Is there no solution to the blue-eyed islander puzzle?

Asked 8 years, 4 months ago Active 6 years, 5 months ago Viewed 13k times

38

28

Text below copied from here

> The Blue-Eyed Islander problem is one of my favorites. You can read about it here on Terry Tao's website, along with some discussion. I'll copy the problem here as well.
>
> There is an island upon which a tribe resides. The tribe consists of 1000 people, with various eye colours. Yet, their religion forbids them to know their own eye color, or even to discuss the topic; thus, each resident can (and does) see the eye colors of all other residents, but has no way of discovering his or her own (there are no reflective surfaces). If a tribesperson does discover his or her own eye color, then their religion compels them to commit ritual suicide at noon the following day in the village square for all to witness. All the tribespeople are highly logical and devout, and they all know that each other is also highly logical and devout (and they all know that they all know that each other is highly logical and devout, and so forth).
>
> [For the purposes of this logic puzzle, "highly logical" means that any conclusion that can logically deduced from the information and observations available to an islander, will automatically be known to that islander.]
>
> Of the 1000 islanders, it turns out that 100 of them have blue eyes and 900 of them have brown eyes, although the islanders are not initially aware of these statistics (each of them can of course only see 999 of the 1000 tribespeople).
>
> One day, a blue-eyed foreigner visits to the island and wins the complete trust of the tribe.
>
> One evening, he addresses the entire tribe to thank them for their hospitality.
>
> However, not knowing the customs, the foreigner makes the mistake of mentioning eye color in his address, remarking “how unusual it is to see another blue-eyed person like myself in this region of the world”.
>
> What effect, if anything, does this faux pas have on the tribe?

The possible options are

Argument 1. The foreigner has no effect, because his comments do not tell the tribe anything that they do not already know (everyone in the tribe can already see that there are several blue-eyed people in their tribe).

Argument 2. 100 days after the address, all the blue eyed people commit suicide. This is proven as a special case of

Proposition. Suppose that the tribe had n blue-eyed people for some positive integer n. Then n days after the traveller's address, all n blue-eyed people commit suicide.

Proof: We induct on n. When $n = 1$, the single blue-eyed person realizes that the traveler is referring to him or her, and thus commits suicide on the next day. Now suppose inductively that n is larger than 1. Each blue-eyed person will reason as follows: "If I am not blue-eyed, then there will only be $n - 1$ blue-eyed people on this island, and so they will all commit suicide $n - 1$ days after the traveler's address". But when $n - 1$ days pass, none of the blue-eyed people do so (because at that stage they have no evidence that they themselves are blue-eyed). After nobody commits suicide on the $(n - 1)^{st}$ day, each of the blue eyed people then realizes that they themselves must have blue eyes, and will then commit suicide on the n^{th} day.

It seems like no-one has found a suitable answer to this puzzle, which seems to be, "which argument is valid?"

My question is... Is there no solution to this puzzle?

logic induction puzzle

Share Cite Edit Follow Close Flag

edited Apr 13 '17 at 12:21
Community ♦
1

asked Nov 15 '12 at 22:55
picakhu
4,758 1 26 56

3 Why would it not have a solution? You stated the perfectly valid proof (that the 100 blue-eyed people kill themselves after 100 days) yourself. – Gregor Botero Nov 15 '12 at 23:02

@Gregor, the problem is... "what additional information did the outsider give the tribe?" – picakhu Nov 15 '12 at 23:04

3 The day after the blue-eyed people commit suicide, everyone else gathers in the square and commits suicide, leaving the stunned foreigner alone on the island. – Neal Nov 15 '12 at 23:06

It must be really depressing when you know your own eyecolor :D – Nikola Milinković Nov 15 '12 at 23:09

1 The new information is burrowed in the chain of thought inferences the people do (precisely at the end - the $n = 1$ case). The people on the island are quick and logical thinkers, they are *not* mind-readers. For sure, it does not have to be an outsider who speaks, it could as well be one of the persons on the island. That is why there is this law prohibiting discussing eye-colors. – Gregor Botero Nov 15 '12 at 23:28

In a way, the speech act ensures that everyone has the same information or at least can do the same deductions. – Gregor Botero Nov 15 '12 at 23:36

1 This problem is discussed on [this wikipedia page on common knowledge](). From what I understand, I believe this type of problem cannot be handled with our usual

logic, and requires the use of modal logic, but this out of my qualifications. – Vincent Nivoliers Nov 15 '12 at 23:37

8 As is common in these sorts of things, this isn't a logic problem, it's a modeling problem. As soon as you write down a specific logic to represent how the villagers reason, the outcome becomes indisputable. The "paradox" comes from the fact that there is no well-known model for this sort of reasoning. If you make the logic strong enough, the villagers will kill themselves even without the foreigner's comment. If you make it weak enough, they will not kill themselves even with the foreigner's comment. Thus it is entirely a question of which logic the villagers use to reason about knowledge. – Carl Mummert Nov 16 '12 at 13:16

Wouldn't everyone commit suicide then? If you don't know you have brown eyes, you could potentially rationalize just like those with blue eyes that you may have blue eyes and kill yourself just the same, thinking you know your own eye color. I suppose the factor here is whether or not the tribesmen can see each other prior to committing mass suicide. – Neil Feb 20 '13 at 14:36

I want to know is this puzzle can be solved logically? This is difficult question for me. If the answer is positive, I will try to solve this. – user79462 May 25 '13 at 20:14

The following interesting corollaries occurred to me when I first encountered this puzzle: 1. From the moment the foreigner makes his statement every islander knows the date of his or her death, but not the date of anyone else's. 2. Any sharing of this knowledge would be tantamount to informing the listener of his or her eye colour, and would therefore be taboo. – MartinG Oct 6 '14 at 22:09

So, doesn't that require that all people know the total number of eye colors on the island? If it doesn't then the +n can't determine if he should off himself on the +nth day. If this is true, then on the blue+1 day don't all the browns have to die? – nick Dec 16 '14 at 21:41

Both of these arguments are wrong, because they do not take into account the knowledge that all the islanders have of the minimum possible number of blue-eyed islanders. I have written up a general form answer to this puzzle on Puzzling.SE: puzzling.stackexchange.com/a/37673/20907 – Jed Schaaf Jul 14 '16 at 20:40

I'm reminded of the 'surprise exam' problem. A logic professor tells students that they are definitely receiving an exam during the next n future days, and when they do, it will be on a surprise day, not predictable in advance. Obviously n can't be zero, because then the exam can't exist - , if n=1 the exam has to take place on that day, but not by surprise. If n=2, then the professor has to avoid failing to have the exam on day 1, or else he has created the previous n=1 condition where the exam is not a surprise, so the exam can't be on day 2, but it can't be a surprise on day 1 – Cato Dec 26 '17 at 10:21

- this argument can be repeated inductively until n=100, but then there is no problem with the prof picking a day between 10 and 20 himself, and announcing the surprise exam on that day. So although n=1,2,3.. seem to have 'some' problem, it vanished by n=100. Similarly to me, in the blue eyes problem, few people have problems with n=1,2,3 - but some people feel a problem has 'emerged' at n=4 -does the original logic break down? Can it be inductively extended to n=100? With the blue eye problem already having the 'no information' paradox, it kind of looks even less likely to be 'logical' – Cato Dec 26 '17 at 10:22

6 Answers

Active | Oldest | Votes

46

Argument 1 is clearly wrong.

Consider the island with only *two* blue-eyed people. The foreigner arrives and announces "how unusual it is to see another blue-eyed person like myself in this region of the world." The induction argument is now simple, and proceeds for only two steps; on the second day both islanders commit suicide. (I leave this as a crucial exercise for the reader.)

Now, what did the foreigner tell the islanders that they did not already know? Say that the blue-eyed islanders are A and B. Each already knows that there are blue-eyed islanders, so this is *not* what they have learned from the foreigner. Each knows that there are blue-eyed islanders, but *neither* one knows that the other knows this. But when *A* hears the foreigner announce the existence of blue-eyed islanders, he gains new knowledge: he now knows *that B knows that there are blue-eyed islanders*. This is new; A did not know this before the announcement. The information learned by B is the same, but mutatis mutandis.

Analogously, in the case that there are three blue-eyed islanders, none learns from the foreigner that there are blue-eyed islanders; all three already knew this. And none learns from the foreigner that other islanders knew there were blue-eyed islanders; all three knew this as well. But each of the three does learn something new, namely that all the islanders now know that (all the islanders know that there are blue-eyed islanders). They did not know this before, and this new information makes the difference.

Apply this process 100 times and you will understand what new knowledge was gained by the hundred blue-eyed islanders in the puzzle.

Share Cite Edit Follow Flag

edited Sep 3 '14 at 14:25

answered Nov 16 '12 at 0:02

MJD

58.8k 34 246 451

4 Argument 2 claims "if there are n blue eyed islanders and everyone knows 'everyone knows there is a blue eyed islander' then they will commit suicide after n days". Accepting that as given, if there are at least 4 blue eyed islanders, they can all deduce 'everyone knows there is a blue eyed islander' already. This is because each of them knows that anyone who can see at least 2 blue eyed islanders has deduced 'everyone knows there is a blue eyed islander', and if there are 4 with blue eyes then each of them knows that everyone can see at least 2 blue eyed islanders. – Carl Mummert Nov 16 '12 at 13:35

1 In other words the arguments 1 and 2 are not in tension. 1 is correct: if there are 4 or more, every one of them already knows 'everyone knows there is someone with blue eyes' so the foreigner gave them no new information. But argument 2 does not argue that the foreigner *did* give them new information, it just argues about the consequences of them having that information. – Carl Mummert Nov 16 '12 at 13:41

8 No. If there are 4 blue-eyed islanders, then: 1. Every islander knows that there are blue-eyed islanders. 2. Every islander knows (1). 3. Every islander knows (2). 4. But *no* islander knows (3). And the foreigner's statement informs all the islanders of 3. To make this clearer, reconsider the case with 2 blue-eyed islanders and an unspecified number of brown-eyed islanders. Clearly (1) is true, but (2) is not. But after the foreigner's statement, (2) *is* true. – MJD Nov 16 '12 at 14:12

1 @Graphth: they might not know what the other ones know. For example, with three people, if Cathy has blue eyes then Alice knows that Bob sees Cathy has blue eyes, so Alice knows "Bob knows someone has blue eyes". But Alice does not know her own eye color, and Bob does not know his own eye color, so Alice does not know whether Bob knows "Cathy knows someone has blue eyes"; the only way Bob would know that Cathy sees blue eyes is if Alice has them, and Alice does not know her own eye color. – Carl Mummert Nov 16 '12 at 17:20

1 I have nothing else to say about this. – MJD Sep 10 '13 at 17:47

|

8

This isn't a solution to the puzzle, but it's too long to post as a comment. If one reads further in the post (second link), for clarification:

In response to a request for the solution shortly after the puzzle was posted, Terence Tao replied:

> I don't want to spoil the puzzle for others, but the key to resolving the apparent contradiction is to understand the concept of "common knowledge"; see http://en.wikipedia.org/wiki/Common_knowledge_%28logic%29

Added much later, Terence Tao poses *this question*:

> [An interesting moral dilemma: the traveler can save 99 lives after his faux pas, by naming a specific blue-eyed person as the one he referred to, causing that unlucky soul to commit suicide the next day and sparing everyone else. Would it be ethical to do so?]

Now that is truly a dilemma!

Added: See also this alternate version of same problem - and its solution, by Peter Winkler of Dartmouth (dedicated to Martin Gardner). See problem/solution (10).

Share Cite Edit Follow Flag

edited Nov 16 '12 at 16:43

answered Nov 15 '12 at 23:38

amWhy
200k 124 255 473

3 The traveler can say on closer inspection they are more a blue/violet color. The next day 1000 tribesmen show up to kill themselves, which gives them a chance to realize that they can't trust the traveler. Who has ideally left town by then. – psr Nov 16 '12 at 2:30

7

Easiest case to show what's wrong with the solution offered here and all other solutions like on wikipedia etc is to consider case of 4 blue eyed people.

Proposition 0: If there are 4 blue eyed people than everyone sees other 3 and also all of them can conclude(even without knowing their own eye color) that all of them know(not a typo) that there are at least 2 blue eyed people on the island. So if blue eyed people are A,B,C & D than A can say that B knows that C knows that D knows that A knows that there are at least 2 blue eyed, B can make similar assumption etc.

This would be enough for a pure mathematical proof, since if all of them know that all of them know that there are at least 2 blue eyed people than visitors anouncement that there is one blue eyed among them does not introduce new knowledge.

Let's find out what's wrong with the proof:

General argument for any kid of proof: If proof/solution assumes something that is impossible or contradicts to initial conditions than every conclusion based on such an assumption is useless.

All solutions I've seen so far include some variations of :

Suggestion A: Suppose/if there was only person (same as suppose n=1)
Suggestion B: Suppose/If there was only one blue eyed person

Clearly if there are 4 blue eyed person we can't suggest that there is only one person on the island so Suggestion A is clearly wrong and all proof based on such a suggestion are wrong. This also invalidates all recursive proofs that assume that n=1 / day one . One can abstract from the concrete example and say suppose n=1/day one, but than you can't imply the knowledge that on day one you knew that there is one blue eyed person.

Now what's wrong with the suggestion B? remember proposition 0. Everybody knows that all of them know that there are at least 2 blue eyed people on the island. But suggestion B says suppose there is only one blue eyed person. This Suggestion is also wrong since we know that there are at least two of them. So, all profs/conclusions based on suggestion B are also wrong.

That's it.

If you find a proof that does not use some variation of A or B than we can reopen this discussion.

Share Cite Edit Follow Flag

edited Feb 20 '13 at 14:11

answered Feb 14 '13 at 18:10

imachabeli
87 1 2

2 This is like that one problem where a professor tells his students that he will give a "surprise" exam in the next week. – Joe Z. Feb 14 '13 at 18:18

1 The foreigner doesn't introduce the literal information "there is someone with blue eyes here". If the foreigner told only one person that then that person would never kill themselves, for the reason you stated. It *has* to be made common knowledge. You've only gone up one level in the "knowledge hierarchy" with your four person example, which isn't what the foreigner does. Do they "know that the others know that they know"? How deep can their knowledge of *other people's* knowledge go? That's the crux of the argument, and you don't seem to address it here. – Robert Mastragostino Feb 17 '13 at 19:13

5 In case of four, knowledge goes to all hierarchical levels. I can see 3 blue eyed, I also can see that everybody sees at least 2 (my eye color does not count). So I can say that all of them know that all of them know(up to any hierarchical level if you wish) that there are at least 2 blue eyed people – imachabeli Feb 19 '13 at 22:45

2 With his statement, the foreigner tells them the following: "You will all start reasoning NOW'. This is something new, which they didn't know before. And the key to the "induction argument" is the fact that all can reason in the same time. – N. S. Dec 11 '13 at 22:30

1 The flaw to your argument is the following: You can see 3 blue eyed people, you can also see that everyone can see two blue eyed people. But what you cannot know, and that is what the visitor is introducing, is that everyone can reason now that 2 is not possible...... To put it simple, before the visitor came there is no way for you to know that everyone else now knows that you are thinking: "I can see 3 blue eyed, I also can see that everybody sees at least 2 (my eye color does not count)"... – N. S. Dec 11 '13 at 22:42

2

Here's my answer as to why the outsider gives new info. I'm considering the situation of (A)lice, (B)ob and (C)athy as mentioned in a post above, where all 3 have blue eyes. For argument's sake, I'll be Alice.

- I know there are at least two people with blue eyes
- I know Bob and Cathy know at least one person has blue eyes
- Here's the tricky part: What do I know about what Bob knows about Cathy's knowledge?

- As far I know, I may have brown eyes, and Bob may think he has brown eyes. So when Bob looks at Cathy, he may think she sees no one with blue eyes. So from Bob's point of view, Cathy may not know there are blue eyed people.

Simply put, I know there are blue eyed people. I know that Bob and Cathy know. But I don't know that Bob knows that Cathy knows. When the outsider announces it, I now know that everyone knows, and everyone knows that everyone knows.

It's also easier to imagine that everyone assumes they have brown eyes. So when Alice thinks about Bob's thoughts on Cathy's thoughts, everyone down the line assumes they themselves have brown eyes. When Alice thinks of Bob's thoughts, she's assumes he thinks they both have brown eyes, and only Cathy has blue. When Alice thinks of Bob's thoughts on Cathy's thoughts, she imagines Bob will conclude that Cathy will not see blue eyes.

This scales up to four or more, and it's easier (for me) to think of the fourth being 'above' A (Alice), let's say Omega. So Omega must imagine what Alice thinks that Bob is thinking about what Cathy thinks, not just what Cathy herself thinks.

So when Omega imagines Alice's thoughts, he's assuming she thinks they both have brown eyes. When he imagines Alice's thoughts on Bob's thoughts, same thing, everyone assumes they have brown eyes in their own mind. That's why Omega concludes that Alice may imagine Bob could think that Cathy may not know there are blue eyed people.

Very deep and difficult, but the 'everyone assumes they have brown eyes' POV helped me wrap my head around it.

Share Cite Edit Follow Flag

edited Oct 18 '13 at 17:33

answered Oct 17 '13 at 20:18

Shawn

21 2

1

The foreigner is only used as a point of reference for counting days. That is his role in the solution as it was thought. The puzzle merely says: look, they start their reasoning at this point in time. I don't agree with the solution though. If their goal was to find their eye color (such as in a different formulation of the puzzle), maybe this can be done by finding a solution which minimizes the amount of time that passes. This would allow them to agree without communicating. In this case, it is entirely possible to reach the conclusion that they all wait to see what happens in the first 99 days. If their goal is to not discover their eye color, then obviously, they will implicitly agree to do nothing and will stay alive for ever, never knowing their eye color.

Share Cite Edit Follow Flag

answered Apr 8 '14 at 8:41

Victor

11 1

-4

it looks like an administrator here deleted my answer. i just want to say that if you consider it important in any way to be any sort of ambassador to the layperson, then you need to seriously reconsider your attitude. theres a post on here that literally is written by a salesman, talking about the psychological reasons of why the 100 people would know. that one didnt get deleted. i wrote a clearly stated, reasoned response to the best of my ability, and it was snarkily ridiculed with no explanation, then deleted. terrible attitude dude. i am extremely disappointed.

if there is a problem with my logic that i am unaware of, then let me know. i am interested in discussion. i just dont think it is helpful to simply tell someone they havent answered the question without any explanation other than: its a math problem. it might be listed in the math section of this site, but as it is stated, it isnt a math problem. its a logic problem. the whole point of the problem is to determine what can and cannot be known with what information exists. if there is a flaw in my logic, please point it out.

i have re-answered the question, with a few clarifications. i expect you will delete it again mr. hardmath. but i just wanted you to know that if you do, it is very petty and uncool, especially considering you left up the answer from the salesman.

ANSWER:

the correct answer is option one: there would be no change. everyone already knows that everyone knows there are people with blue eyes. as long as there are at least 2 people amongst the islanders with blue eyes, everyone would see at least one other person with blue eyes, but not know if they themselves have blue eyes or not. in this case there are 100 people with blue eyes, and 900 without. but it doesnt matter how many people there are with blue eyes as long as there are more than 2 - they would have no way of knowing their own eye color, EVER. since there is no new information, why would anything change?

the problem that i see with the reasoning in the answers supporting option 2 is that it claims that the new knowledge is that the islanders now know that all of the other islanders know. this is not logical! all of the islanders must know that all the others know already! what one person sees, the others can see. they already understand that what they see the others can see, so therefore they already know that everyone already knows they already know. or whatever.

to take the above example of bob, cathy, and alice: all three must know that someone has blue eyes, and they all must know that they all know, because they all have blue eyes, therefore bob knows cathy knows and cathy knows alice knows bob knows, etc.

therefore, there is no new information. therefore, behavior does not change.

if someone disagrees with this, please explain. i have read through all of the answers and i really cannot understand why anyone thinks that the new information is that now all the islanders know that all the islanders know.

Share Cite Edit Follow Flag

edited Oct 6 '14 at 23:28

answered Oct 6 '14 at 22:43

noah

9 2

1 In your example of Alice, Bob and Cathy. Say that these three all have blue eyes. What does Alice know about what Bob knows? Alice knows that Bob knows that Cathy has Blue eyes, but Alice doesn't know that Bob knows that she (Alice) has blue eyes. The outsider saying that people have blue eyes now tells everyone nothing on day 1 or day 2. But by day 3, Alice now knows that not only does Bob know that Cathy has blue eyes, but also that Alice has blue eyes. The "not killing themselves" is the element that the outsider brings in, when people don't kill thenselves, they reveal information. – picakhu Oct 7 '14 at 13:46