved in less than three questions. In this paper we solve Uzquiano's puzzle in three questions and show why there is no solution in two. Finally, to cement a tradition, we introduce a puzzle of our own.

Recall Uzquiano's puzzle and his guidelines for solving it.

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 or whether Random speaks at all 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

G. Wheeler (✉) · P. Barahona
CENTRIA - Centre for Artificial Intelligence, The New University of Lisbon, Lisbon, Portugal
e-mail: grw@fct.unl.pt

will answer in their own language, in which the words for 'yes' and 'no' are 'da' and 'ja', in some order. You don't know which word means which. ([4]: 44)

The guidelines:

1. It could be that some god gets asked more than one question.
2. What the second question is, and to which god it is put, may depend on the answer to the first question.
3. Whether Random answers 'da' or 'ja' or whether Random answers at all should be thought of as depending on the toss of a fair three-sided dice hidden in his brain: if the dice comes down 1, he doesn't answer at all; if the dice comes down 2, he answers 'da'; if 3, 'ja'.

What distinguishes these puzzles from each other are three different specifications for Random's behavior. Boolos [1] allows for Random to speak truly or lie, albeit randomly, whereas Rabern and Rabern [2] stipulate that Random answers 'da' and 'ja' randomly. Both Uzquiano's solution strategy as well as Rabern and Rabern's exploit a common trait in the first two puzzles, which is that there are yes/no questions that True cannot answer and yes/no questions that False cannot answer, but no question that Random will fail to answer. Uzquiano [4] eliminates this particular asymmetry from his version of the puzzle by granting Random the option of remaining silent. This modification, need it be said, is what makes Uzquiano's the hardest logic puzzle ever.

There are three parts to our solution to Uzquiano's puzzle. First, we assume that A, B, and C agree to answer our questions in English, and we show how to solve the puzzle in three questions. Next, we show how to solve the puzzle without this assumption. It turns out that most scenarios covered by our solution resolve the identities of the gods in two questions, but there is one case where information from a third question is necessary. Our final step is to show that there is always at least one stray case to scupper any two-question solution strategy.

We begin by observing that there are 6 state descriptions that correspond to the possible identities of the three gods.

(P1)	A-True	B-False	C-Random
(P2)	A-True	B-Random	C-False
(P3)	A-False	B-True	C-Random
(P4)	A-False	B-Random	C-True
(P5)	A-Random	B-True	C-False
(P6)	A-Random	B-False	C-True

The puzzle is solved when one possibility remains, revealing the true identities of all three gods. Here is the first of our three questions.

(Q1) Directed to god A: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?*

If B is Random, then god A must be either True or False, in which case A cannot answer and will remain silent. If A is Random, he can either answer or remain silent. If A is True and B is False, then A will answer 'no'. Finally, if A is False and B is True, then A will answer 'yes'. In more detail, we have the following possibilities.

If **A answers 'yes' to Q1**, then three possibilities remain: Either A is False, B is True, and C is Random, or A is Random.

(P3) A-False   B-True   C-Random
(P5) A-Random   B-True   C-False
(P6) A-Random   B-False   C-True

If **A answers 'no' to Q1**, then three possibilities remain: Either A is True, B is False, and C is Random, or A is Random.

(P1) A-True   B-False   C-Random
(P5) A-Random   B-True   C-False
(P6) A-Random   B-False   C-True

If **A gives no answer to Q1**, then four possibilities remain: Either A is Random or B is Random.

(P2) A-True   B-Random   C-False
(P4) A-False   B-Random   C-True
(P5) A-Random   B-True   C-False
(P6) A-Random   B-False   C-True

So, whatever the outcome of the first question, we may identify a god who is not Random. If A answers either 'yes' or 'no', then B is not Random, and if A remains silent, then C is not Random. Now, turn to the second question.

(Q2) Put to B or C we now know not to be Random: *Would you and the god not questioned thus far give the same answer to the question of whether Lisbon is south of Oxford?*

The possible answers to Q2 depend on how A answers Q1.

If **A answers 'yes' to Q1**, then, given the open possibilities in this case, we know that B is not Random. So, the second question put to B is this: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?* However B responds, we may determine the identify of each god since B answers 'yes' if and only if (P6) is actual, B answers 'no' if and only if (P5) is actual, and B is silent if and only if (P3) is actual.

If **A answers 'no' to Q1**, then, given the open possibilities in this case, we know that B is not Random. So, the same question is put to B, namely, *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?* Here again B answers 'yes' if and only if (P6) is actual, B answers 'no' if and only if (P5) is actual, and B is silent if and only if (P1) is actual.

If **A gives no answer to Q1**, then, given the open possibilities in this case, we know that C is not random. So, the second question is directed to C: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?* Here C answers 'yes' if and only if (P5) is actual and C answers 'no' if and only if (P6) is actual, but if C is silent then either (P2) is actual or (P4) is actual.

So, two questions suffice to solve the puzzle *unless* one fails to elicit an answer to both Q1 and Q2. To resolve the uncertainty between (P2) and (P4) in this case, a third question is required.

(Q3) Put to A: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?*

Since we know that B is Random, A answers 'yes' if and only if state (P4) is actual and A answers 'no' if and only if state (P2) is actual.

To turn this argument into a solution to Uzquiano's puzzle, where A, B, and C will only answer 'da' or 'ja' to our questions, if they answer at all, we will make use of two lemmas. First, let Q be some yes-no question and, in the spirit of Roberts [3], suppose that Q + is a question template of the following form,

Q+: *Would you answer 'ja' to Q?*

If the response 'ja' indicates affirmation in the gods' language, and the correct answer to Q is 'yes', then True will answer 'ja' to Q and thus 'ja' to Q+, whereas False will answer 'da' to Q and thus 'ja' to Q+. If instead the response 'ja' indicates denial in their language, then True will answer 'da'to Q and subsequently 'ja' to Q+, whereas False will answer 'ja' to Q and then 'ja' to Q+. Hence, both gods True and False will answer 'ja' to Q+, for a yes-question Q, regardless of whether 'ja' means 'yes' or 'no'. Adapting this reasoning for a no-question, and putting these observations together yields the embedded question lemma in Rabern and Rabern [2]: 108,

(EQL) When Q+ is posed to either True or False, each replies 'ja' if and only if the correct answer to Q is 'yes'.

This lemma is insufficient to discriminate True from False, however. For this, consider two new question templates, Q' and Q*.

Q': Would you answer a word meaning 'yes' in your language to whether Q?
Q*: Would you answer 'ja' to Q'?

If the response 'ja' indicates affirmation and the correct answer to Q is 'yes', then in reply to Q, Q' and Q*, True will always answer 'ja' whereas False will answer 'da' to Q, 'ja' to Q' but 'da' to Q*. Similarly, if the response 'ja' indicates affirmation and the correct answer to Q is 'no', then in reply to Q*, True will always answer 'da' to Q, Q' and Q* and False will answer 'ja' to Q, 'da' to Q' and again 'ja' to Q*. By an analogous line of reasoning, if instead the response 'ja' indicates denial and the correct answer to Q is 'no' ('yes'), then in reply to Q* True will answer 'ja' ('da') and False will answer 'da' ('ja'), respectively.

Putting these observations together yields a decoder lemma,

(DL) True and False use the responses 'ja' and 'da' to reply to Q* precisely as they each, according to their natures, would use 'yes' and 'no' to reply to Q, respectively.

To solve Uzquiano's puzzle, we replace Q in our question template Q' by Q1, Q2, and Q3, yielding the following three questions:

(Q1*)   Directed to god A: *Would you answer 'ja' to the question of whether you would answer with a word that means 'yes' in your language to the question of whether you and B would give the same answer to the question of whether Lisbon is south of Oxford?*

(Q2*)   Put to one of B or C we now know not to be Random: *Would you answer 'ja' to the question of whether you would answer with a word that means 'yes' in your language to the question of whether you and the god not*

addressed by (Q1) *give the same answer to the question of whether Lisbon is south of Oxford?*

(Q3*)   Put to A: *Would you answer 'ja' to the question of whether you would answer with a word that means 'yes' in your language to the question of whether you and C give the same answer to the question of whether Lisbon is south of Oxford?*

Finally, in discussing his puzzle, Uzquiano considers what would happen if both True and False have the oracular ability to predict Random's answers even before the coin lands in Random's brain. The wrinkle introduced by this possibility is that, even though Random remains random to our eyes, he is entirely predictable to his peers. Fortunately, Uzquiano's construction for this scenario can be applied to our solution, too. Following Uzquiano [4], we make use of the following observation:

(L): Would you answer 'ja' to the question whether you would answer 'da' to L?

Neither True nor False will answer L, since each is required to answer 'ja' if and only if his answer is 'da', and it is prohibited by their natures to answer undecidable questions. However, whether Random answers 'ja' or 'da', or whether Random answers at all, is entirely a random matter.

We can now ask a question that separates the states into the same clusters of our previous solution.

(Q1**) Put to A: *Would you answer 'ja' to the question whether B is Random and you would answer 'da' to Q1**?*

If B is not Random the first conjunct is false, which suffices to render Q1** false. Therefore, if the answer is 'ja' then A must be either False or Random; if the answer is 'da' then A is either True or Random; if the answer is silence, then A is Random. If B is Random, then A is either True or False and in both cases he will remain silent.

So when A is silent, then either A is Random or B is Random. But this yields exactly the 3 clusters we observed after question Q1, so a similar three-question strategy can be pursued.

If **A answers 'ja' to Q1****, then three possibilities remain. Either A is False, B is True, and C is Random, or A is Random.

(P3)   A-False   B-True   C-Random
(P5)   A-Random   B-True   C-False
(P6)   A-Random   B-False   C-True

If **A answers 'da' to Q1****, then three possibilities remain. Either A is True, B is False, and C is Random, or A is Random.

(P1)   A-True   B-False   C-Random
(P5)   A-Random   B-True   C-False
(P6)   A-Random   B-False   C-True

If **A gives no answer to Q1****, then four possibilities remain: Either A is Random or B is Random.

(P2)   A-True   B-Random   C-False
(P4)   A-False   B-Random   C-True

(P5)   A-Random   B-True   C-False
(P6)   A-Random   B-False   C-True

As before, whatever the outcome of the first question, we may identify a god who is not Random. If A answers either 'ja' or 'da', then B is not Random, and if A remains silent, then C is not Random. Now, the second question:

Q2** Put either to B or C whom we now know not to be Random: *Would you answer 'ja' to the question whether A is not Random and you would answer 'da' to Q2**?*

Finally, a third question is required only if we know that B is Random.

Q3** Put to A: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?*

Thus, the hardest logic puzzle ever can be solved in three questions. But, even though our solution requires three questions, how do we know that there isn't a clever two-question solution waiting in the wings? To resolve this worry, we appeal to a lemma from Information Theory:

(QL) If a question has N possible answers, these N answers cannot distinguish M > N different possibilities.

Initially, any question put to any god has three possible answers, 'ja', 'da', and no response, and there are six different possibilities to distinguish, (P1) through (P6). So, by (QL), it follows trivially that the puzzle cannot be solved in one question. To have a shot at solving the puzzle in two questions, it is necessary for there to be a first question that reduces the number of possibilities to no more than three. Our first question fails this condition precisely when there is no response, since this leaves four possibilities to consider. To see that this limitation is a feature of any solution and not merely our own, observe that whatever first question is posed to whichever god, say to A, we cannot exclude the possibility that A is Random, since Random may answer 'ja', 'da', or remain silent. But, if A were Random, there would be no information in his response that could be used to distinguish between states (P5) and (P6). So, for each of the three possible answers to the first question, states (P5) and (P6) will not be eliminated. So, the best we can do is to split the four remaining states by the three possible replies. Thus, there is always a scenario in which at least four possibilities remain after the first question. Because any question has three distinct answers, 'ja', 'da', and no answer, the scenario in which there are four possibilities cannot be distinguished by a single question. For this case a third question is necessary.[1] [2]

---

[1] More formally, (QL) says that an answer can decrease the initial entropy of an agent's knowledge about the identity of the gods by the maximum quantity of information an answer may convey. When there are four equally probable states, the entropy of the agent is $H = -4(1/4) \times \log2(1/4) = 2$ bits. The maximum information contained in a 3-valued answer is $I = -3(1/3) \times \log2(1/3) = 1.585$ bits, so the entropy can only decrease to $2–1.585 = 0.415$ bits. Thus, our agent is not guaranteed certain knowledge after two questions, because that result requires probability 1 or 0 to every state j, such that j's contribution to the whole entropy is $H_j = p_j \times \log2\left(p_j\right)$ is 0.

[2] Unbeknown to us, Landon Rabern also has given an impossibility argument at http://landon–rabern. blogspot.com/.

Finally, there is one loose thread to tie up. Our general strategy avoids posing direct questions about the identities of the gods, which we regard as a virtue, but our strategy for handling Uzquiano's oracular gods does not meet this standard: in Q1** and Q2** we ask whether some god or another is Random. To force this option off the table, suppose the gods all refuse to snitch on one another. According to their code of omertà, if a god is asked a direct question about his identity or the identity of another god, he remains silent rather than answer according to his nature. To make our general strategy truly general, we may jettison self-referential questions altogether and instead pose questions that exploit differences in each god's behavior over time.

(∀Q1**) Put to A: *Will B always answer differently from you to the question of whether Lisbon is south of Oxford?*

(∀Q2**) Put either to B or C, we now know not to be Random: *Would you and the god not asked so far always answer differently to the question of whether Lisbon is south of Oxford?*

Then, for the sake of clarity, we may simply recycle Q3**:

(∀Q3**) Put to A: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?*

Our solution to the hardest logic puzzle ever trades on identifying the pair of non-Random gods and exploiting their symmetric behavior in answering only decidable questions, regardless of their scruples about naming names. Alas, this suggests an even harder variation on the puzzle if we replace the god False with another named Devious:

Three gods, A, B, and C are called in some order, True, Random, and Devious. True always speaks truly, and whether Random speaks truly or falsely or whether Random speaks at all is a completely random manner. Devious always speaks falsely, if he is certain he can; but if he is unable to lie with certainty, he responds like Random. 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 in their own language, in which the words for 'yes' and 'no' are 'da' and 'ja', in some order. You don't know which word means which.

And here are our guidelines:

1. It could be that some god gets asked more than one question.
2. What the second question is, and to which god it is put, may depend on the answer to the first question.
3. Whether Random answers 'da' or 'ja' or whether Random answers at all should be thought of as depending on the toss of a fair three-sided dice hidden in his brain: if the dice comes down 1, he doesn't answer at all; if the dice comes down 2, he answers 'da'; if 3, 'ja'.
4. When Devious is able to lie he does so; but if Devious cannot be sure of telling a lie, then rather than remain silent, he responds randomly like Random, i.e., there is a fair

three-sided dice in Devious' brain that is tossed when he is not certain to lie. If the dice comes down 1, he doesn't answer at all; if the dice comes down 2, he answers 'da'; if 3, 'ja'.

Here is one possible solution. Initially, there are 6 possible scenarios, corresponding to the permutations of the three gods.

(S1)	A-True	B-Devious	C-Random
(S2)	A-True	B-Random	C-Devious
(S3)	A-Devious	B-True	C-Random
(S4)	A-Devious	B-Random	C-True
(S5)	A-Random	B-True	C-Devious
(S6)	A-Random	B-Devious	C-True

Our first question:

(QQ1) Directed to god A: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?*

Here are the possible answers A gives to QQ1 under each scenario:

(S1)	'no'
(S2)	silence
(S3)	'yes'
(S4)	'yes'/'no'/silence
(S5)	'yes'/'no'/silence
(S6)	'yes'/'no'/silence

We may rearrange this information into three columns, one for each of the possible responses, and list the set of states belonging to each group:

Answer: 'yes'	Answer: 'no'	Answer: silence
(S3) D T R	(S1) T D R	(S2) T R D
(S4) D R T	(S4) D R T	(S4) D R T
(S5) R T D	(S5) R T D	(S5) R T D
(S6) R D T	(S6) R D T	(S6) R D T

Now we show that the identities of the gods can always be determined in three questions by working through each column.

Case 1: 'yes'.

If **A answers 'yes' to QQ1**, then the second question is:

(QQ2-y) Directed to god C: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?*

Then the outcomes for each answer to QQ2-y are the following sets of possible states:

Answer: 'yes'	Answer: 'no'	Answer: silence
(S3) D T R	(S3) D T R	(S3) D T R
(S5) R T D	(S6) R D T	(S4) D R T

For each column, a third question is posed to separate the states:

(QQ3-yy) Column 1, Directed to god B: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?* Since B is True, there are only two possible responses he can give: 'no', in which case state (S5) holds; silence, in which case (S3) holds.

(QQ3-yn) Column 2, Directed to god B: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?* Since Devious must answer but True must remain silent, again there are only two possible responses B can give: 'no', in which case (S6) holds; silence, in which case (S3) holds.

(QQ3-ys) Column 3, Directed to god A: *Are you **certain** that you and B answer differently as to whether Lisbon is south of Oxford?* Since A is Devious and he must lie if he is able, he will provide a non-random answer: 'no', when (S3) holds; 'yes' when (S4) holds.

Case 2: 'no'.
If **A answers 'no' to QQ1**, then

(QQ2-n) Directed to god C: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?*

The outcomes for each answer to QQ2-n are separated by a similar procedure:

Answer: 'yes'	Answer: 'no'	Answer: silence
(S1) T D R	(S1) T D R	(S1) T D R
(S5) R T D	(S6) R D T	(S4) D R T

Then, just as before, for each column a third question is posed to resolve the identities of the gods.

(QQ3-ny) Column 1, Directed to god B: *Would you and A give the same answer to the question of whether Lisbon is south of Oxford?* Since Devious must answer but True must remain silent, there are only two possible responses B can give: 'yes', in which case (S1) holds; silence, in which case (S5) holds.

(QQ3-nn) Column 2, Directed to god B: *Are you **certain** that you and C answer differently as to whether Lisbon is south of Oxford?* Since B is Devious and he must lie if he is able, he will provide a non-random answer: 'no', when (S6) holds; 'yes' when (S1) holds.

(QQ3-ns) Column 3, Directed to god A: *Would you and C give the same answer to the question of whether Lisbon is south of Oxford?* Since Devious must answer but True must remain silent, again there are only two possible responses A can give: 'yes', in which case (S4) holds; silence, in which case (S1) holds.

Case 3: silence.
If **A is silent to QQ1**, then

(QQ2-s) Directed to god C: *Is Oxford south of Lisbon?*

Here we have less information to sort through.

For each column a third question is posed to separate the states,

Answer: 'yes'	Answer: 'no'	Answer: silence
(S2) T R D	(S4) D R T	–
(S5) R T D	(S6) R D T	–

(QQ3-sy) Column 1, Directed to god C: *Are you **certain** that you and B answer differently as to whether Lisbon is south of Oxford?* If C replies 'yes', then (S2) holds. If C replies 'no', (S5) holds.

(QQ3-sn) Column 2, Directed to god C: *Would you and B give the same answer to the question of whether Lisbon is south of Oxford?* If C replies 'no', then (S6) holds. If C is silent, (S4) holds.

Since three questions suffice to identify the gods in all possible cases, therefore we have a 3-question solution for the Devious puzzle. Moreover, the lemma (DL), when restricted to decidable questions, can be used to adapt our solution to the situation where the gods only speak in their mother tongue.

However, there is a catch. Our solution hinges on the gods, unlike us, being aware of who their neighbors are, but like us in not having the oracular ability to predict with certainty what Random or Devious will say beforehand whenever the chance device in each god's brain is set off. In other words, although our solution respects the gods' code of omertà, it nevertheless depends on the gods being like us in viewing the outcome of a 3-sided dice throw as a chance event. But if instead we assume that True and Devious can predict with certainty the outcome of any particular dice throw set off in some other god's head, then the question template[3]

(□Q) *Are you certain that* Q?

will fail to exploit a god's knowledge about which of his peers are disposed to answer some question stochastically or not, as our solution demands, but instead will simply track the god's confidence in his ability to predict his peer's answer to Q. To flatten this wrinkle, we may exploit a limitation to the gods' predictive powers. In place of QQ1, ask:

(QQ1t) Directed to god A: *for some future time t of my choosing, would you and B give the same answer to the question of whether Lisbon is south of Oxford if you were asked at t?*

Even though each god may predict with certainty the sequence of answers any of their peers would give to repeated queries of some yes-no question, QQ1t effectively blinds the gods by randomizing which point in the sequence of queries we will ask whether he and one of his neighbors would agree to our yes-no question. Since the gods cannot predict which point in time we will choose, QQ1t suffices (in the long run) to restore to Devious and Random their stochastic temperaments.

---

[3] Compare questions QQ3-ys, QQ3-nn and QQ3-sy.

Next, replace ($\Box$Q) by the question template

($\forall$Q) *Is it always the case that* Q*?*

Then, regardless of whether Devious views Random as stochastic or not, rather than ask 'Are you certain that you and god X answer differently as to whether Lisbon is south of Oxford?', ask 'Is it always the case that you and god X answer differently as to whether Lisbon is south of Oxford?'

Finally, flatly asking True whether he would answer the same as some other god to a yes-no question will fail to distinguish his addressing Random from his addressing Devious. Question QQ3-yy is an example. Here, rather than ask B 'Would you and C give the same answer to the question of whether Lisbon is south of Oxford?' ask instead 'Would you and C always answer differently as to whether Lisbon is south of Oxford?' Then, if QQ3-ny and QQ3sn are likewise changed to 'always answer differently' and QQ3-ns changed to 'always answer the same', then the revised question strategy yields a solution for when the gods have the oracular ability to predict with certainty how Random and Devious will answer a question.

In summary, a solution strategy based on the modal question template ($\Box$Q) is sensitive to whether the gods view their peers as stochastic or not, whereas the temporal question template ($\forall$Q) is not sensitive to this difference (in the long run). Both of these quantified question strategies offer an alternative that is at least as strong as strategies which appeal to self-referential statements. It is an open question whether the Devious puzzle can be solved without appealing to quantified questions, namely with only self-referential questions like (L). In any case, we are certain that the puzzle cannot be solved in less than three questions.[4]

References

1. Boolos, G. (1996). The hardest logic puzzle ever. *The Harvard Review of Philosophy, 6*, 62–65.
2. Rabern, B., & Rabern, L. (2008). A simple solution to the hardest logic puzzle ever. *Analysis, 68*, 105–112.
3. Roberts, T. S. (2001). Some thoughts about the hardest logic puzzle ever. *Journal of Philosophical Logic, 30*(4), 609–612.
4. Uzquiano, G. (2010). How to solve the hardest logic puzzle ever in two questions. *Analysis, 70*, 39–44.