

Halting Problem for Humans

Source: <http://sci.tech-archive.net/Archive/sci.logic/2006-10/msg00607.html>

- *From:* stevendaryl3016@xxxxxxxxxx (Daryl McCullough)
 - *Date:* 21 Oct 2006 11:22:23 -0700
-

The Halting problem is ultimately about the failure of prediction. No program can have perfect ability to predict the future behavior of all other programs. Here I'm going to illustrate the problems with perfect prediction using humans instead of programs.

Suppose that we have game in which two players, call them "Peter" and "Daryl" are asked yes/no questions. The rules stipulate that the only legal answers are "yes" and "no" (if neither answer seems appropriate, then the players must respond with silence). Assume that the players are intelligent and will attempt to answer the questions correctly, but most importantly they will never intentionally answer a question incorrectly—if they are not sure about the answer, then they should refrain from answering, rather than risk making an incorrect answer.

The rules are that both players get to see both questions. Afterwards, they are taken to separate answer rooms to figure out their answers. No interaction between the players is allowed. So there is no possibility that one player's answer can influence the answer given by the other. Each player must answer only using the knowledge that he brings into the answer room.

Okay, that's the background. Now here are the questions:

Peter is asked: Will Daryl answer "yes"?
Daryl is asked: Will Peter answer "no"?

I think it is clear that there is nothing paradoxical about the questions. Peter must use his knowledge about Daryl in order to answer the question, and Daryl must use his knowledge about Peter.

We can analyze the possibilities here (and presumably Peter and Daryl can do the analysis).

1. If Daryl and Peter give the same answer (either both "yes" or both "no") then Daryl is mistaken, and Peter is correct.
2. If they give opposite answers (one "yes" and one "no") then Peter is mistaken, and Daryl is correct.

Halting Problem for Humans

3. In light of 1&2, if both answer, then one of them is mistaken.

So, it is clearly impossible for both Daryl and Peter to have perfect ability to predict the future behavior of the other.

This formulation seems to be completely about the limitations of knowledge and reasoning ability. There is nothing paradoxical going on. And in particular, there doesn't seem to be the escape clause: Daryl (or Peter) knows the answer, but he can't say it. If he knows the answer, there is nothing at all preventing him from saying it. One player's answer has no causal effect on the other player.

This game is similar to the following one-player game: Someone asks Daryl to answer the following question with "yes" or "no" (and to make no other responses)

Will the next answer you give be "no"?

In this case, Daryl cannot correctly answer the question, but it is possible for him to *know* the answer (and keep it to himself). He can refuse to answer at all, in which case, he can know that the correct answer is "no".

In my two-player version, there is nothing preventing a player from saying the answer if he knows it.

--

Daryl McCullough
Ithaca, NY

.