[DRAFT] All Solutions to the Hardest Logic Puzzle Ever

Sander Beckers

1 Introduction

The Hardest Logic Puzzle ever is a puzzle introduced by Boolos (1996), which has attracted quite a lot of attention from philosophers. The puzzle is formulated as follows.

> Three gods A, R, 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, R, 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.

We here present a systematic solution to the puzzle in its most natural interpretation, i.e., the interpretation that we take to be the *intended* interpretation. Concretely, we take it that $Random$ truly answers randomly, and we require that all questions be truly yes-no questions. The latter means that in any given circumstance, for any god, there is always a correct answer to the question.

For anyone unfamiliar with the literature on this puzzle, it might seem as if our interpretation is simply a restatement of part of the puzzle. However some authors found aspects of the puzzle that are left somewhat ambiguous in Boolos' formulation, which has led the way to a range of variants on this puzzle that fill in this ambiguity in different ways. [note: give references (Rabern and Rabern, 2008; Uzquiano, 2010; Wheeler and Barahona, 2012), and say something about paradoxical questions, and the behaviour of Random.]

[note: (wikipedia) Rabern and Rabern (2008) suggest making an amendment to Boolos' original puzzle so that Random is actually random. The modification is to replace Boolos' third clarifying remark with the following:[5] "Whether Random says ja or da should be thought of as depending on the flip of a coin hidden in his brain: if the coin comes down heads, he says ja; if tails, he says da." This is also our interpretation of Random.]

Any solution to the puzzle consists of a strategy on which questions to ask, where the second and third question depend on the answers given to the previous questions. The fact that we will solve the puzzle systematically, implies that in some sense we will describe all possible solutions to it. In some sense, because there will still be a large degree of freedom as to exactly how one formulates the questions. More precisely, what we shall do for each of the three questions, is to characterize to which statement they have to be equivalent.

Name the three gods X, Y and Z; their descriptions T, F and R (short for $True$, $False$, and $Random$); and the three questions Q_1, Q_2 and Q_3. We pose Q_1 to X. There are six possibilities regarding the gods' identities, which we note as follows.

$$(X, Y, Z) \tag{1}$$
$$a = (R, F, T) \tag{2}$$
$$b = (R, T, F) \tag{3}$$
$$c = (T, F, R) \tag{4}$$
$$d = (F, T, R) \tag{5}$$
$$e = (T, R, F) \tag{6}$$
$$f = (F, R, T) \tag{7}$$

For each $i \in \{1, 2, 3\}$, let A_i denote the answer given to Q_i. Also, $\mathcal{A}_i$ denotes the answers given so far, i.e., $\mathcal{A}_i = \{A_j | j \leq i\}$. Lastly, let $S_{\mathcal{A}_i}$ denote the set of all remaining possible solutions after hearing the answer to Q_i. Further, we define $S_{\mathcal{A}_0}$ as the initial set of possibilities $\{a, b, c, d, e, f\}$, and s is the solution we are looking for, i.e., $s = S_{\mathcal{A}_3}$.

2 The First Question

We start with making the following obvious observation.

Observation 1 (Optimal Split)**.** *The best any binary question can do is to guarantee that the number of possible solutions is reduced by half.*

Since the puzzle requires that $\#S_{\mathcal{A}_3} = 1$, Observation 1 implies that we need $\#S_{\mathcal{A}_1} \leq 4$.

Next we make the following observation.

Observation 2 (Random Split)**.** *Given a god G and a question Q_i, for any solution $s \in S_{\mathcal{A}_{i-1}}$ such that $G = Random$, it holds that $s \in S_{\mathcal{A}_i}$.*

In other words, posing a question to $Random$ does not give you any information. Since $X = R$ in $\{a, b\}$, Observation 2 tells us that $\{a, b\} \subset S_{\mathcal{A}_1}$.

If $X \neq R$, then we are faced with one of the other 4 possibilities $\{c, d, e, f\}$. In this case, by Observation 1, the best we can aim for is to choose Q_1 so that only 2 possibilities of these 4 will remain, i.e., we cannot do better than $\#S_{\mathcal{A}_1} = 4$.

Combining these two conclusions, we get that $\#S_{\mathcal{A}_1} = 4$, and $\{a, b\} \subset S_{\mathcal{A}_1}$.

Applying Observation 1 to the fact that $\#S_{\mathcal{A}_3} = 1$ and $\#S_{\mathcal{A}_1} = 4$, we find that $\#S_{\mathcal{A}_2} = 2$.

Given Observation 2, if we pose Q_2 to $Random$, then there will be at least one solution that is compatible with both possible answers. But that violates the requirement that we get from $\#S_{\mathcal{A}_1} = 4$ to $\#S_{\mathcal{A}_2} = 2$. Therefore we have to make sure that we do not pose Q_2 (or Q_3) to $Random$.

Since we know that $\{a, b\} \subset S_{\mathcal{A}_1}$, we may not pose Q_2 to X. Hence Q_1 has to be such that it splits up the possibilities into $Y \neq R$ on the one side and $Z \neq R$ on the other.

Therefore Q_1 should be such that:

$$A_1 = da \Rightarrow S_{\mathcal{A}_1} = \{a, b, c, d\} . \tag{8}$$

$$A_1 = ja \Rightarrow S_{\mathcal{A}_1} = \{a, b, e, f\} . \tag{9}$$

Obviously the roles of da and ja can be reversed as well.

If $X = R$ then Q_1 does not matter, since the above will hold any way. So for the moment let us assume that $X \neq R$, giving $S_{\mathcal{A}_0} = \{c, d, e, f\}$ and:

$$A_1 = da \Rightarrow S_{\mathcal{A}_1} = \{c, d\} . \tag{10}$$

$$A_1 = ja \Rightarrow S_{\mathcal{A}_1} = \{e, f\} . \tag{11}$$

[note: the previous and following parts still need to be connected nicely.]

A tempting suggestion would be to simply stipulate to the gods that regardless of the actual meaning of “da” and “ja”, they should respond as if $“da'' = yes$ and $“ja'' = no$. This simple trick would resolve the language barrier. But of course that would be cheating, because the puzzle states that the gods answer in their own language, not in some artificial language that we have just invented. The good news is that there’s a way we can implement this trick without cheating.

Any yes-no question is semantically equivalent to a proposition followed by a question mark, so we can write Q_1 as $P_1?$, where P_1 is some proposition.

In order to implement our trick, we first consider a simpler setup. Imagine that we’re not posing our first question to one of the gods, but to an English-speaking, truthful person who knows the solution. In that case P_1 would simply be: $s \in S_1$, for some subset of solutions S_1. If the person answers “yes”, then $S_{\mathcal{A}_1} = S_1$, else $S_{\mathcal{A}_1} = S_{\mathcal{A}_0} \setminus S_1$.

Of course one can come up with many different questions that achieve this, but all of them are merely different formulations of what is semantically the same question: $P_1?$. The only freedom one has is in choosing an appropriate S_1.

For example, instead of asking $P_1?$, we could ask “Is P_1 true?”, or put differently, instead of asking about the proposition $s \in S_1$, we could ask if $\top \Leftrightarrow s \in S_1$ holds. Or we can reverse the roles played by “yes” and “no”, and ask instead if $\perp \Leftrightarrow s \in S_1$ holds.

We are free to replace $\top$ with any true proposition, for example we can fill in the fact that “yes” means yes: $“yes'' = yes \Leftrightarrow s \in S_1$. In this case, the answer “yes” implies $S_{\mathcal{A}_1} = S_1$ and “no” implies $S_{\mathcal{A}_1} = S_{\mathcal{A}_0} \setminus S_1$.

But we can also choose $“no'' = yes \Leftrightarrow s \in S_1$ and again reverse the roles of “yes” and “no”: now the answer “no” implies $S_{\mathcal{A}_1} = S_1$ and “yes” implies $S_{\mathcal{A}_1} = S_{\mathcal{A}_0} \setminus S_1$.

What these examples show, is that we can formulate our proposition P_1 in such a way that it does not matter whether “yes” means yes and “no” means “no”. In both cases, we have as P_1 a statement of the form $“X'' = yes \Leftrightarrow s \in S_1$. Regardless of the meaning of X, for this question it functions exactly as “yes” does for the statement $S_{\mathcal{A}_1} = S_1$.

Therefore we do not need to know the meaning of “da” and “ja”, since we can formulate our proposition in a manner so that “da” functions as yes no matter what: $da = yes \Leftrightarrow s \in S_1$.

Recall that the purpose of Q_1 is to find a god of whom we can be certain that he is not $Random$. Say we want to ask if $Y \neq Random$. This comes down to taking S_1 as $\{a, b, c, d\}$. That is, if we were asking our question to a truthful person, then that would be our choice of S_1. The trouble is of course that we have to take into account the possibility that $X = F$ or $X = R$.

If $X = R$, then we’re safe no matter what question we ask, since we will not pose our second question to X. Therefore we always meet our goal of posing our second question to a god who is not $Random$.

If $X = T$, then the solution is either c or e. In c we have that $Y \neq R$, whereas in e on the other hand $Y = R$. Therefore any choice of S_1 so that $c \in S_1$ and $e \notin S_1$ is fine. If the answer is da, then $Y \neq R$, and otherwise $Z \neq R$.

If $X = F$, then the solution is either d or f. In d we have that $Y \neq R$, whereas in f on the other hand $Y = R$. The fact that $X = F$ means that we should choose S_1 so that $f \in S_1$ and $d \notin S_1$. Again this ensures that answering da implies that $Y \neq R$, and else $Z \neq R$.

Combining the three possibilities for X, we get the following two conditions: $\{c, f\} \in S_1$, and $\{e, d\} \notin S_1$. So the easiest choice for P_1 would be:

"$da'' = yes \Leftrightarrow s \in \{c, f\}$.

Boolos makes the same choice in his article, be it that the roles of da and ja are reversed. Rabern & Rabern choose $S_1 = \{a, b, d, e\}$ and $S_2 = \{a, b, c, f\}$. [NOTE: I do intend to explain this in a longer version.]

3 The Second Question

On to the second question. We have found out after Q_1 that either $Y \neq R$, or that $Z \neq R$. Let's assume $Y \neq R$, the reasoning is completely analogous for $Z \neq R$. In other words, we assume that $A_1 = da$, so that $S_{\mathcal{A}_1} = \{a, b, c, d\}$.

As with the first question, only statements that are equivalent to a statement of the following form are allowed:

"$da'' = yes \Leftrightarrow s \in S_2$.

We have to reduce the number of possible solutions from 4 to 2. Hence we have to choose a set S_2 so that the answer da corresponds to two members of $\{a, b, c, d\}$, and ja corresponds to the remaining two.

This excludes trying to figure out if $X = T$, $X = F$, $Z = T$, or $Z = F$, since there is only member of $\{a, b, c, d\}$ that satisfies each of those. Of course we also have to exclude focussing on whether $Y = R$, since that would give us no information at all. As a result, $X \neq R$ is the same as $Z = R$, and $Z \neq R$ is the same as $X = R$.

That leaves us with $Z = R$, $X = R$, $Y = T$, or $Y = F$.

The first two of these are analogous to the situation we faced for the first question. For example, we can figure out if $Z = R$ by making sure that $\{a, d\} \subset S_2$ and $\{b, c\} \not\subset S_2$.

The last two options are somewhat different. We could apply the same procedure as before, but there's an easier solution. The fact that we're asking Y a question to figure out if he answers truthfully allows us to forget about S_2 and focus instead of using any statement that we know to be true. Ignoring the langauge barrier for a second, the easiest to way to figure out if Y always speaks the truth or always lies is to just ask him if $p = p$, or any other true statement. Combining this insight with our earlier trick, we get that the most straightforward strategy is to choose P_2 as:

"$da'' = yes \Leftrightarrow \top$.

If he answers da, then $S_{\mathcal{A}_2} = \{b, d\}$, else $S_{\mathcal{A}_2} = \{a, c\}$.

4 The Third Question

After hearing the second answer, we either know that $Y = T$, $Y = F$, $X = R$, or $Z = R$. For all of these we have that $\#S_{\mathcal{A}_2} = 2$, as required. In each case one can apply one of the strategies we discussed for the second question to find the unique solution after the third question.

For example, in case $X = R$, we can try to ask our third question to either Y or Z in order to see if he speaks truthfully. In case $Y = T$, we can pose our third question to Y as well, using the same structure as before and choosing a set S_3 so that either $b \in S_3$ and $d \notin S_3$, or the other way around.

References

Boolos G (1996) The hardest logic puzzle ever. The Harvard Review of Philosophy 6(1):62–65

Rabern B, Rabern L (2008) A simple solution to the hardest logic puzzle ever. Analysis 68(298):105–112

Uzquiano G (2010) How to solve the hardest logic puzzle in two questions. Analysis 70(1):39–44

Wheeler G, Barahona P (2012) Why the hardest logic puzzle ever cannot be solved in less than three questions. Journal of Philosophical Logic 41(2):493–503