A Reductio for Reliabilism

Brian Weatherson, Rutgers/Arché

July 17, 2009

In "A Reliabilist Solution to the Problem of Promiscuous Bootstrapping", Hilary Kornblith (2009) proposes a reliabilist solution to the bootstrapping problem. I'm going to argue that Kornblith's proposal, far from solving the bootstrapping problem, in fact makes the problem much harder for the reliabilist to solve. Indeed, I'm going to argue that Kornblith's considerations give us a way to develop a quick reductio of a certain kind of reliabilism.

Let's start with a crude statement of the problem. The bootstrapper, call them S, looks at a device D_1 that happens to be reliable, though at this stage S doesn't know this. We assume that S is a reliable reader of devices. S then draws the following conclusions.

(1a) D_1 says that p_1 at t_1 (since at t_1 D_1 appears to say p_1).
(2a) p_1 is true at t_1 (since D_1 says that p_1 at t_1).
(3a) D_1 is accurate at t_1 (by deductive inference from (1a) and (2a)).

Note that S need not, and in the story does not, know that the grounds for (1a) and (2a) are good grounds. S does, we'll suppose, know that (1a) and (2a) entail (3a). If reliabilism is true, then it seems S knows, or at least justifiably believes, that (1a), (2a) and (3a) are true, since each belief was produced by a reliable process. S then repeats the process a few more times, each time going through a version of the following triptych of inferences.

(1z) D_1 says that p_{26} at t_{26} (since at t_{26} D_1 appears to say p_{26}).
(2z) p_1 is true at t_{26} (since D_1 says that p_{26} at t_{26}).
(3z) D_1 is accurate at t_{26} (by deductive inference from (26a) and (2a)).

From (3a) through (3z), S infers (4).

(4) D_1 has been accurate the last 26 times I've used it.

And from (4), S infers (5).

(5) D_1 is generally accurate.

It isn't clear whether Kornblith thinks the problem for reliabilism is that it lets S infer (4) or that it lets S infer (5). As we'll see, the response he offers is a reason for S to not infer either (4) or (5).

In any case, it isn't clear that the inference from (4) to (5) is in any sense reliable. It's true that S has a lot of information that D_1 has worked well. On the other hand, if D_1 were disfunctional, S would not have that information. In fact, if D_1 had been unreliable at, say, t_8, then S would not have any evidence about how accurate D_1 was at t_8, since the relevant step, (2h) as it turns out, would fail. Since S's information harvesting technique can only produce evidence when D_1 works, and not when it doesn't work, it isn't clear that the fact that the evidence base includes only cases where D_1 works is evidence that D_1 is generally reliable. So I'll focus here on the worry that S shouldn't be able to infer (4) by using D_1 itself. (The points in this paragraph draw on some considerations raised by Jonathan Weisberg (forthcoming).)

Kornblith's point is that the process S uses to get to (4) is, overall, an unreliable process. For imagine that S simultaneously goes through the same reasoning with D_2, which is in fact unreliable.

(6a) D_2 says that q_1 at t_1 (since at t_1 D_2 appears to say q_1).
(7a) q_1 is true at t_1 (since D_2 says that p_1 at t_1).
(8a) D_2 is accurate at t_1 (by deductive inference from (1a) and (2a)).
...

(6z) D_2 says that q_{26} at t_{26} (since at t_{26} D_2 appears to say q_{26}).
(7z) p_1 is true at t_{26} (since D_2 says that q_{26} at t_{26}).
(8z) D_2 is accurate at t_{26} (by deductive inference from (26a) and (2a)).
(9) D_2 has been accurate the last 26 times I've used it (deductive inference from (8a)-(8z)).

Kornblith says that S uses the same process to get to (9) as to get to (4). And that process is clearly an unreliable process, since it doesn't distinguish between (4) and (9). So S isn't justified in believing (4), since the process that produced (4) is unreliable.

I don't think this is properly responsive to the problem. The worry was that S was using a reliable process to derive (4), and hence S's belief that (4) was unjustified, contrary to a strong intuition that it is not. Kornblith's response is that S is using an unreliable process to derive (4), and so S's belief is unjustified. But what the reliabilist needs is that S is not using a reliable process, not merely that S is using an unreliable process. The former claim does not follow from the latter, if S is using more than one process. And, worryingly for the reliabilist, that seems to be exactly what's going on here.

If I'm preparing beans and rice to eat, there is a process I run through to prepare the meal. That process has some subprocesses; there is at least one for the beans and one for the rice. Let's assume that I make the beans first, then the rice. Then when I finish preparing the rice, I finish two processes at once, the rice-preparing process and the meal-preparing process. There's nothing particularly special about this case; any particular action can be the conclusion of any number of processes. For similar reasons, any particular belief can be the termination of any number of cognitive processes. Some of these may be quite short processes, some of them longer processes.

That seems to be what's going on in S's case. When S reaches step 4, two processes terminate. One is the very long, and very unreliable, process of figuring out whether D_1 is reliable by comparing D_1's outputs with what we know about the world via D_1. Another is the very short, and very reliable, process, of drawing logical conclusions from what S has come to know through (2z).

Let's assume, with Kornblith, that the long process is really a process. And let's also assume the following two conditionals, (R1) and (R2), which are characteristic of a certain kind of reliabilism.

(R1) Any belief is justified if it is the outcome of a reliable process.
(R2) Any belief is unjustified if it is the outcome of an unreliable process.

I claim these assumptions lead to a contradiction. For S's belief in (4) is the outcome of two processes, the last two steps of which overlap, one of which is reliable, and the other of which is unreliable. So it is both the outcome of a reliable process, and the outcome of an unreliable process. That is no contradiction, any more than it is a contradiction to say that the belief is the outcome of both a long process and a short process. What is a contradiction is to say that the belief is both justified and unjustified. But that's just what follows from our earlier conclusion, plus the conditionals connecting reliability to justification. So I conclude that (R1) and (R2) can't both be true.

To be sure, that doesn't mean that reliabilism is doomed in all shapes, since there are plenty of reliabilist theories that don't endorse both (R1) and (R2). Alvin Goldman (forthcoming), for instance, says that justification requires reliability and evidential support. That implies that (R2) is true, but (R1) is false. Of course, any theory that rejects (R1) has no need for a fancy solution to the bootstrapping problem, since they are not required to say that (2a) was a legitimate step.

So I conclude that Kornblith's attempt to save reliabilism from the bootstrapping problem in fact leads strong versions of reliabilism, those that accept both (R1) and (R2), into contradiction. Weaker versions of reliabilism, including versions which qualify (R1), avoid the contradiction, but also have more direct means of avoiding the bootstrapping problem in the first place.

To finish, I want to briefly comment on how the problem I've discussed here relates to the well-known generality problem for reliabilism (Conee and Feldman 1998). The short version is that the two problems relate to different aspects of the evaluation of a belief.

The generality problem starts with the observation that token processes can't be reliable or unreliable, except in the trivial sense that they can yield a single true belief or a single false belief. Classes of processes can be reliable or unreliable, and some of them in interesting ways. The problem is that any process will be a member of some very reliable classes of processes, and some very unreliable classes of processes. To get a useful measurement of reliability for token processes, we need to match each process with a class of processes, and look at the frequency of success from the processes in that class.

Even if we succeed in that task, the puzzle I'm raising will still remain. The puzzle here isn't about matching token processes with classes of processes, but rather, about matching token beliefs with token processes. The bootstrapping problem suggests that a single belief may be the outcome of multiple processes. That suggests that any theory which accepts (R1) and (R2) will have to ensure, somehow, that those different processes 'match up' in terms of reliability. Perhaps some solution to the reference class problem will ensure that. But further reflection on the bootstrapping problem suggests that the prospects of such a solution are dim. Intuitively, the process that leads from (1a) to (4) should be grouped with the broad class of processes that involve testing a measuring device against its own outputs. And that is a very unreliable class. And the process that leads from (2z) to (4) should be grouped with the class of processes that involve deduction from known premises. And that is a very reliable class. So the natural solution, in this case, to the reference class problem, i.e., the natural mapping from token processes to the classes of processes that is relevant for their reliability, does nothing to solve the problem I'm raising.

References

Conee, Earl and Richard Feldman (1998) "The Generality Problem for Reliabilism" *Philosophical Studies* 89(1): 1-29.

Goldman, Alvin (forthcoming) "Toward a Synthesis of Reliabilism and Evidentialism? Or: Evidentialism's Problems, Reliabilism's Rescue Package," to appear in T. Doughterty, ed., *Evidentialism and Its Discontents*, Oxford University Press.

Kornblith, Hilary (2009) "A Reliabilist Solution to the Problem of Promiscuous Bootstrapping", *Analysis* 69: 2, 263-267.

Weisberg, Jonathan (forthcoming) "Bootstrapping in General", presented at the 2009 Rutgers Epistemology Conference.